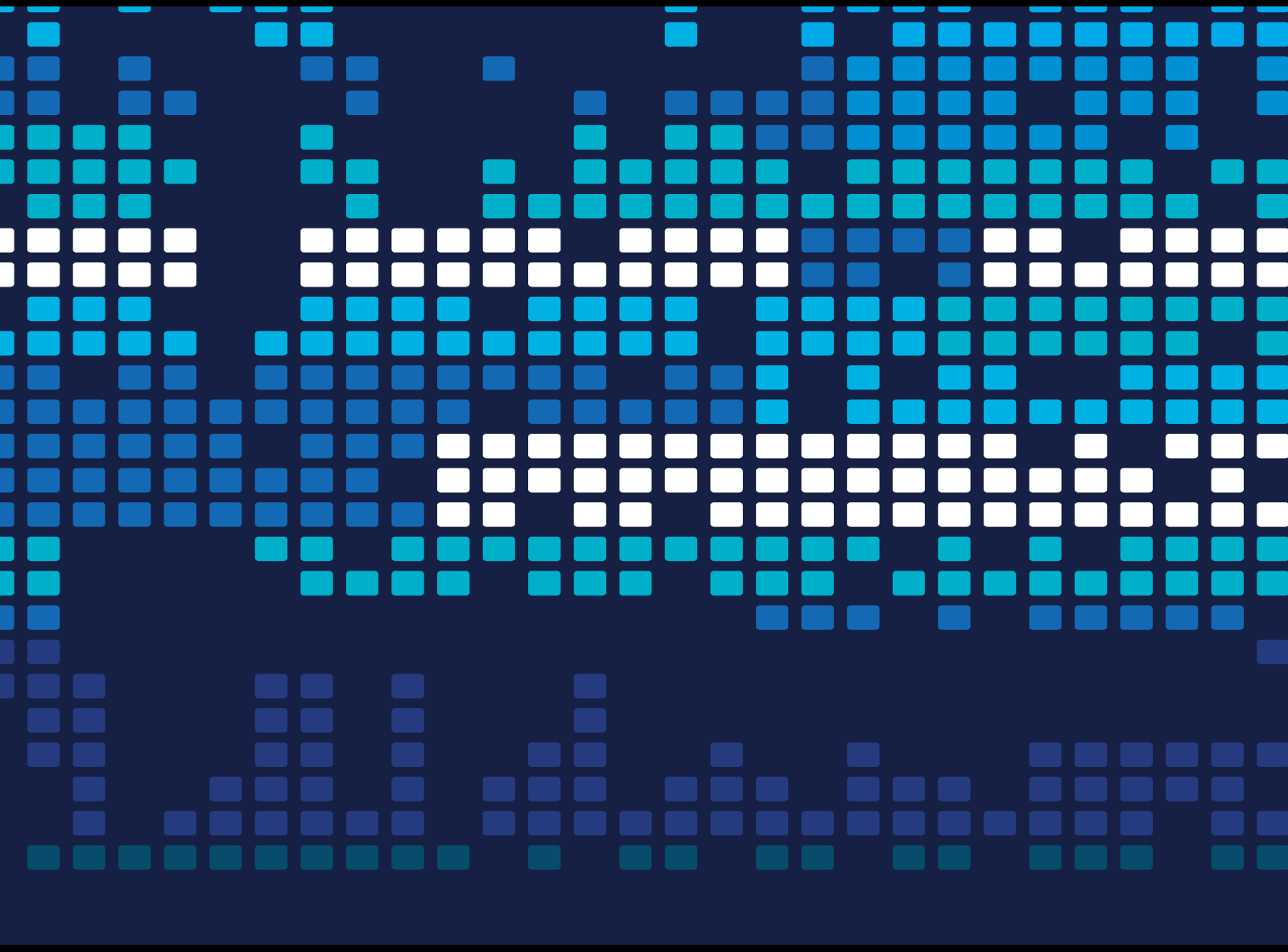
Scientific Programming

Managing Big Data, Visualization and its Analytics in Healthcare Based on Scientific Programming 2021

Lead Guest Editor: Shah Nazir
Guest Editors: Habib Ullah Khan, Iván García-Magariño, and Shaukat Ali



Managing Big Data, Visualization and its Analytics in Healthcare Based on Scientific Programming 2021

Scientific Programming

Managing Big Data, Visualization and its Analytics in Healthcare Based on Scientific Programming 2021

Lead Guest Editor: Shah Nazir
Guest Editors: Habib Ullah Khan, Iván García-Magariño, and Shaukat Ali

Contents

Contents

Contents

Hindawi
Scientific Programming
Volume 2023, Article ID 9827619, 1 page
https://doi.org/10.1155/2023/9827619

Retraction

Retracted: Analysis of Tai Chi Ideological and Political Course in University Based on Big Data and Graph Neural Networks

Scientific Programming

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

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1].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

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

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

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

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

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

References

[1] C. Y. Li and L. Zheng, "Analysis of Tai Chi Ideological and Political Course in University Based on Big Data and Graph Neural Networks," *Scientific Programming*, vol. 2021, Article ID 9914908, 9 pages, 2021.

Hindawi
Scientific Programming
Volume 2023, Article ID 9815683, 1 page
https://doi.org/10.1155/2023/9815683

Retraction

Retracted: Research on Deep Learning Algorithm in Cultural and Creative Product Design

Scientific Programming

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

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

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

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

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

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

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

References

[1] W. Liu, F. Kou, and H. Huang, "Research on Deep Learning Algorithm in Cultural and Creative Product Design," *Scientific Programming*, vol. 2021, Article ID 3212882, 15 pages, 2021.

Hindawi
Scientific Programming
Volume 2022, Article ID 7253832, 7 pages
https://doi.org/10.1155/2022/7253832

Research Article

Financial Risk Control Model Based on Deep Neural Networks

Jinghong Xu[1,2] and Daguang Yang[1]

[1]*Business School, Northeast Normal University, Changchun 130117, China*
[2]*Business School, Changchun Humanities and Sciences College, Changchun 130117, China*

Correspondence should be addressed to Daguang Yang; yangdg797@nenu.edu.cn

Received 26 April 2021; Revised 29 May 2021; Accepted 5 June 2021; Published 10 March 2022

Academic Editor: Shah Nazir

Under the background of global economic integration, commercial banks are facing more and more complex business environment. As one of the major financial risks faced by commercial banks, liquidity risk determines and reflects the safety and profitability of bank operation. Based on joint-stock commercial banks as the research object, this paper, respectively, from the angle of the static and dynamic measurements and projections for liquidity risk and based on the current situation of four joint-stock commercial banks liquidity level study, tries to explore the change law of commercial banks liquidity risk and financial risk control of commercial bank and puts forward reasonable suggestions. In this paper, an AHP neural network model combining subjective and objective methods is proposed. This method can not only overcome the defects of the single evaluation method but also improve the data accessibility by using the qualitative data and quantitative data of AHP. In the aspect of financial risk control system, this paper tries to establish a more comprehensive and practical financial risk control model by combining the previous research of scholars, the business model process, and the experience of practical workers.

1. Introduction

Although the overall liquidity of the banking industry has been in a state of excess for a long time, it does not mean that commercial banks do not have liquidity risks [1, 2]. The overall liquidity level of the banking system cannot completely represent the liquidity status of all individual banks, and there is no inevitable relationship between the two. In fact, commercial banks have many potential liquidity risks [3]. Specific manifestations are as follows:

(1) At the present stage, the good liquidity situation of commercial banks is not based on their own efficient risk management [4, 5] but is caused by external macro factors, which is highly unstable. With the full implementation of the opening up of the financial industry, the macro environment faced by banks will become complicated. Only by optimizing and improving the internal risk management level of banks can we fundamentally guarantee the sustainable and stable liquidity level of banks.

(2) China's state-owned banking system enables commercial banks to be protected by national credit all the time, so there is less possibility of bank runs. However, the excess liquidity situation brought by such institutional reasons will not last for a long time. With the continuous development of the financial market, the banking industry will gradually move towards the mode of independent management. In order to improve their comprehensive competitiveness, banks must attach great importance to the role of internal risk management and comprehensively improve the system of risk management.

(3) With the rapid development of the capital market, financial instruments continue to innovate, financial wealth management products are becoming more abundant, customers who originally belonged to banks will transfer funds to nonbank financial institutions, and the phenomenon of "financial disintermediation" has gradually become clear. According to this, the liquidity of the banking system

will be absorbed and dissolved in a large amount, and the stability of funds within the bank will also be challenged.

(4) With the continuous progress of financial technology [6–8], the credit rating system of commercial banks will be constantly updated and upgraded, and the ways of information transmission will become increasingly diversified and fast. Accordingly, the public will put forward higher requirements for the liquidity risk management level of commercial banks. Once the confidence crisis occurs, all negative information about the bank will be fully disclosed and widely spread in a very short time. The "domino effect" caused by this will expand the scope of influence of the crisis, and the bank will face significant risk of solvency crisis and run.

Based on the above observations, this article will combine the current status of commercial banks' liquidity and adopt a more advanced AHP neural network algorithm [9–12] to study the liquidity risks of joint-stock commercial banks. The research results are of great significance to the liquidity financial risk control of commercial banks. Following are the main innovations points of this paper:

(i) In this paper, the neural network based on the depth of the financial risk control model helps the bank's risk control department to improve the accuracy of risk prediction, so that they can early take corresponding measures to prevent risk and thus effectively reduce the probability of liquidity crisis, to reduce the potential risk of harm degree and the scope and eventually improve the overall operating performance of the bank.

(ii) In this paper, the subjective and objective methods are combined to evaluate financial risks, which can not only overcome the defects of a single evaluation method but also provide timely feedback when a single method deviates, and the AHP neural network model proposed in this paper can effectively predict financial risks.

2. Related Work

2.1. Commercial Bank Liquidity Risk Measurement. Effective measurement of liquidity risk is the core link and important premise of liquidity risk prediction [13]. However, there is no unified and comprehensive liquidity measurement index system in theory and practice, and the effective measurement of liquidity level is still a difficult problem worthy of in-depth discussion. To sum up, there are three main methods to measure liquidity risk at present: first, the static liquidity index measurement method; second, market signal system forecasting method; third, dynamic liquidity measures [14, 15].

The so-called static liquidity index measurement method mainly means that commercial banks use a series of economic indicators to measure the stock level of liquidity at a certain point in time. According to the basic characteristics of liquidity, liquidity can be measured from two aspects: assets and liabilities. Among them, the asset liquidity index is used to measure the degree of difficulty of the bank's asset realization. The so-called market signal system prediction [16, 17] method mainly analyzes bank liquidity from the external generation mechanism of risk. This method believes that the indicator system for measuring liquidity risk mainly includes changes in deposits, securities prices, premiums of debt instruments issued by banks, and asset realization. The dynamic measurement methods [18, 19] adopt frame analysis to include more factors affecting liquidity. The main methods include liquidity gap, net liquid assets, cash flow, and financing gap [20, 21].

2.2. AHP. AHP is a decision-making method that quantifies qualitative issues. It decomposes decision-making issues into goals, criteria, and plans. On this basis, qualitative analysis and quantitative analysis are carried out to hierarchize and quantify people's thinking processes. It also uses mathematics to provide a quantitative basis for analysis, decision-making, forecasting, or control. The simplicity and effectiveness of the analytic hierarchy process have made it quickly accepted by people. After decades of development, the current analytic hierarchy process has been widely used in energy system analysis, system evaluation, planning scheme selection, scientific research management [22], business management [23], policy analysis, regional planning [24], etc. Various applications of the AHP are present in literature [25].

2.3. BP Neural Network. BP neural network [26, 27] generally has a network structure of three layers or more. Adjacent neurons are fully connected, and each connection has a corresponding weight. Neurons in the same layer have no connection and no weight. The overall structure generally includes an input layer, an intermediate layer, and an output layer. The intermediate layer may have several layers, and the input layer and the output layer are both single layers. The structure of the BP neural network with a three-layer structure is shown in Figure 1.

In the same neural network model, in order to improve the computational accuracy of the neural network, the method of increasing the number of hidden layers or the number of neurons in the hidden layer is generally adopted. Increasing the number of hidden layers makes the model more complex, and the time needed to train the weight is greatly increased and the efficiency is low. Therefore, scholars basically do not adopt this method. In practice, regulating the number of neurons in the hidden layer is the method chosen by most people, but there is no clear way to determine the optimal number of neurons. An excessive number of neurons will lead to a longer training time, and the target of optimal precision may not be achieved. A too small number of neurons will lead to the weak training neural network, and the fault tolerance is poor. At present, there is no universal method to determine the number of neurons. The only way to determine the number of neurons

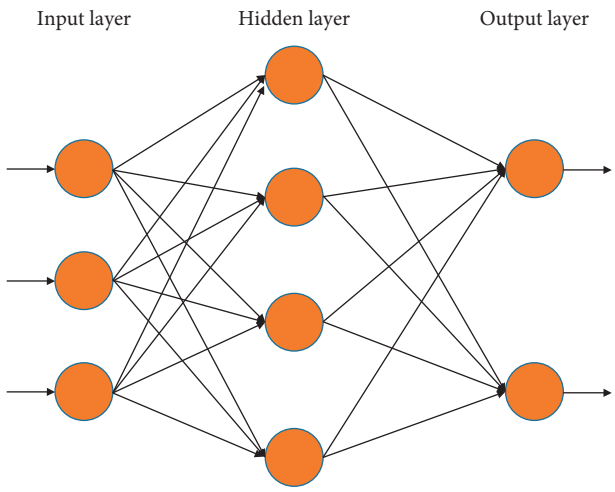


Figure 1: BP neural network structure.

is to train different numbers of neurons and then increase the allowance appropriately.

3. Methodology

3.1. AHP-NN. The training of neural network requires relatively high sample data. It is difficult to achieve full coverage of absolutely large amount of globally representative data, and it is difficult to ensure the accuracy, robustness, and stability of the training network in the training of simple neural network with small samples. In the real logistics financial risk assessment, many qualitative problems are difficult to be expressed with quantitative data, so it is difficult to grasp the accuracy of the samples that can be used for neural network neuron training. The samples obtained by AHP method and proved by practice first ensure the accuracy of neural network training. In addition, expert evaluation makes it possible to quantify qualitative data, which objectively guarantees the source of neural network sample data. Finally, the neural network based on AHP can learn the knowledge and experience of experts in a small sample to maximize the learning efficiency of the network.

The shortcomings of the AHP method are that subjective factors have a great influence on the evaluation results, the objectivity is poor, and the evaluation operations are more repetitive. In addition, the AHP evaluation method does not accumulate existing work experience; that is, historical data is difficult to have a positive impact on subsequent evaluations. The characteristic of the neural network is that the neural network stores information or knowledge in a large number of neurons or the entire system. It has the characteristics of holographic association, high-speed computing capability, strong adaptability, and self-learning and self-organizing capabilities. The self-learning ability and self-adaptability of the neural network method make the empirical accumulation of historical data retained in the network. As the sample data continues to expand, the accuracy of the model will become higher and higher.

The AHP-NN model is an evaluation model that combines the two methods of AHP and neural network. This model not only simply combines the two but also gives full play to their respective advantages and complements their shortcomings. Therefore, it has the advantages of subjective and objective analysis methods and can overcome their own shortcomings, and the evaluation model has good stability and accuracy.

3.2. AHP-NN Modeling

3.2.1. Positive Index. The bigger the indicator data, the better, such as sales growth rate and net profit growth rate. The standardized equation of the positive index is as follows:

$$x_{ij}^{*} = \frac{x_{ij} - x_{\min}}{x_{\max} - x_{\min}}, \tag{1}$$

where x_{ij}^{*} is the standardized data, x_{ij} is the original data, $x_{\max}$ is the maximum value in the data, and $x_{\min}$ is the minimum value in the data.

3.2.2. Negative Index. The smaller the indicator data, the better. The standardized calculation equation for negative indicators is as follows:

$$x_{ij}^{*} = \frac{x_{\max} - x_{ij}}{x_{\max} - x_{\min}}. \tag{2}$$

The basic steps of using analytic hierarchy process to evaluate financial risks are as follows:

(1) Establish the hierarchical structure of the system
(2) Perform pairwise comparisons of the elements at the same level to construct a pairwise judgment matrix
(3) Perform a consistency check on the calculated weights
(4) Calculate the composite weight, reorder the levels, and check the overall consistency
(5) Construct an evaluation model and calculate the evaluation results

3.3. Neural Networks Model. Neuron is the most basic unit of artificial neural network [28–31]. The neuron of the next layer receives the neuron information of the previous layer as input and obtains its output result after being stimulated by the excitation function, and the output result is used as the input of the neuron of the next layer. The input and output results of the neurons in the input layer of the neural network are the same, which means that the neurons do not do processing. The data of the neurons in the hidden layer and the output layer are stimulated by the transfer function. When the mean square error between the result of the output layer and the target result is lower than the set error value, it indicates that the training set learning has met the requirements. If the mean square error is greater than the set error value, weight correction

is required, and the information must be propagated backwards.

This paper chooses the hyperbolic tangent function as the excitation function. In the forward information process, the operation of the hidden layer and the output layer is as follows:

$$Y_{kj} = f\left(\sum_{i=1}^{n} W_{(k-1)i,kj} Y_{(k-1)i}\right),$$
$$f = \frac{e^{\lambda(\text{net}-\theta)} - e^{-\lambda(\text{net}-\theta)}}{e^{\lambda(\text{net}-\theta)} + e^{-\lambda(\text{net}-\theta)}}, \tag{3}$$

where net is the weighted sum of the input signal and θ is the offset.

When the root mean square error between the output result and the target value is greater than 0.1, the neural network training does not meet the standard, so the weight of the neural network needs to be corrected. At this time, reverse propagation is performed. Assuming the P-th sample, the input formula of the j-th node (neuron) is

$$\text{net}_{pj} = \sum_i W_{ji} O_{pj}. \tag{4}$$

The output formula of the P-th sample and the j-th node (neuron) is

$$O_{pj} = f\left(\text{net}_{pj}\right). \tag{5}$$

Then, the root mean square error of each input sample and target sample is

$$E = \sum_P E_P = \frac{\left(\sum_P \left(d_{pj} - O_{pj}\right)^2 p\right)}{2}, \tag{6}$$

where d_{pj} is the target value of the p -th sample output unit j.

The weight correction formula for neural network reverse transmission is

$$W_{ji} = W_{ji}(t) + \eta \sigma_{pj} O_{pj}, \tag{7}$$

where W_{jt} is the updated weight, $W_{jt}(t)$ is the original weight, η is the learning rate, and σ_{pj} is the input and output error. Among them, the error calculations of output layer neurons and hidden layer neurons are different. The calculation equation of the neurons in the output layer is as follows:

$$\sigma_{pj} = f'\left(\text{net}_{pj}\right)\left(d_{pj} - O_{pj}\right). \tag{8}$$

The error calculation formula of the hidden layer neuron is

$$\sigma_{pj} = f'\left(\text{net}_{pj}\right) \sum \sigma_{pk} W_{kj}. \tag{9}$$

Usually an inertia parameter α needs to be added to the weight correction formula. The new calculation equation is as follows:

$$W_{ji} = W_{ji}(t) + \eta \sigma_{pj} O_{pj} + \alpha\left(W_{ji}(t) - W_{ji}(t-1)\right), \tag{10}$$

where α is a constant, which represents the influence of the last weight update on this update.

4. Experiments and Results

4.1. Evaluation Methods. The quality of experimental results is generally verified by experimental results, so certain evaluation standards and criteria are needed to measure our experimental results. Even for the same model, different results may appear under different evaluation systems. This experiment selects several evaluation criteria that are frequently used in the financial industry.

4.1.1. Confusion Matrix. The confusion matrix is also called the error matrix, which is generally used to distinguish the quality of the classification model. The labels predicted by the model are represented by the columns of the matrix, and the actual labels of the samples are represented by rows. Take the binary classification problem as an example; it is a $2 * 2$ square matrix, as shown in Figure 2.

According to the true value of the sample and the prediction result of the model, four commonly used indicators can be obtained: true positive, false positive, false negative, and true negative. Commonly used predictors are precision, true positive rate, false positive rate, and specificity.

4.1.2. F-Measure. The precision index and the recall index sometimes have contradictions. The most common way to deal with this situation is to introduce the F-Measure that integrates these two indexes. F-Measure is the weighted harmonic average of precision index and recall index. Its calculation equation is as follows:

$$F = \frac{\left(b^2 + 1\right) P * R}{b^2 (P + R)}, \tag{11}$$

and when the parameter $b = 1$, it is the $F1$ score, and its calculation equation is as follows:

$$F = \frac{2 * P * R}{P + R}. \tag{12}$$

4.1.3. ROC and AUC. The ROC curve is plotted with the characteristic true positive rate (TPR) and false positive rate (FPR) as the vertical and horizontal axes, which shows a "game" between positive and negative examples. AUC is the area enclosed by the ROC curve. Its meaning is explained as follows. A sample A is randomly selected from all positive examples, and a sample B is randomly selected from all negative examples. The classifier judges A as a positive example with a higher probability than B. The probability of judging as a positive example is high.

4.2. Experimental Results. The AHP-NN neural network consists of an input layer, a hidden layer, and an output layer. The number of nodes in the input layer must be consistent with the dimension of the input data. We set the number of nodes in the input layer to 45, and the number of

True positive (TP)	False positive (FP)
False negative (FN)	True negative (TN)

Figure 2: Confusion matrix.

nodes in the output layer must be the same as the output layer. The categories of the results remain the same. The output results of this article have only two states: Pass and Reject, so the number of output nodes is 1, as shown in Table 1.

Regarding the setting of the number of hidden layers and the number of nodes, the method for determining the number of layers and nodes is more complicated. Currently, there is no general solution. Usually, systematic experimental methods are used to select the appropriate number of layers and nodes for a specific data set. This article adopts the method of selecting fewer layers or fewer nodes at the beginning, and then gradually increasing the complexity of the network structure, and the basic principle is to correctly reflect the relationship between output and input.

In the process of debugging the number of network layers, this article locks other hyperparameters: the number of hidden layer nodes is 20, the activation function uses Sigmoid, the error function uses the log loss function, the optimization function uses Adam, the number of batches is 100, and the number of iterations is set as 100. The number of network layers is set to contain one hidden layer (one), two hidden layers (two), three hidden layers (three), and four hidden layers (four). For the above in the four scenarios, the data in this article is used to test, and the loss curve on the test set is shown in Figure 3.

According to Figure 3, when the number of layers reaches 4, the loss value reaches a low value, and the loss value does not decrease significantly as the number of layers increases. After the model training is completed, the final loss and AUC values are shown in Table 2.

It can be seen from Table 2 that when the number of hidden layers is 4, both AUC and KS reach high values. Accordingly, the number of hidden layers selected in this paper is 4. In addition, the curve of AUC is shown in Figure 4.

To further verify the effectiveness of AHP-NN as a financial risk control model, the P-R diagram and ROC diagram are drawn as reference, as shown in Figures 5 and 6.

Table 1: Type of output node data.

Approval results	Type of output
Refuse	1
Agree	0

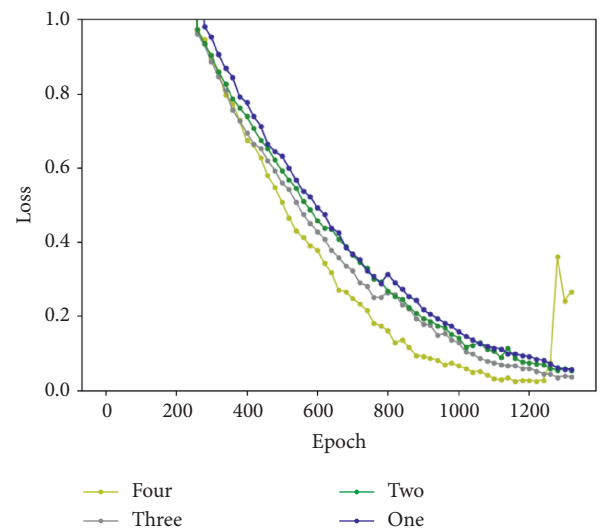


Figure 3: Loss curve graphs under different layers.

Table 2: LOSS and AUC values under different layers.

Layers	Loss	AUC
1	0.2695	0.8789
2	0.2069	0.8854
3	0.1896	0.9152
4	0.1752	0.9355

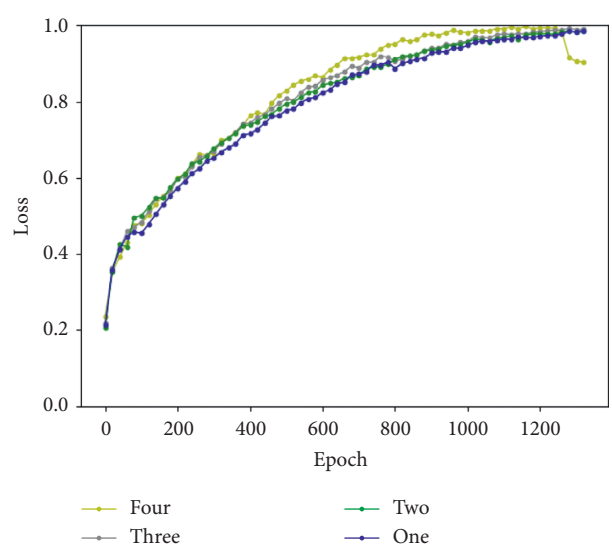


Figure 4: AUC curve graphs under different layers.

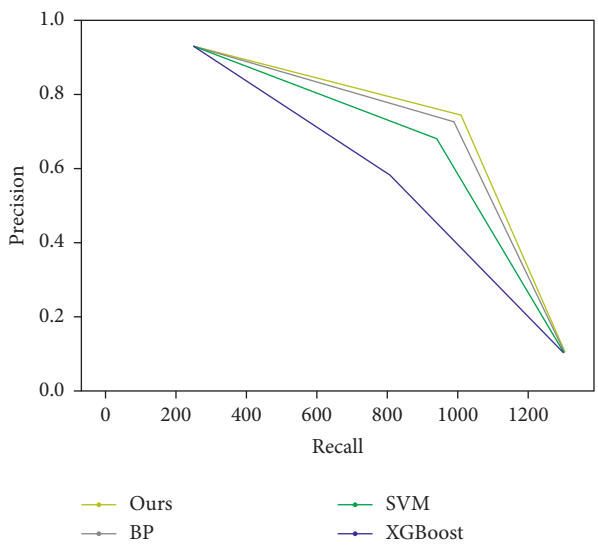


Figure 5: PR curve chart.

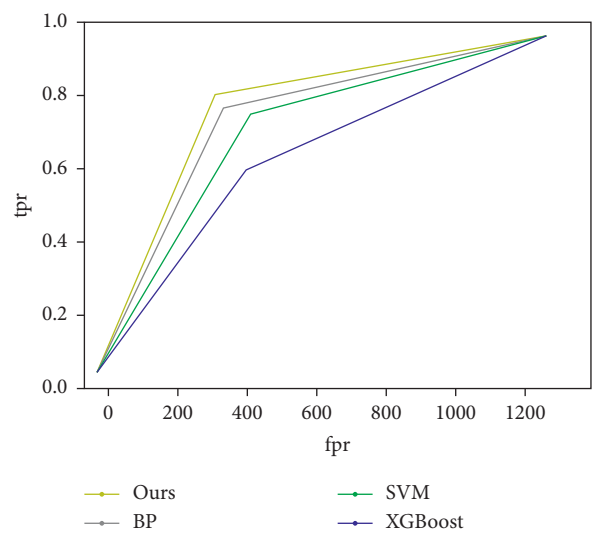


Figure 6: ROC curve graphs under different layers.

5. Conclusion

This paper takes joint-stock commercial banks as the research object, measures and predicts liquidity risks from static and dynamic perspectives, and conducts research based on the current level of liquidity of four joint-stock commercial banks in China, trying to explore the changes in commercial banks' liquidity risks laws, and then puts forward reasonable suggestions on the financial risk control of commercial banks. This paper proposes a neural network model of analytic hierarchy process that combines subjective and objective methods. This method can not only overcome the shortcomings of a single evaluation method but also use AHP's qualitative and quantitative data to significantly improve the availability of data. In terms of financial risk control system, this article combines the previous scholars' research, business model process, and the experience of practical workers to establish a more comprehensive and practical financial risk control model. Experimental results of the study shows the effectiveness of the proposed study under consideration.

Data Availability

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

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work was supported by the Key Project of Jilin Provincial Department of Education in 2020 "Study on Improving the Financing Efficiency of Small- and Medium-Sized Enterprises in Jilin Province" (JJKH20201298SK) and the Social Science Foundation of Jilin Province in 2020 "Research on the Fiscal and Financial Policies to Support the Survival and Development of Small-, Medium-, and Micro-Sized Enterprises in Jilin Province under the Normal Epidemic Condition" (2020C027).

References

[1] R. Aydemir and B. Guloglu, "How do banks determine their spreads under credit and liquidity risks during business cycles?" *Journal of International Financial Markets, Institutions and Money*, vol. 46, pp. 147–157, 2017.

[2] A. Abdul-Rahman, A. A. Sulaiman, and N. L. H. Mohd Said, "Does financing structure affects bank liquidity risk?" *Pacific-Basin Finance Journal*, vol. 52, pp. 26–39, 2018.

[3] M. Grill, C. Lambert, P. Marquardt, G. Watfe, and C. Weistroffer, "Counterparty and liquidity risks in exchange-traded funds," *Financial Stability Review*, vol. 2, 2018.

[4] A. A. Elamer, C. G. Ntim, and H. A. Abdou, "Islamic governance, national governance, and bank risk management and disclosure in MENA countries," *Business & Society*, vol. 59, no. 5, pp. 914–955, 2020.

[5] H. Valipour and M. Sohouli Vahed, "Risk management and forecasting macro-variables influences on bank risk," *International Journal of Business and Management*, vol. 12, no. 6, pp. 137–150, 2017.

[6] E. Abad-Segura, M.-D. González-Zamar, E. López-Meneses, and E. Vázquez-Cano, "Financial technology: review of trends, approaches and management," *Mathematics*, vol. 8, no. 6, p. 951, 2020.

[7] D. H. B. Phan, P. K. Narayan, R. E. Rahman, and A. R. Hutabarat, "Do financial technology firms influence bank performance?" *Pacific-Basin Finance Journal*, vol. 62, Article ID 101210, 2020.

[8] V. Sangwan, P. Prakash, and S. Singh, "Financial technology: a review of extant literature," *Studies in Economics and Finance*, vol. 37, pp. 71–88, 2019.

[9] Y. Tong, L. Yu, S. Li, J. Liu, H. Qin, and W. Li, "Polynomial fitting algorithm based on neural network," *ASP Transactions on Pattern Recognition and Intelligent Systems*, vol. 1, no. 1, pp. 32–39, 2021.

[10] Z. Chu, M. Hu, and X. Chen, "Robotic grasp detection using a novel two-stage approach," *ASP Transactions on Internet of Things*, vol. 1, no. 1, pp. 19–29, 2021.

[11] Z. Huang, P. Zhang, R. Liu, and D. Li, "Immature apple detection method based on improved Yolov3," *ASP Transactions on Internet of Things*, vol. 1, no. 1, pp. 9–13, 2021.

[12] X. Ning, Y. Wang, W. Tian, L. Liu, and W. Cai, "A biomimetic covering learning method based on principle of homology continuity," *ASP Transactions on Pattern Recognition and Intelligent Systems*, vol. 1, no. 1, pp. 9–16, 2021.

[13] D. Prodan-Palade, "Bankruptcy risk prediction models based on artificial neural networks," *The Audit Financiar Journal*, vol. 15, no. 147, pp. 110–121, 2017.

[14] H. P. Ramos and M. B. Righi, "Liquidity, implied volatility and tail risk: a comparison of liquidity measures," *International Review of Financial Analysis*, vol. 69, Article ID 101463, 2020.

[15] K. Chang, R. Chen, and J. Chevallier, "Market fragmentation, liquidity measures and improvement perspectives from China's emissions trading scheme pilots," *Energy Economics*, vol. 75, pp. 249–260, 2018.

[16] M. Ananthi and K. Vijayakumar, "Stock market analysis using candlestick regression and market trend prediction (CKRM)," *Journal of Ambient Intelligence and Humanized Computing*, vol. 12, pp. 1–8, 2020.

[17] I. Parmar, N. Agarwal, S. Saxena et al., "Stock market prediction using machine learning," in *Proceedings of the 2018 First International Conference on Secure Cyber Computing and Communication (ICSCCC)*, pp. 574–576, IEEE, Jalandhar, India, December 2018.

[18] S. Artese, V. Achilli, and R. Zinno, "Monitoring of bridges by a laser pointer: dynamic measurement of support rotations and elastic line displacements: methodology and first test," *Sensors*, vol. 18, no. 2, p. 338, 2018.

[19] M. Jiang, L. Jiang, D. Jiang et al., "Dynamic measurement errors prediction for sensors based on firefly algorithm optimize support vector machine," *Sustainable Cities and Society*, vol. 35, pp. 250–256, 2017.

[20] N. Yoshino, F. Taghizadeh–Hesary, and M. Nakahigashi, "Modelling the social funding and spill-over tax for addressing the green energy financing gap," *Economic Modelling*, vol. 77, pp. 34–41, 2019.

[21] D. Domeher, G. Musah, and N. Hassan, "Inter-sectoral differences in the SME financing gap: evidence from selected sectors in Ghana," *Journal of African Business*, vol. 18, no. 2, pp. 194–220, 2017.

[22] H. Dźwigoł and M. Dźwigoł-Barosz, "Scientific research methodology in management sciences," *Financial and Credit Activity: Problems of Theory and Practice*, vol. 2, no. 25, pp. 424–437, 2018.

[23] M. Radovic Markovic and A. Salamzadeh, "The importance of communication in business management," in *Proceedings of the 7th International Scientific Conference on Employment, Education and Entrepreneurship*, Belgrade, Serbia, October 2018.

[24] J. Harrison, D. Galland, and M. Tewdwr-Jones, "Regional planning is dead: long live planning regional futures," *Regional Studies*, vol. 55, no. 1, pp. 6–18, 2021.

[25] L. Wang, Y. Ali, S. Nazir, and M. Niazi, "ISA evaluation framework for security of internet of health things system using AHP-TOPSIS methods," *IEEE Access*, vol. 8, pp. 152316–152332, 2020.

[26] L. Huang, G. Xie, W. Zhao, Y. Gu, and Y. Huang, "Regional logistics demand forecasting: a BP neural network approach," *Complex and Intelligent Systems*, pp. 1–16, 2021, In press.

[27] G. Zhao, Y. Zhang, Y. Shi, H. Lan, and Q. Yang, "The application of bp neural networks to analysis the national vulnerability," *Computers, Materials and Continua*, vol. 58, no. 2, pp. 421–436, 2019.

[28] W. Cai, Z. Wei, R. Liu, Y. Zhuang, Y. Wang, and X. Ning, "Remote sensing image recognition based on multi-attention residual fusion networks," *ASP Transactions on Pattern Recognition and Intelligent Systems*, vol. 1, no. 1, pp. 1–8, 2021.

[29] X. Ning, K. Gong, W. Li, L. Zhang, X. Bai, and S. Tian, "Feature refinement and filter network for person Re-identification," *IEEE Transactions on Circuits and Systems for Video Technology*, 2020, In press.

[30] R. Liu, X. Ning, W. Cai, and G. Li, "Multiscale dense cross-attention mechanism with covariance pooling for hyperspectral image scene classification," *Mobile Information Systems*, vol. 2021, Article ID 9962057, 15 pages, 2021.

[31] W. Cai, B. Liu, Z. Wei, M. Li, and J. Kan, "TARDB-Net: triple-attention guided residual dense and BiLSTM networks for hyperspectral image classification," *Multimedia Tools and Applications*, vol. 80, no. 7, pp. 11291–11312, 2021.

Hindawi
Scientific Programming
Volume 2023, Article ID 9815683, 1 page
https://doi.org/10.1155/2023/9815683

Retraction

Retracted: Research on Deep Learning Algorithm in Cultural and Creative Product Design

Scientific Programming

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

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

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

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

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

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

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

References

[1] W. Liu, F. Kou, and H. Huang, "Research on Deep Learning Algorithm in Cultural and Creative Product Design," *Scientific Programming*, vol. 2021, Article ID 3212882, 15 pages, 2021.

Hindawi
Scientific Programming
Volume 2021, Article ID 3212882, 15 pages
https://doi.org/10.1155/2021/3212882

Research Article

Research on Deep Learning Algorithm in Cultural and Creative Product Design

Wenming Liu, Fusheng Kou, and Hengbo Huang

Shenyang Jianzhu University, Shenyang, Liaoning 110168, China

Correspondence should be addressed to Fusheng Kou; koufusheng@sjzu.edu.cn

Received 30 April 2021; Revised 13 August 2021; Accepted 26 August 2021; Published 27 September 2021

Academic Editor: Shah Nazir

Cultural and creative design is a new design mode based on the Internet platform, which gathers the design wisdom of the public and provides online solutions for design tasks. Evaluation indicators are used to score cultural and creative design schemes that require manual construction of an evaluation matrix, which makes the evaluation efficiency very low and increases the product design cycle. Therefore, based on the deep belief network of cultural creation, the method of constructing the evaluation model of the scheme is designed in this paper. According to the characteristics of the deep belief network, the k-medoids clustering method is adopted, and the clustering key value k is set. Meanwhile, the clustering result is calculated, and k central point elements are obtained. Moreover, project the central point elements back to the cultural and creative design plan to obtain the characteristic plan, and use manual evaluation to calculate the comprehensive score. Besides, the feature scheme data and the data simulation formula are used to construct the training set of the deep belief network, which are used to train the evaluation model of the deep belief network to solve the problem of the length of time that requires a large amount of manual evaluation to construct the output of the training set. In addition, according to the number of evaluation indicators, the number of nodes in the input layer of the deep belief network is determined to define the initialization method of network weights and biases, formulating experimental procedures to determine the optimal network structure. In the simulation experiment, through the test analysis of the comparison model, the function synchronization rate in this paper is 4.2% better than that of the comparison method. After multiple iterations, the test effect is 7 times higher than that of the comparison method, realizing the optimal evaluation of the design schemes that have not been manually evaluated in the same design scheme set.

1. Introduction

In recent years, the term cultural and creative products have repeatedly appeared in people's sight, and the new trend of fashionable life has been felt. Meanwhile, the creativity and design of cultural and creative products at domestic and foreign countries are constantly increasing. Therefore, the design of cultural and creative products is combined with relevant theoretical knowledge, design practice, research, and specific themes. In addition, aiming at how to make cultural and creative products better integrate with culture and the commercial market, ideas are proposed in this paper, hoping to attract others and give some inspiration to the design and development of cultural and creative products [1].

The experience value of cultural and creative products can not only meet consumers' material needs, but more importantly, they can meet their psychological needs. Moreover, cultural and creative products not only possess the general characteristics of ordinary commodities, but also possess cultural, regional, commemorative, practical, and contemporary characteristics [2]. Therefore, excellent cultural and creative products not only have the external form of artistic appreciation, but also can bring pleasant feelings to the audience through the appreciation of the inner spirit.

The research and design of cultural and creative products can connect culture and people. The old objects and life scenes are the source of design, and cultural customs are the core of design. Through fashionable and trendy cultural and creative products, interaction with young groups can be

achieved in modern society, promoting the dissemination and promotion of culture [3]. What is more, research and practice provide reference for the design and innovation of cultural and creative products, which have practical foundation and reference value. The regional cultural elements are rich in various forms, only part of which carries information related to the theme. Therefore, cultural and creative products will ultimately be materialized and visualized, and typical regional cultural elements with strong symbolic sense and clear forms should be first selected [4].

The choice of the design method and how to use it in the design of cultural and creative products are particularly important, which provides direction for how to reflect the cultural and artistic, regional, and national characteristics of cultural and creative products. Therefore, it is clearly recognized that product innovation is both the purpose and means of cultural and creative design, which is at the core of cultural and creative product activities. Meanwhile, it develops ideas for the design of cultural and creative products [5].

At present, the deep learning method has made great progress in data classification, fitting, regression, and prediction, especially in image recognition and reconstruction, speech recognition, and other fields [6]. In the process of data analysis using the deep learning model, it is mainly divided into three stages: data preprocessing, model structure training, and model testing and optimization, and each stage will affect the results of data analysis. In the aspect of deep learning network application, the test and optimization of model structure has become the main research field of many scholars, including proposing better adaptive network structure, such as Gaussian–Boltzmann machine, and proposing new model training methods, such as using the simulated annealing algorithm to optimize network structure parameters. These methods effectively improve the accuracy of model data analysis and provide a reference for subsequent research. Chen [7] divided product design into two target optimization problems of design and manufacturing through concurrent product and process design and proposed a satisfaction index to guide the selection of schemes with the help of the minimum and geometric average operators as the baseline. Besides, the comprehensive satisfaction of the two teams is calculated based on game theory. Vairaktarakis [8] adopted the quality house of product function (QFD) to consider the demand information of all customers and determined the combination of new products to meet the expected constraints and match or exceed the performance expectations of all customers in the target market.

For the cultural and creative design scheme, many designers participate in the design, which makes the data volume of the scheme large. At the same time, due to the consistent design goals, the data of cultural and creative design scheme also have the characteristics of strong data similarity [9]. The input of cultural and creative scheme data for deep learning network training is a part of the decision matrix of cultural and creative design scheme set, and the output is the comprehensive design result that must be calculated by the design team and combined with the weight of each index, which makes the manual work still very heavy.

In order to solve the problem of time-consuming for deep learning network to prepare training data set, the clustering method is used to analyze the data distribution law of cultural and creative design scheme, so as to reduce the labor time and reasonably construct training set to speed up the efficiency. This paper mainly studies the evaluation methods of the cultural and creative design schemes in the result output part of the cultural and creative design and studies the evaluation index selection and weight calculation methods of the cultural and creative design schemes. Moreover, the analytic hierarchy process is used to build a hierarchical structure for the cultural and creative program, and the evaluation index system is determined to obtain the vector representation of the cultural and creative design program. Besides, construct a judgment matrix, calculate the allocation matrix of index level and subcriterion level in turn, and check the consistency.

Structure of this paper is organized as follows:

(1) In Section 1, the research content, innovations, and organizational structure to be carried out are introduced in this paper.

(2) In Section 2, related work is introduced.

(3) In Section 3, the clustering model of the training set of the cultural and creative design scheme of the deep network is designed.

(4) In Section 4, the evaluation method of cultural and creative design plan is designed.

(5) In Section 5, experiments verify the research results proposed in this paper.

(6) Finally, summarize the full text and look forward to the future work.

The main innovations in this paper are as follows:

(1) The evaluation indicators are used to model the consistency of cultural and creative design schemes, which is conducive to data analysis of the deep belief network.

(2) According to the data characteristics of cultural and creative design plans, a deep belief network evaluation model is constructed, and the deviation standardization method is used to preprocess the data set of cultural and creative design plan.

(3) An evaluation method of cultural and creative design schemes using clustering is proposed, and the k-medoids clustering method is used to cluster the cultural and creative design plan data.

2. Related Work

In this paper, related technologies are introduced from the technical aspects of deep learning and restricted Boltzmann machines.

2.1. Deep Learning. With the vigorous development of the Internet big data, the research of deep learning has gradually become a mainstream trend and applied to various scenes in life [10].

The main network structures of deep learning models include auto-encoders, deep belief networks, and convolutional neural networks. Among them, the automatic encoder encodes and decodes the original data signal, which can be used for data noise reduction and data dimensionality reduction according to the number of hidden layers and the number of hidden layer nodes. Moreover, the deep belief network is composed of multiple restricted Boltzmann machines, which has a wide range of applications in the field of processing data classification. Since it also has the characteristics of fast convergence, it can also be used to initialize the structural parameters of the neural network to shorten training time. In addition, convolutional neural networks are often used in target detection and face recognition due to their huge advantages in image processing [11].

When deep belief network is used as the parameter initialization of neural network structure, it has the characteristics of fast training speed and not easy to fall into local extreme value. Therefore, the deep belief network is used in this paper to initialize the parameters of the evaluation network structure, and then the backpropagation algorithm is adopted to adjust the error of the model layer by layer. Finally, the scoring results of the cultural and creative design of the evaluation model are applied to making predictions [12].

2.2. Restricted Boltzmann Machine. The Boltzmann machine has a hidden layer and an input layer; the nodes between the layers are connected one by one, and the nodes in the layer are connected to each other. Due to the characteristics of the node connection in the layer, the time consumption in the training process is greatly increased, which brings difficulties to the use of the maximum likelihood method to train the network model structure [13]. Moreover, in this model, the connection between the nodes in the visible layer and the hidden layer is cancelled, and only the interlayer nodes are connected, which solves the problem of difficult training and is widely used. In addition, the restricted Boltzmann machine is a two-way undirected graph model, which consists of a visible layer v and a hidden layer h. What is more, the nodes between the visible layer and the hidden layer are connected one by one. The nodes in the layers are not connected, and the simplified structure of RBM is shown in Figure 1.

The restricted Boltzmann machine is a probability graph model that satisfies the Boltzmann distribution and the definition of the thermodynamic energy function. If a restricted Boltzmann machine has n visible layer nodes and m hidden layer nodes, the vector v can be used to represent the visible layer node, and the vector h can be used to represent the hidden layer node. For the restricted Boltzmann machine, the energy function $E(v, h|\theta)$ is defined as

$$E(v, h|\theta) = -\sum_{n}^{i=1}\sum_{m}^{j=1} w_{ij}v_i b_i b_i - \sum_{n}^{i=1} a_i v_i - \sum_{m}^{j=1} b_j h_j, \qquad (1)$$

where $\theta = \left\{w_{ij}, a_j, b_{ij} 1 \le i \le n, 1 \le j \le m\right\}$ is the parameter of the restricted Boltzmann machine and w_{ij} refers to the weight between the visible node i and the hidden node j. b_i is the bias of the visible node i, andb_j is the bias of the hidden node j [14].

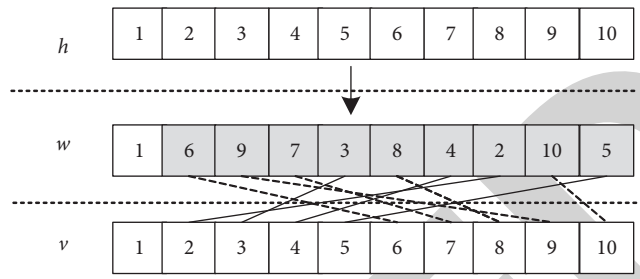


Figure 1: Structure of the restricted Boltzmann machine.

3. Clustering Model of Training Set of Cultural and Creative Design Schemes Based on Deep Network

According to the problems of sampling and reconstruction work in traditional cultural and creative design schemes, a training set clustering model of deep network is designed, and the model structure and training method are optimized.

3.1. Insufficiency of Cultural and Creative Design Schemes. The deep learning network model for cultural and creative design scheme can achieve a high degree of accuracy under the condition of sufficient training data preparation and can improve the work of cultural and creative design scheme, but this method also faces the following two problems:

(1) The training of deep learning network can be divided into two steps: pretraining and optimization [8]. For the pretraining, the training of deep learning network model requires t cultural and creative design schemes $P_i(i = 1, 2, \ldots, t)$ as the input v of deep learning network. The contrast divergence method is used for data sampling and reconstruction. Finally, the training results of each RBM and the parameter values of each layer are obtained. The overall structure of deep learning design scheme is shown in Figure 2.

In the process of pretraining, the data are obtained after the unified index screening and the calculation of index weights at all levels, which takes less time and is relatively easy to obtain. In the optimization stage of network structure parameters, it is necessary to use the design result K of each cultural and creative design scheme to input back from the deep learning network and use the backpropagation algorithm of neural network training to adjust the error of the whole network structure parameters, so as to optimize the data analysis ability of the model [15]. In the optimization stage, the data used must be screened based on indicators to get the results of each cultural and creative design scheme. The time consumption depends on the number of cultural and

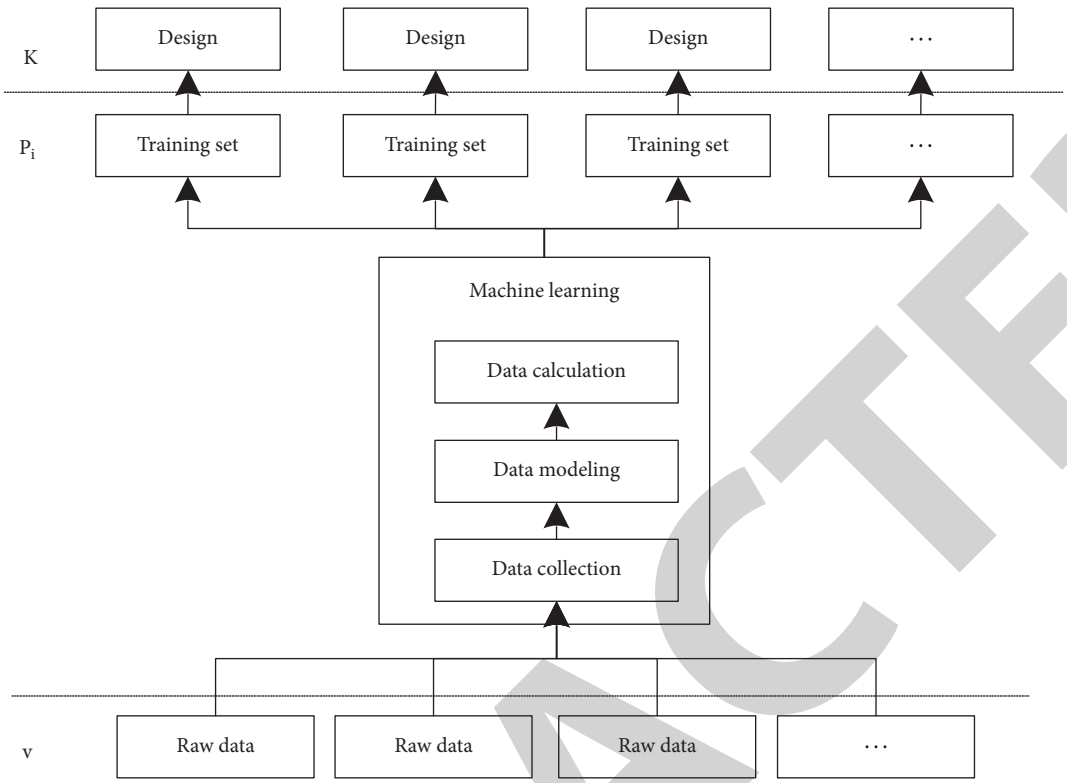


Figure 2: Overall structure of the deep learning design scheme.

creative design schemes used for training and the time taken for a single cultural and creative design scheme to be screened by the design team. It is relatively difficult to obtain the data, which greatly reduces the efficiency of the whole model [16].

(2) Randomness of training data selection: in the process of adopting the deep learning network cultural and creative design scheme, the cultural and creative design scheme selected as the training data set is randomly selected, without considering the distribution law and similarity degree of the cultural and creative design scheme data, which makes the training of the model have a strong contingency; for the cultural and creative design scheme with high similarity, it has the feature of repeated manual. Moreover, it will reduce the model efficiency, resulting in great errors in the data of cultural and creative design schemes with large differences [17].

Therefore, it is necessary to improve the selection process of training data of deep learning network, so as to improve the accuracy of the model and the efficiency of training data preparation, and solve the problem of long time.

3.2. The Clustering Model of the Training Set of Cultural and Creative Design Schemes. For a large number of data, clustering is a data mining method to obtain data distribution rules. For the cultural and creative design scheme under the same cultural and creative design, due to the same design objectives and the same defined index system, the similarity between the data is more likely [18]. This paper analyzes the cultural and creative design schemes in a unified dimension by using the clustering method and divides them into H clusters; then it takes the central data of each cluster as the feature scheme to obtain the criteria, and the structure of the clustering scheme is shown in Figure 3.

Although the method of feature scheme improves the efficiency, it reduces the number of training concentration schemes of the deep learning model, which is not conducive to the training and testing of the model. In order to train the model with accurate results, the training set should be expanded. Because of the strong randomness of the design scheme, the normal distribution random number is used to fit the original data fluctuation space. The training set of cultural and creative design scheme for training and the test set of cultural and creative design scheme for testing are reasonably constructed as the data input of the deep learning model [19].

The following three steps can be used to cluster the training set of cultural and creative design schemes:

(1) Cluster analysis of cultural and creative design schemes: for all cultural innovation design scheme $P_i\,(i = 1, 2, \ldots, t)$, if its index number is k, it is projected into the k-dimensional space to obtain a set of data point clouds of cultural innovation design scheme. Set a reasonable cluster number H, use the K-means clustering algorithm to perform clustering calculation on the data of cultural and creative clustering design scheme, and get the center of each cluster [20].

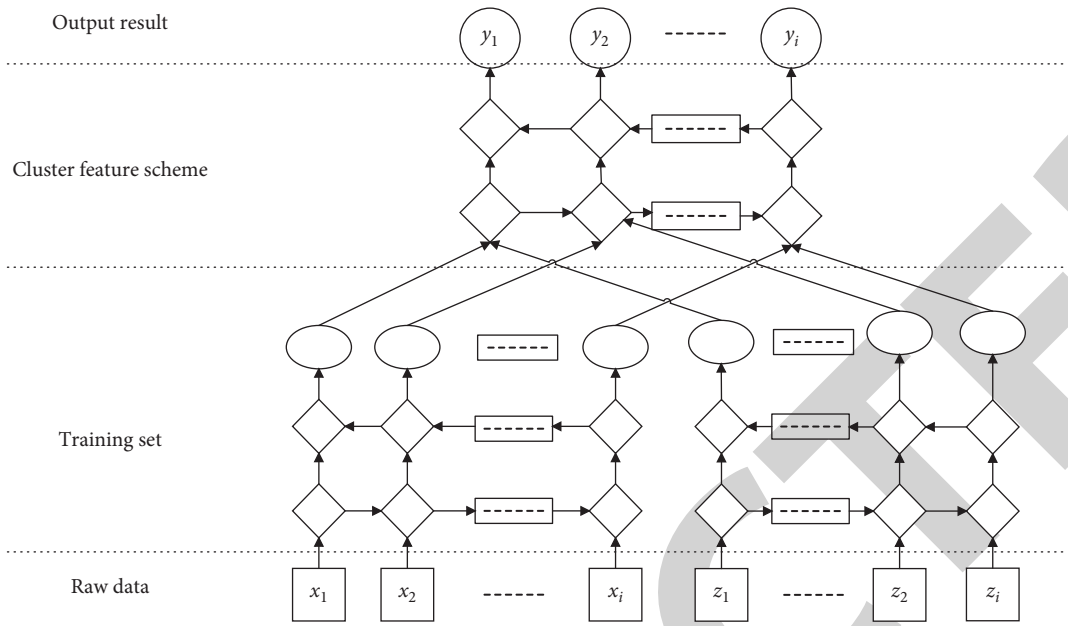


Figure 3: Clustering scheme based on the deep learning network model.

(2) Feature scheme and construct training set: take the center of the cluster group, calculate with the European distance calculation equation, and construct the preliminary set Pre by using the distance of all schemes within the group and the minimum scheme. The preliminary set is used as the sample for technical guidance, the matrix is obtained, and the comprehensive score of the feature scheme is calculated based on the weight distribution results. The value characteristics of each index were analyzed, and the training set of deep learning network was constructed by using the data simulation equation [21].

(3) Training and optimization of the deep learning network model: the deep learning network model is used as the model, the constructed training set is used to train the deep learning network model, and then the original cultural and creative design scheme data are used as the test set to test the model and verify the training effect of the simulation training set on the model [22].

3.3. Cultural and Creative Design Plan Data Set Preprocessing. For any cultural and creative design scheme $P_t = (\alpha_{i1}, \alpha_{i2}, \ldots, \alpha_{ik})$ taken from the set of cultural and creative design schemes $S = \{P_1, P_2, \ldots P_i \ldots P_N\}$, it is assumed that after the evaluation of the scheme based on the evaluation index, the score obtained is y_t. Meanwhile, if t cultural and creative design schemes $p_t\,(t = 1, 2, \ldots, n)$ are used as the input V of the deep belief network, and the scoring resulty_t is used as the output of the deep belief network, the network model will be trained to obtain the evaluation rule of the scheme, and the trained model can be used to evaluate the unevaluated cultural and creative design plan $p_i\,(i = t + 1, t + 2, \ldots, t + n)$, so that the scoring result y_i^o of the deep belief network model can be obtained. As long as y_i^o meets the scoring error requirements, the evaluation result can be used to replace the manual evaluation result, reduce the evaluation work of cultural and creative design schemes, and improve the evaluation efficiency [23].

Due to the limitation of the activation function sigmoid in the restricted Boltzmann machine, the value range of each node of the input layer v is [0, 1]. In addition, to reduce the computational cost of the RBM training reconstruction process and achieve better training results, it is necessary to perform normalization processing before inputting the cultural and creative design plan data into the model, so that the model algorithm can run better and achieve the desired effect [24].

Before normalizing the data of the cultural and creative design plan, it is necessary to consider whether to reduce the dimensionality of the cultural and creative design plan data to remove the secondary features and select the main features to make the evaluation results of the model better. Nowadays, data dimensionality reduction is mainly selected by manual feature extraction and principal component analysis dimensionality reduction [25]. Moreover, for data whose dimensionality fc is more than hundreds or thousands, using principal component analysis is a very effective method to reduce dimensionality. Besides, for the cultural and creative design data, since the dimension of the design data is controlled by the number of evaluation indicators, and the primary and secondary relationship of the features has been considered in the selection process of the evaluation indicators, the dimensionality reduction and non-dimensionality reduction can be selected according to the specific situation [26].

Nowadays, the normalization method for mapping data to the interval of [0, 1] is the min-max standardization method, which has also become the dispersion

standardization method. This method uses the maximum and minimum values of the data under the unified index to linearly transform all the data and map them to the target interval, which can also be scaled according to the needs of the interval [27]. If the evaluation index parameter α_{ij} of any cultural and creative design plan p_i, whose maximum value in the sample data set is $\alpha i_{j\max}$, the minimum value is $\alpha_{ij\min}$, and the standardized evaluation index parameter is α'_{ij}, using the deviation standardization method, the calculation method of a'_{ij} will be

$$a'_{ij} = \frac{a_{ij} - a_{ij\min}}{\alpha_{ij} + na_{ij}}. \quad (2)$$

For all evaluation indicators α_{ij}, after normalized with the scoring results, if $v_i = \alpha'_{ij}$, the visual layer input $v = (v_1, v_2, \ldots, v_k)$ and output y'_i of the deep belief network can be obtained.

3.4. Evaluation Model Network Structure and Parameter Selection. The construction of a deep neural network mainly includes the following: the number of network layers and nodes of each layer, the initial structure parameter value of the network, the learning rate and momentum of weight updating, the number of iterations of reconstruction, and BP [28].

For the deep neural network model used for the evaluation of cultural and creative design schemes, the number of visual layer nodes v is equal to the number of evaluation indexes, and the value is K. The number of nodes in the output layer is equal to the number of design results, with the value of 1. According to the characteristics of deep neural network, the number of hidden layers is small, and the effect of feature extraction is not different from the function of the neural network. Too many hidden layers will lead to a large deviation between the reconstructed data of the later layers and the original data, resulting in data distortion. For the value of nodes of each hidden layer, there are three methods: decreasing layer by layer, increasing layer by layer, increasing then decreasing first, and decreasing then increasing first. Since there is only one design result of the cultural innovation design scheme and the number of output nodes is small, the value method of decreasing then increasing and increasing layer by layer cannot be adopted. Therefore, the layers of the deep neural network model need to be analyzed through experiments to determine the optimal choice [29].

The initial value of network structure parameters will have an influence on the result of scheme evaluation, and the reasonable choice of parameter initialization method is conducive to improving the accuracy of scheme evaluation [30]. At present, there are three methods to initialize structural parameters of the deep neural network, namely, Xavier initialization method, Gaussian initialization method, and uniform random initialization method. The specific initialization table is

(1) The Xavier initialization method: $w \sim U\left(-\sqrt{(6/n_{\text{in}} + n_{\text{out}} + 1)}, \sqrt{(6/n_{\text{in}} + n_{\text{out}} + 1)}\right)$, where n_{in} and n_{out} represent the number of neurons in the input layer and output layer connected by the weight and U represents the uniform distribution.

(2) Gaussian initialization method: $w \sim N(0, 0.01)$, N is the Gaussian distribution.

(3) Uniform random initialization: $w \sim U\left(-\sqrt{(1/d)}, \sqrt{(1/d)}\right)$, where d represents the number of input layer neurons connected by weights.

Because the data distribution of cultural and creative projects is similar to the Gaussian distribution, the Gaussian initialization method is used to initialize the network structure parameters. When using the contrast divergence method to update RBM weights, learning rate and momentum are two key training parameters, which determine the update rate of network structure weights, and can accelerate the speed of model convergence and training. In the selection process of learning rate, if the learning rate is too large, it will cause the solution of the parameters to diverge near the best point, resulting in the increase of training error, the weight will be too large, and the evaluation effect of the model is not ideal; although the selection of small learning rate can effectively avoid the above problems, it will also lead to slow convergence speed and longer training time of the evaluation model [31]. Adding momentum element in the process of training can make each weight update not only related to the gradient, but also related to the last updated weight, so as to improve the evaluation effect. Generally, the learning rate is between 0 and 0.1, and the momentum is below $0.5 \sim 0.8$.

Hinton proposes the contrast divergence method, and the reconstruction times of the constrained Boltzmann machine is generally 1, which can get good results. The iteration times of the backpropagation algorithm used in the structural parameter optimization process of the rating model can be obtained by the BP error convergence of the actual model [32].

3.5. Training Method of Evaluation Model. The training process of the deep neural network model for the evaluation of cultural and creative design schemes is shown in Figure 4.

(1) Preprocess the parameter, assigning a value to visual layer V. The range normalization method is used to preprocess all the cultural and creative evaluation schemes $P_1, P_2, \ldots, P_i, \ldots, P_N$ in the cultural and creative design scheme data set S, and the partially normalized cultural and creative design scheme $Pi(1 = 1, 2, \ldots, t)$ is used as the visual layer of the deep neural network model to input into the training network model; the Gauss initialization method is used to initialize the network structure parameter 0, in which the connection weight w and the visual layer offset a are used. The values of hidden layer bias b are all 0.1; the parameter learning rate is set to 0.1, the momentum p is set to 0.5, the number of RBM reconstruction is set to 5, and the number of BP reverse iteration is set to 500.

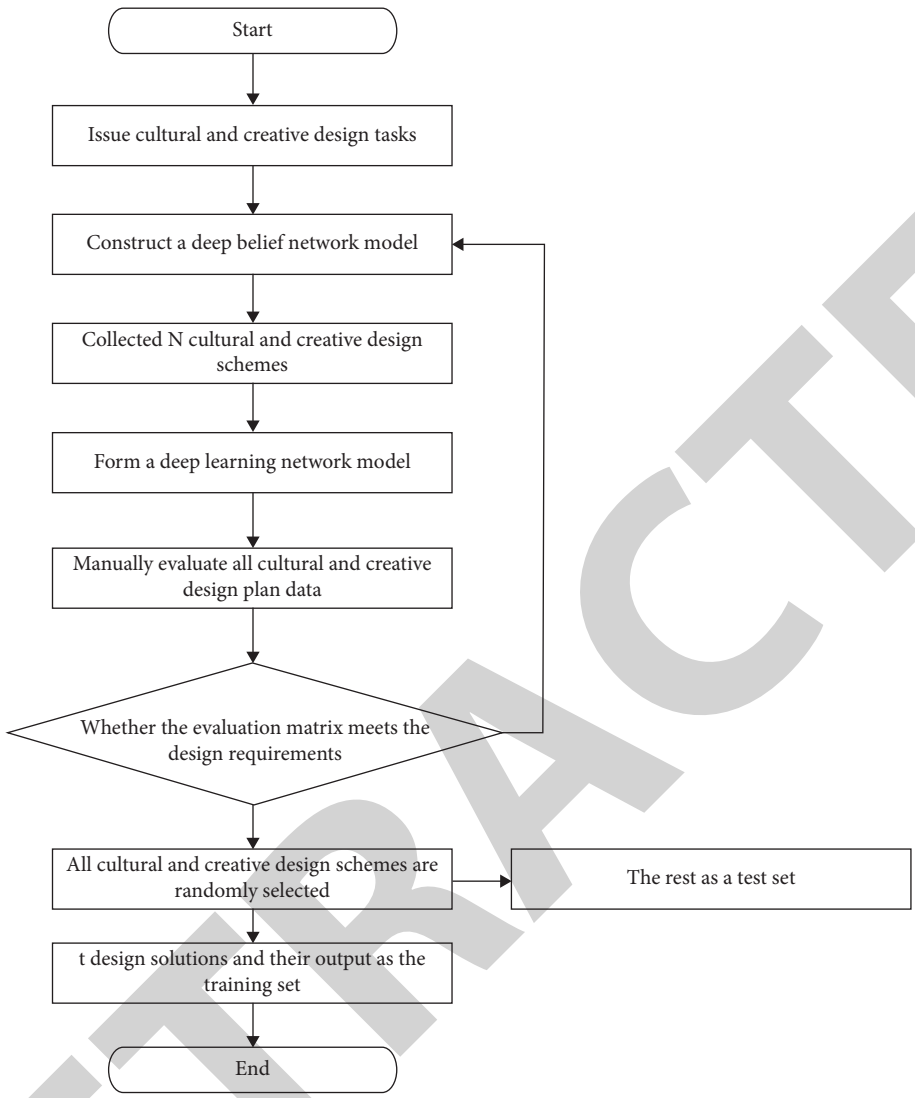


Figure 4: Cultural and creative design scheme process based on the deep learning network model.

(2) The hidden layer node value h is calculated. The activation probability $P(h_i = 1|v\colon 0)$ of hidden layer node h_i is calculated according to the following equation:

$$P(h_i = 1|v, \theta) = \text{sigma}\left(\sum_{n}^{i=1} w_{ij} v_i + b_j\right), \tag{3}$$

where θ is the activation function of sigmoid.

The value of the hidden layer node is determined by a randomly generated set of 0-1 random numbers. The value of the hidden layer node h distribution is obtained by judging the condition that the value is 1 or 0.

(3) Calculate and reconstruct visual layer v'. After calculating the value distribution result of the calculated hidden layer, equation (2) is used to calculate the value of each node of the visible layer, and the reconstruction is completed to obtain the value distribution h of the visible layer:

$$P(v_i = 1|h, \theta) = \sigma\left(\sum_{m}^{j=1} w_{ij} h_j + a_i\right). \tag{4}$$

(4) Calculate the weight update value. The adjusted values of various parameters are calculated according to the following equation:

$$\Delta w_{ij} \leftarrow \rho \cdot \Delta w_{ij} + \eta\left(P\left(h_j = 1|v\right) \cdot v - P\left(h_j - 1|v\right) \cdot v\right). \tag{5}$$

where ρ is the momentum and $\eta$is the learning rate.

(5) RBM is trained layer by layer. Record the calculated RBM structure parameters, and input the hidden layer h as the visual layer of the next restricted Boltzmann machine into the network structure. Use Step 2, Step 3, and Step 4 to calculate the structure parameters of the second restricted Boltzmann machine until all the structure parameters of the

restricted Boltzmann machine are calculated, and complete the pretraining process.

(6) Backpropagation algorithm optimization. The trained DBN is taken as a neural network with the same structure, and the structure parameters of each RBM are taken as the initial parameters of the neural network. The corresponding design result Y_i of the cultural and creative design scheme $P_i\,(i = 1, 2, \ldots, t)$ used for training is taken as the label, which is input in the output layer, and the BP algorithm is used for reverse training to adjust the parameter error of the whole structure.

(7) Test and evaluate the accuracy of the model. The untrained cultural and creative design scheme data $P_i\,(i = t + l, t + 2, \ldots, N)$ are tested to obtain the output design result y_i' of the deep training model. The y_i deviation degree of y_i' from the known manual evaluation result was calculated. If the deviation degree is within the allowable range, the training is completed; otherwise, the parameter structure is changed and the training is reconducted.

4. Evaluation Method of Cultural and Creative Design Plan

Clustering is to divide the data objects in the data space into different classes or clusters according to a unified standard, so that the similarity of data objects in the same class is as large as possible, and data objects that are not in the same cluster are as diverse as possible. After clustering, the data of the same category should be gathered together as much as possible, while the data of different categories should be separated as much as possible. Moreover, the clustering of cultural and creative design schemes can be described as follows: in the data space, N cultural and creative design schemes constitute the data set S, and the cultural and creative design scheme data point $P_i = (a_{i1}, a_{i2}, \ldots, a_{ik})$. Each attribute of P_i is numeric. In addition, the ultimate goal of clustering is to divide the data set S into H partitions $d_f\,(f = 12, \ldots, h)$. It is also possible that some data objects do not belong to any partition, and these are noise S. The union of all these divisions and noise is the data set S, and there is no intersection between these divisions, namely,

$$\begin{cases} S = D_1, \\ D_1 v_j = Q, \quad (i \neq j). \end{cases} \tag{6}$$

These divisions d_f are the clustering results of the cultural and creative design plan.

The number of evaluation indicators for cultural and creative design schemes is massive, and the total amount of design schemes is large. To make the results of clustering reflect the laws of cultural and creative design schemes, the clustering algorithm for cultural and creative design schemes should meet the following three requirements:

(1) Adopt clustering algorithms that can handle the same amount of cultural and creative design plan data sets. Some clustering algorithms can perform clustering well on smaller data sets with data objects within dozens of them. Sometimes to reduce the number of data objects to be processed, sampling methods are used. Sampling can improve the efficiency of clustering, which will also affect the results of clustering and even cause errors. Therefore, sampling methods should be adopted reasonably.

(2) The clustering method that is not sensitive to individual data has a good classification effect on the data distributed near the center of the class, but for the data far away from the center of the class, the clustering results are very different. For example, under different input orders, some data objects will be divided into different clustering groups with the help of the same clustering algorithm, which is not conducive to further analysis and processing. Therefore, the selected clustering algorithm should be able to obtain the same clustering results after clustering all the data.

(3) It must have the ability to process high-dimensional data. The data dimension of the cultural and creative design plan is determined by the number of evaluation indicators, and the number of evaluation indicators is more than 10. In addition, compared with two-dimensional data and three-dimensional data, the clustering of cultural and creative design schemes requires the use of clustering algorithms that can process high-dimensional data and analyze the distribution of data.

The k-medoids clustering algorithm is called the k-center point algorithm, which is an algorithm that uses the object closest to the center in the data object cluster to represent the forbidden algorithm. It is an improvement of the k-means algorithm. Moreover, the k-means algorithm uses the data centroid of each cluster to classify. Compared with noise and outlier data, a very large value will have a greater impact on the calculation result of the centroid. Meanwhile, the k-center point algorithm replaces the center of mass with the center point, which can effectively eliminate this effect. The processing process of the k-center point algorithm can be divided into the following four steps:

(1) Determine the number of clustered data clusters K, and initialize the center point. First determine the number of clustered data clusters K, and randomly select K cultural and creative design data points from the cultural and creative design data set $S = \{P_1, P_2, \ldots P_j, \ldots P_K\}$ as the initial center point.

(2) Divide the data points of all cultural and creative design schemes into clusters. For any cultural and creative design scheme $P_j \in S\,(j = 1, 2, \ldots, N)$, calculate the distance $d_f = \|p_j - D_f\|$ from p_j to each center point. If the distance $f = t$ obtains the minimum value $d_{\min}$ at time $f = t$, divide the data point p_j into the t-th cluster. Among them, the calculation of distance d_f adopts Euclidean distance, and its formula is

$$d_f = \left\| p_j - D_f \right\| = \left(\sum_{y=1}^{k} \left| P_{jy} - D_{fy} \right|^2 \right)^{1/2}. \quad (7)$$

(3) Calculate the square error criterion function E. After all the sample clustering of cultural and creative design schemes are completed, a unified criterion should be used to describe the excellence of the clustering results, namely, the compactness of the data clusters. What is more, the square error criterion E is the judgment index obtained by summing the distance squares from the currently selected center point to all sample points in the cluster. It can be considered that the smaller the value of E is, the more compact the cluster will be. When E achieves the minimum value, the center point at this time is the characteristic scheme. The calculation formula of the square error criterion is

$$E = \sum_{j}^{N} \left\| P_j - D_f \right\|. \quad (8)$$

(4) Update the center point and find the square error criterion function E until E takes the minimum value. The current center point D_f is arbitrarily taken from the cultural and creative design plan data set S, and its square error criterion function E is not necessarily the minimum value. Take data points from each cluster to update the center point, and repeat Steps 2 and 3 to obtain the value of the smallest square error criterion E in the cluster, and use this point as the new center point to recluster. Repeat the above steps until the new center point set is the same as the original center point set, and the algorithm terminates.

The center point plan of the final clustering result $D_1, D_2, \ldots, D_h$ is selected to construct the feature plan set $s' = \{p_1', p_2', \ldots, p_j', \ldots, p_h'\}$, and the feature plan set is used as the sample of learning evaluation rules for manual evaluation.

5. Simulation Experiment Analysis

Based on the evaluation method of cultural and creative design plan proposed in this paper, a customized cultural and creative design platform is developed. Taking the cultural and creative design plan as an analysis, the performance of the method in this paper is tested to evaluate the cultural and creative design plan in MATLAB.

5.1. Model Test Conditions. The software and hardware environment of the model test is shown in Table 1.

5.2. Training Data and Test Data of the Model. The analytic hierarchy process is used to construct a judgment matrix for the upper-level indicators and the lower-level indicators to calculate the weight of each evaluation indicator. The judgment matrix constructed for the upper-level indicators is

Table 1: Model testing hardware and software environment.

Item	Configuration	Model
Hardware configuration	CPU	i7-7700 HQ
	GPU	GTX 1050 Ti
	RAM	32 GB
	Hard disk	1T SSD
Software configuration	OS	Win 10 x64
	IDE	PyCharm
	Programming language	Python 3.5
	Deep learning framework	TensorFlow 1.11

$$A_1 = \begin{cases} 1 & 1 & 2 & 4 \\ 1 & 1 & 2 & 4 \\ \frac{1}{2} & \frac{1}{2} & 1 & 2 \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{2} & 1 \end{cases}. \quad (9)$$

The eigenvector of the judgment matrix is solved by the 1-order column average solution method, and T_i is obtained by accumulating each column first:

$$T_1 = \left(\frac{11}{4}, \frac{11}{4}, \frac{11}{2}, 11 \right). \quad (10)$$

After normalizing the columns of the judgment matrix A_i, the judgment matrix D of the type is obtained:

$$D = \begin{cases} \frac{4}{11} & \frac{4}{11} & \frac{4}{11} & \frac{4}{11} \\ \frac{4}{11} & \frac{4}{11} & \frac{4}{11} & \frac{4}{11} \\ \frac{2}{11} & \frac{2}{11} & \frac{2}{11} & \frac{2}{11} \\ \frac{1}{11} & \frac{1}{11} & \frac{1}{11} & \frac{1}{11} \end{cases}. \quad (11)$$

Accumulate the elements of each row, calculate the sum value, and get the eigenvector R:

$$R_1 = \left(\frac{16}{11} \frac{16}{11} \frac{8}{11} \frac{4}{11} \right)^T. \quad (12)$$

After normalization, the weight distribution matrix W_i of the superior index is obtained:

$$W_i = (0.3636, 0.3636, 0.1818, 0.0909)^r. \quad (13)$$

After calculating the weight distribution matrix, check the consistency of the judgment matrix A_i. Then, the judgment matrix and the weight distribution matrix are used

to calculate the characteristic root matrix of the judgment matrix A_i according to formulas (2)–(9):

$$AW_i = (0.8408, 0.8408, 1.6816, 0.3633)^r. \quad (14)$$

Then calculate the maximum characteristic root according to formulas (2)–(10) $\lambda_{\max} = 4$.

For the value of RI corresponding to each level of judgment matrix, when $n = 4$, $R_i = 0.89$, and $CI = 0, CR = 0 < 0.1$ can be obtained. Then the weight result of the superior evaluation index will be credible. With the help of the above method, construct the judgment matrices A_2, A_3, A_4, and A_5 for the secondary evaluation index, and calculate the weight distribution results W_2, W_3, W_4, and W_5 of the judgment matrix.

The judgment matrix A_2 constructed for the secondary evaluation index of the structural index is

$$A_2 = \begin{cases} 1 & 2 & 2 & 2 \\ 1/2 & 1 & 1 & 1 \\ 1/2 & 1 & 1 & 1 \\ 1/2 & 1 & 1 & 1 \end{cases} \quad (15)$$

The weight distribution matrix W_2 is calculated as

$$W_2 = (0.3333, 0.1667, 0.1667, 0.1667, 0.1667)^T. \quad (16)$$

Finally, the consistency test and summary of each judgment matrix are performed, and the final weight distribution results are calculated and summarized in Table 2.

After selecting the evaluation index and calculating the weight of the evaluation index on the collected 100 groups of cultural and creative design schemes, cluster analysis of the cultural and creative design schemes can be performed.

Matlab's tsne function is used to reduce the dimension of all cultural and creative design plan data to 2 dimensions for visual analysis, and the element dimension data are still used in the process of clustering to avoid distortion. The two-dimensional distribution of 100 groups of cultural and creative data schemes is shown in Figure 5.

The k-medoids algorithm is used for clustering analysis of the cultural and creative design plan data, and a different cluster number k is set. Meanwhile, the clustering result is analyzed by the elbow method to determine the optimal cluster number k. In addition, the elbow clustering uses sum of the squared errors (SSEs) to analyze the clustering results. The smaller the SSE is, the better the clustering results will be. When the value of k is larger, the value of SSE will become smaller, but as the number of classes increases, the change of the SSE value will change. When the value of k increases for the last time, the SSE decreases significantly, which is where the optimal value of k is located. Figure 6 shows the clustering SSE of the cultural and creative design scheme of the injection parts of the injection molding machine with the k value.

5.3. Specific Experimental Model Design. In order to verify the proposed model in this paper, approximating the model input variables x_1, x_2, x_3, x_4 and output response Y_1, Y_2, Y_3 contained between the low-order and high-order power function, trigonometric function, inverse trigonometric function, and exponential function elementary function, such as strong nonlinear mapping relations, can be done through the simplified design depth model CDMT01 cross validation method.

The shallow models CSMT02 and CSMT03 are designed as their comparison models. The shallow model CSMT02 has the same training conditions, functions, and the total number of hidden neurons of the network compared with the depth model CDMT01, but only the number of model layers is different. The shallow model CSMT03 has the same training conditions and the selection of neuron activation function compared with the shallow model CSMT02, but only the number of model layers is different. The number of neurons in the hidden layer of the model is increased to 2 times, and the specific superparameters of each model are shown in Table 3.

The reason why the depth model in the model test only contains CDMT01 and no other depth model is designed as a comparison model is as follows:

(1) One of the main roles of the depth model CDMT01 in model testing is to compare the performance of the shallow models CSMT02 and CSMT03 as comparison models and quantitatively show its comparative advantages;

(2) If the depth model CDMT01 can directly achieve great advantages only through the BP algorithm in model testing (which is difficult to achieve for most DNN models designed by other rules), then this itself can fully prove that the guiding theory of DNN model design in this paper is correct, scientific, and reasonable; otherwise the DNN model will not be successful. Compared with the shallow model with mature theory, the performance of the model and the convergence speed of the parameter training are superior.

5.4. Performance Comparison of Each Experimental Model. The Adam optimization algorithm is selected as the model training algorithm, and the test data set is used to test the current training effect and performance of the model after each iteration training. After programming and operation under TensorFlow, the training results of each model are shown in Table 4 and Figure 7.

In Figure 7, the synchronous approximation effect of each experimental model for the three test functions is observed: from the analysis in the degree of approximation of function fitting, shallow model CSMT02 after 10000 iterations training for test function Y_1 is relatively the best fitting effect.

The trend chart of adjusted R-square in each experimental model for the test data set during training is shown in Figure 8.

In Figure 8, all the green lines can quickly reach the highest point (the maximum value of adjusted R-square is 1) and keep close to 1, which shows that both the depth model

Table 2: Matrix weight distribution results.

Superior indicator	Weight	Secondary indicator	Weight	Comprehensive weight
C_1	0.35	B_1	0.34	0.11
		B_2	0.15	0.06
		B_3	0.15	0.06
		B_4	0.15	0.06
C_2	0.40	B_5	0.17	0.06
		B_6	0.17	0.06
		B_7	0.19	0.08
C_3	0.15	B_8	0.25	0.13
		B_9	0.25	0.13
C_4	0.10	B_{10}	0.42	0.14
		B_{11}	0.42	0.14

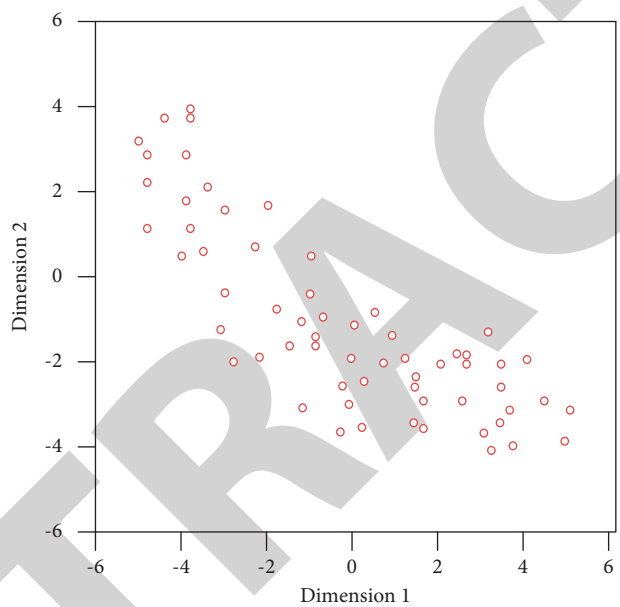


Figure 5: Two-dimensional distribution map of the cultural and creative design plan.

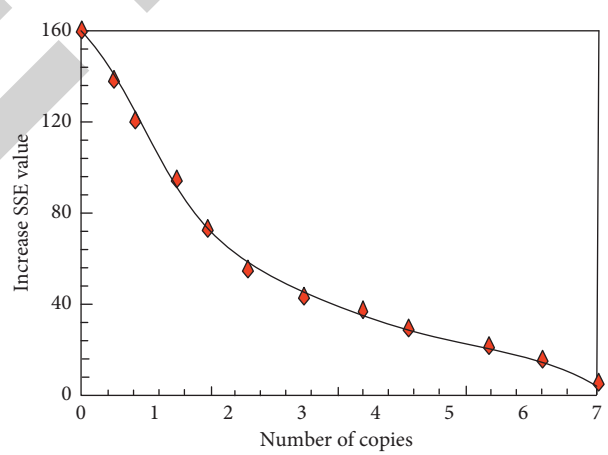


Figure 6: The clustering error squared sum (SSE) varies with the number of clusters (k).

and the shallow model have very good approximation effect for the test function Y_1.

The RMSE variation trend of the test data set during the CDMT01 training of the depth model is shown in Figure 9.

In Figure 9, in the training of approximation, the RMSE approximated by the model to each test function is only achieved after about 150 iterations. In the subsequent training process, the RMSE approximated by the model to each test

Table 3: Overview of experimental model hyperparameters.

Model	Model neuron number distribution	Distribution of neuron activation function in each layer of the model	Total hidden layer neurons
CDMT01	5-160-100-80-50-3	$a - ReLU - cr - ReLU - ReLU$	390
CDMT02	5-390-3	$a - ReLU$	390
CDMT03	5-780-3	$a - ReLU$	780

Table 4: Experimental models adjusted R-square training results for the model test set.

Network model	CDMT01	CDMT02	CDMT03
Neuron number distribution	5-160-100-80-50-3	5-390-3	5-780-3
Total hidden layer neurons	390	390	780
Y_1 test data	0.98154	0.97215	0.97678
Y_2 test data	0.98687	0.88356	0.51459
Y_3 test data	0.99994	0.84569	0.34586
Number of iterations	9900	9999	9999

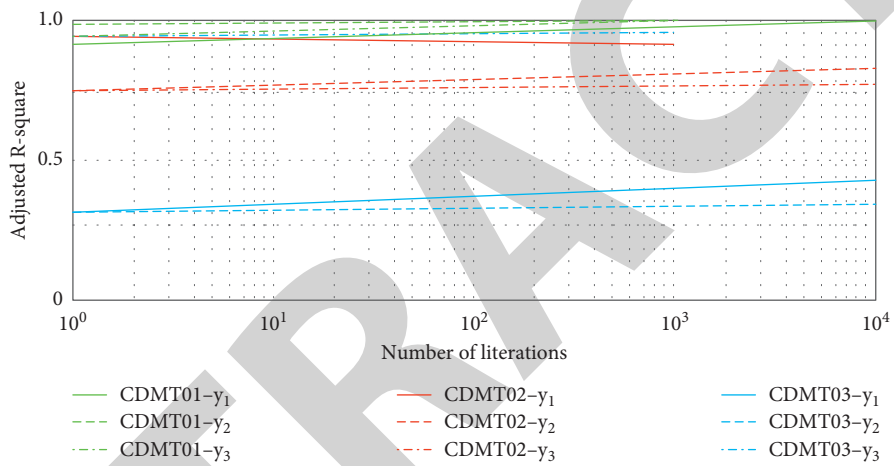


Figure 7: Changes in the adjusted R-square for the test set after the 9000th iteration training.

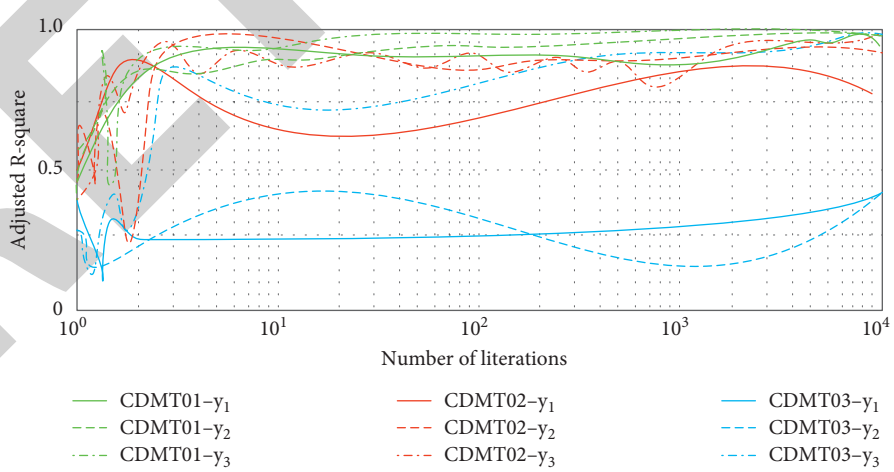


Figure 8: Trend of the adjusted R-square for the model test set during the training of the experimental model.

function is slowly improved, but in the process of continuous training of the model, the RMSE approximated by the model to each test function is also continuously improved. The RMSE variation trend of the test data set during the training of the shallow model CSMT02 is shown in Figure 10.

In Figure 10, the shallow model has reached a relatively low RMSE, with only 200 iterations in the training of approximation of each test function. Compared with the depth model CDMT01, only 150 iterations are about 33.33% slower. In the course of subsequent training, the shallow

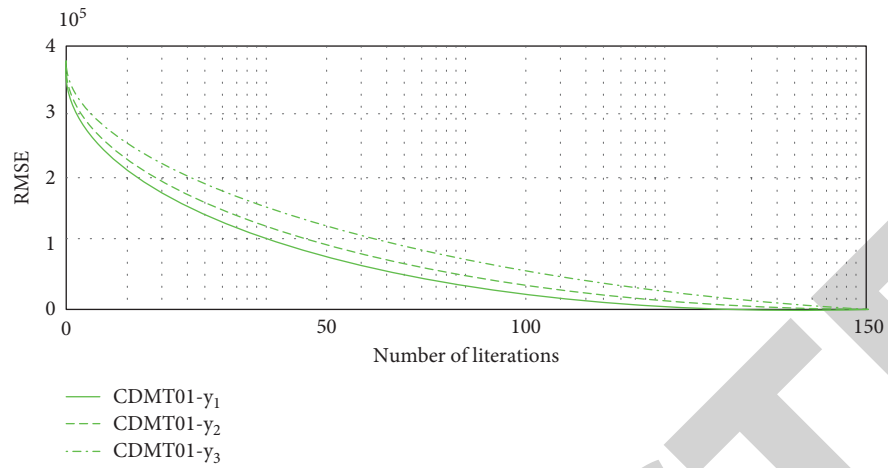


Figure 9: Trend of RMSE for the model test set during the depth model CDMT01 training.

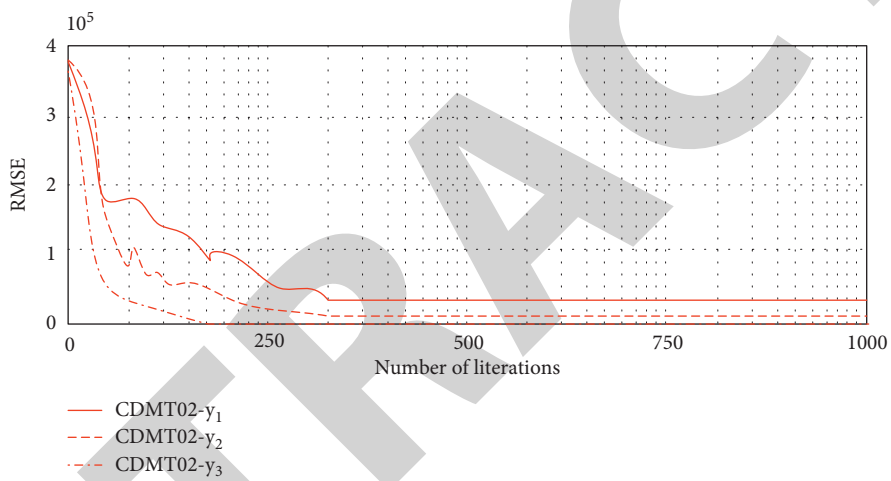


Figure 10: Trend of RMSE for the model test set during the shallow model CSMT02 training.

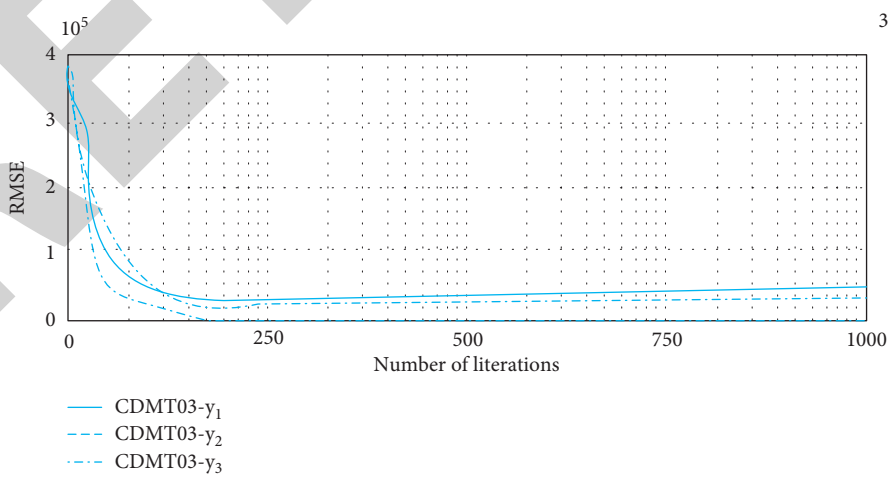


Figure 11: Trend of RMSE for the test set during the training of the shallow model CSMT03.

model is consistent with the depth model. The RMSE variation trend of the test data set during the training of the shallow model CSMT03 is shown in Figure 11.

It can be found in Figure 11 that, from the perspective of training trend, there is basically no difference between the shallow models CSMT03 and CSMT02. Although the

Table 5: Experimental model RMSE training results for the model test set.

Network model	CDMT01	CDMT02	CDMT03
Neuron number distribution	5-160-100-80-50-3	5-390-3	5-780-3
Total hidden layer neurons	390	390	780
Y_1 test data	0.15459	1.26566	0.28956
Y_2 test data	0.01484	3.16544	1.64326
Y_3 test data	0.02459	3.16478	2.16542
Number of iterations	9900	9999	9999

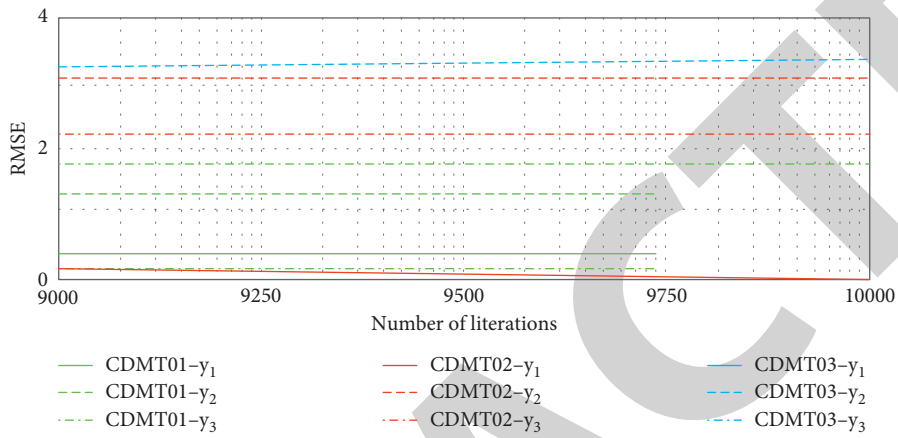


Figure 12: Partial enlargement of the trend of the RMSE for the model test set at the end of the experimental model training.

shallow model CSMT03 has more hidden layer neurons than CSMT02, the model CSMT03 is worse than CSMT02 on the whole in terms of the accuracy result of function approximation of the model.

The synchronous approximation effect of the experimental model to the three test functions is shown in Table 5 and Figure 12.

The test results show that the depth network model can effectively approximate three test functions with different data distribution and difficult function approximation at the same time compared with the traditional response surface. Compared with the shallow model as a contrast model, the adjusted R-square of fitting degree of the test function is nearly 3 times higher than that of the test function. At the same time, the RMSE of fitting accuracy index is even nearly 7 times higher for a test function.

6. Conclusion

Based on clustering and deep belief network model, the evaluation method of cultural and creative design is studied in this paper, which mainly includes cultural and creative design plan analysis and evaluation index weight calculation method, cultural and creative design plan evaluation model based on deep belief network, and cultural and creative design plan evaluation method based on clustering. Moreover, the methods and techniques proposed in this paper are applied to the evaluation of cultural and creative design schemes to simulate the cultural and creative design platform, and the theoretical methods of this paper are verified through tests.

The use of deep belief networks and convolutional neural networks has achieved excellent results in image recognition and classification. However, it is difficult to directly apply to the evaluation of cultural and creative design schemes. If the information can be extracted from the design scheme described by the two-dimensional image, and the deep learning model can be used to obtain the scoring result, the accurate expression of the data can be combined with the accurate output of the evaluation result, which is more conducive to cultural and creative design.

Data Availability

All data used to support the findings of this study are included within the article.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

[1] V. Niaros, V. Kostakis, and W. Drechsler, "Making(in) the smart city: The emergence of makerspaces," *Telematics and Informatics*, vol. 34, no. 7, 2017.

[2] S.-Y. Han, J. Yoo, and H. Zo, "Understanding makers pace continuance:aself-determination perspective," *Telematics and Informatics*, pp. 184–195, 2017.

[3] E. C. Ozcan, S. Unlusoy, and T. Eren, "Acombined goal programming-AHP approach supported with TOPSIS for maintenance strategy selection in hydroelectricpowerplants," *Renewable and Sustainable Energy Reviews*, vol. 78, pp. 1410–1423, 2017.

Hindawi
Scientific Programming
Volume 2021, Article ID 3315695, 12 pages
https://doi.org/10.1155/2021/3315695

Research Article

Iranian COVID-19 Publications in LitCovid: Text Mining and Topic Modeling

Meisam Dastani [1] **and Farshid Danesh** [2]

[1] *Infectious Diseases Research Center, Gonabad University of Medical Sciences, Gonabad, Iran*
[2] *Information Management Department, Regional Information Center for Science and Technology (RICeST), Shiraz, Iran*

Correspondence should be addressed to Farshid Danesh; farshiddanesh@ricest.ac.ir

Received 27 April 2021; Revised 21 July 2021; Accepted 27 August 2021; Published 16 September 2021

Academic Editor: Shah Nazir

COVID-19 is a threat to the lives of people all over the world. As a result of the new and unknown nature of COVID-19, much research has been conducted recently. In order to increase and enhance the growth rate of Iranian publications on COVID-19, this article aims to analyze these publications in LitCovid to identify the topical and content structure and topic modeling of scientific publications in the mentioned subject area. The present article is applied research performed by using an analytical approach as well as text mining techniques. The statistical population is all the publications of Iranian researchers in LitCovid. Latent Dirichlet Allocation (LDA) and Python were used to analyze the data and implement text mining and topic modeling algorithms. Data analysis shows that the percentage of Iranian publications in the eight topical groups in LitCovid is as follows: prevention (39.57%), treatment (18.99%), diagnosis (18.99%), forecasting (7.83%), case report (6.52%), mechanism (3.91%), transmission (3.62%), and general (0.58%). The results indicate that patient, pandemic, outbreak, case, Iranian, model, care, health, coronavirus, and disease are the most important words in the publications of Iranian researchers in LitCovid. Six topics for prevention; four topics for treatment and case report and forecasting; three topics for diagnosis, mechanism, and transmission in general have been obtained by implementing the topic modeling algorithm. Most of the Iranian publications in LitCovid are related to the topic "pandemic status," with 22.47% in the prevention category, and the lowest number of publications is related to the topic "environment," with 11.11% in the transmission category. The present study indicates a better understanding of essential and strategic issues of Iranian publications in LitCovid. The results reveal that many Iranian studies on COVID-19 were primarily on the issues related to prevention, management, and control. These findings provided a structured and research-based viewpoint of COVID-19 in Iran to guide researchers and policymakers.

1. Introduction

In December 2019, an epidemic with mild respiratory infections was first reported to the World Health Organization (WHO) in Wuhan, China's largest metropolitan area in the Hubei Province. Since it was impossible to identify the causative agent, people who suffered from this disease were classified as "pneumonia of unknown etiology." The CDC (Center for Disease Control and Prevention) in China organized an extensive research program to study the extent and prevalence of the disease. The etiology of this disease is currently attributed to a new virus named SARS-CoV-2, belonging to the coronavirus family and causing COVID-19 [1]. The Director-General of the WHO, "Tedros Adhanom Ghebreyesus," on February 11, 2020, announced that the disease is caused by this new coronavirus and named it "COVID-19." According to studies and evidence, COVID-19 is highly contagious and has a high global prevalence rate. At the International Health Regulations (IHR) meeting on January 30, 2020, the outbreak of COVID-19 was identified as a global health threat, since by that time, it was reported that it had spread to 18 countries by human-to-human contact [2].

Because of the increase in the COVID-19 pandemic, the WHO stated on January 30, 2020, that the new coronavirus is the sixth most common public health emergency worldwide, threatening China and all countries [2]. The

disease is also being rapidly spread. By August 06, 2021, global statistics on the coronavirus have indicated that more than 200 countries from all continents were struggling with COVID-19. The Islamic Republic of Iran, located in Southwest Asia, reported the first case of COVID-19 on February 19, 2020, in Qom city [3]. According to the website of Johns Hopkins University Coronavirus Resource Center (https://coronavirus.jhu.edu/map.html), by August 16, 2021, the total COVID-19 cases in Iran were 4467015 (ranked 11th in the world) and total deaths were 98483 (ranked 13th in the world).

To identify the various aspects of COVID-19, the publication and dissemination of results of all scientific activities in COVID-19 have become of particular importance [4], and significant publications in this field have been indexed in international citation databases. The investigations indicated that such a volume of scientific publications had not been observed in any scientific field and in such a short time [5]. In addition to the high number of COVID-19 publications, they are also being increased. However, medical scientists are still faced with many ambiguities; hence, extensive research is being conducted on various aspects of COVID-19 in many countries. Scientists worldwide are conducting significant studies on the aspects and methods of prevention, treatment, and the development of effective vaccines and medications to fight and eradicate COVID-19, and the results of their research projects are published in peer-reviewed journals [6]. The increase in global publications on COVID-19 has been influenced by researchers' events and efforts to fully understand the pandemic circumstances [6].

The reports revealed that scientists and researchers in China, the United States, the United Kingdom, and Italy had published the highest number of publications on COVID-19 [7, 8]. Because of the high volume of reputable publications on COVID-19, many studies with a macro and exploratory approach for policy-making and strategic planning of COVID-19 research have been done [8]. Text mining is one of the methods applied to analyze scientific publications on COVID-19 and involves extraction of information and latent knowledge in texts, especially scientific texts, and converting tacit knowledge into explicit knowledge [9, 10]. The purpose of text mining is to extract knowledge from textual data and has many applications in processing and analyzing scientific publications [11]. Therefore, text mining and text extraction techniques are used to analyze scientific texts and discover latent knowledge and topics discussed in a set of these texts [12–15]. Topic modeling is a statistical method and an essential algorithm in text mining, applied to identify latent topics in a set of textual documents [16, 17]. The results of this algorithm are extensively used for research policy-making and strategic planning [18] and allow analysts and researchers to comprehend better the relationships and variations of the provided topics [19].

After the COVID-19 pandemic, the "Information Explosion" phenomenon is witnessed regarding scientific publications on COVID-19. A review of PubMed revealed that more than 150,000 records related to COVID-19 had been indexed in this medical science database so far. Three researchers named Qingyu Chen, Alexis Allot, and Zhiyong Lu from the United States, who work at the National Center for Biotechnology Information (NCBI), introduced the LitCovid as a specialized COVID-19 research database, which was designed with the support of the US National Institutes of Health's Intramural Research Program (IRP) and its data are being updated daily (Figure 1-Appendix). More than 150,000 records are indexed on LitCovid; this number is increasing and PubMed is the primary data source for LitCovid (Figure 1-Appendix). In LitCovid, research on COVID-19 has been divided into eight categories: general, mechanism, transmission, diagnosis, treatment, prevention, case report, and forecasting [20, 21].

According to the world meter database (https://www.worldometers.info/coronavirus/country/iran/), the I.R. Iran, with a population of 85196159 people, total cases of 4467015, total deaths of of 98483, and total recovered cases of 3757157, among Asian countries, ranks third in terms of total deaths after India and Indonesia and ranks third in terms of total cases and total recovered cases after India and Turkey. Iran ranks 11th in terms of total recovered and total cases and 13th in terms of total deaths compared to other countries. According to the COVID-19 vaccination tracker provided by the New York Times, Iran's data are as follows: 16.2M doses of the COVID-19 vaccines have been given to Iranians.3.15 million people (3.8%) received two doses of the vaccine and were fully vaccinated (https://www.nytimes.com/interactive/2021/world/Covid-vaccinations-tracker.html). All the mentioned data regarding COVID-19 statistics and vaccination were extracted on August 16, 2021.

The number of total cases, deaths, and fully vaccinated population in proportion to the total population of Iran (2.8% fully vaccinated) has made this country one of the most important countries globally in terms of the COVID-19 pandemic. Therefore, it is necessary to conduct a study by text mining methods and topic modeling to determine the topics of Iranian publications on COVID-19. It is necessary to answer the following questions to achieve the main objective of this paper:

(i) RQ1: What are the percentages of global and Iranian publications based on the eight-topic division of LitCovid?

(ii) RQ2: What are the most important words applied based on the repetition rate in Iranian publications in LitCovid?

(iii) RQ3: What are the most important words applied based on TF-IDF weighting in Iranian publications in LitCovid?

(iv) RQ4: What are the main topics and subtopics of Iranian publications on COVID-19 in LitCovid based on the topic modeling algorithm?

(v) RQ5: What are the percentages of Iranian publications on COVID-19 in the eight topics of LitCovid and sub-topics extract from text mining?

2. Related Works

The studies indicate that the COVID-19 pandemic has negatively affected various aspects of the mental and physical

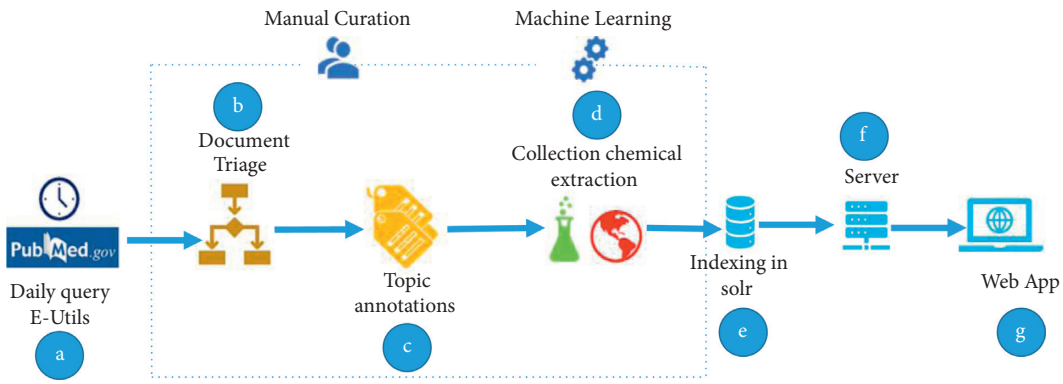


Figure 1: Daily workflow of LitCovid [20].

health of people in the community and has so far killed more than 4 million people worldwide. However, in recent waves of the disease, its adverse effects have diminished, but new disease variants (such as the Delta variant) are still emerging, and the whole world is still in a state of emergency [22]. Undoubtedly, the advancement of medical knowledge and related technologies will improve people's health [23]. The advancement of science is directly related to social needs and cultural elements. Scientists conduct research projects to solve social problems and anomalies.

Many studies have analyzed the publications of medical sciences using the text mining method and its topic modeling algorithms, including studying topic areas of biomedicine [12], personality disorders [24], clinical research [15], AIDS and economic evaluation [25], HIV prevention [26], health information [27], medical informatics [28], and health and cybersecurity [29]. The objective of this paper was the text mining of Iranian publications on COVID-19 in LitCovid. Hence, the text-mining-related works on COVID-19 have been reviewed in the following. It is noteworthy that some studies have analyzed and evaluated COVID-19 articles using scientometrics methods and indicators [6, 30–32]. The authors also applied the text mining method in the present paper.

Tran et al. have used text mining to evaluate the COVID-19 publications by April 23, 2020. In the present paper, countries were classified based on gross domestic product and World Bank report, and ten highly cited documents of each country were analyzed. The most important research topics were "Guidelines for emergency care and surgery," "viral pathogenesis," and "global responses in the COVID-19 pandemic" [33]. In line with previous research, Radanliev et al. also used WOSCC data to analyze articles related to mortality, immunity, and vaccine in the topic area of COVID-19 using data mining techniques. They identified the relationships between concepts and keywords in each scientific topic and collaborations between countries regarding scientific publications [34].

In a study conducted by Dong et al., the publishing features of COVID-19 in CORD-19 (COVID-19 Open Research Dataset) were analyzed by applying topic modeling. They identified important topics in COVID-19 publications and highlighted topics. The results of their study identified eight main topics for COVID-19 publications and indicated that researchers should conduct further investigations on diagnostics, therapeutics, vaccines, viral genomics, and pathogenesis. In another research, the topics of global publications on COVID-19 in the CORD-19 were identified using text mining and topic modeling techniques; this study revealed that the COVID-19 publications have focused more on and have paid less attention to which topics [35]. In another study, the publications related to SARS, MERS, and COVID-19 were evaluated; in this article, the text mining of topics was performed separately, the scientific publications related to each of the viruses were modeled independently, and finally, the results were reviewed using an analytical-comparative approach [36]. In another paper, COVID-19 publications were extracted in the first six months of the pandemic using PubMed, and the topics of COVID-19 articles and their publication trends were extracted using topic modeling and text mining techniques [37].

A literature review indicated that text mining techniques were applied in text analysis and knowledge discovery and extraction in a large volume of texts and analysis of texts published in citation databases and identifying research trends in various scientific fields. According to the objectives and the research population, special text mining techniques have been used in each reviewed literature. Identifying the high-frequency keywords and the essential topics using special text mining algorithms was one of the essential techniques applied in scientific texts. The topic modeling algorithm, because of its high accuracy, was one of the most important algorithms employed in previous research to determine clusters and topics. The searches conducted in national and international citation databases and a review of foreign and domestic literature indicated that no similar research was observed in the topic area of Iranian publications on COVID-19 in LitCovid.

3. Materials and Methods

The present article deals with applied research conducted by an analytical approach and also the text mining technique. The statistical population was all Iranian publications of COVID-19, which were extracted from LitCovid in February 2021. LitCovid has been designed to track and rapidly access COVID-19 publications. Given that LitCovid is updated

daily and extracts its data from PubMed, it allows researchers to review the latest and most reputable scientific publications in the field of COVID-19 [20, 21]. All articles indexed in LitCovid are classified into eight categories: general, mechanism, transmission, diagnosis, treatment, prevention, case report, and forecasting. In the present research, text mining and topic modeling of titles, abstracts, and keywords of Iranian publications on COVID-19 have been performed separately in each of the eight categories of LitCovid.

3.1. Text Mining and Topic Modeling. Text mining analysis included three main steps: preprocessing, text mining, and post-processing, generally after collecting textual data [38]. Preprocessing texts included the selection of documents, extraction of the words used in texts, unification of texts using the manual analysis of articles' keywords and unification of synonymous words, and removing meaningless words and stop-words. In the present study, the most important words were identified by the Term Frequency–Inverse Document Frequency (TF-IDF) method, which indicated the importance of the word in a document or a set of documents. The main objective of this statistical method was to show the importance of the word in the text [28, 39]. In order to perform the preprocessing mentioned on the data and before extracting the most important words based on TF-IDF, the words used in the text of the articles were stemmed by Porter stemmer [40, 41]. In the next step, the frequency and weighting of the words, implementation of topic modeling, and visualization were done through different text mining techniques, and in the final step, the knowledge was extracted and interpreted. The LDA algorithm has been applied for topic modeling in all categories of LitCovid in the present paper.

In the present research and the mentioned preprocessing on the data and implementing the LDA topic modeling algorithm, the Bigrams and Trigrams have also been extracted from the texts to extract valuable data [42]. The Bigram was a set of two adjacent words, and Trigram was three adjacent words. For instance, Machine and Learning were two separate words, but their Bigram is extracted as Machine Learning [36].

Figure 2 indicates the idea behind the LDA logic. This method assumed that the set of documents consists of many topics, each containing many words (left Figure 2). Hence, it can be imagined that each document is created as follows: first, distribution of topics is selected (Figure 2), then a topic assignment (colored circles in the figure) is selected for each word, and finally, the desired word is picked from the relevant topic. It should be noted that the topics and topic assignments shown in Figure 2 were only presented as illustrative examples and were not derived by applying them to real data [17].

The basic idea of this process was that a combination of topics models each document, and each topic was a discrete probable dispersion that determined how the probability of each word appeared in the topic. These topic probabilities provided an accurate representation of the document. A "document" was a "bag of words" with no structure beyond word and topic statistics.

Since the LDA algorithm did not determine the number of desirable topics, the logarithmic (log) UMass Coherence criterion has been applied in the present study to determine the number of desirable topics [43]. The UMass logarithm proposed different values of topics. In the present article, a different number of specific topics have been extracted for each of the topic categories of Iranian COVID-19 publications in the LitCovid. Guidance has been taken from medical sciences specialists to achieve the desired number of topics in each category. It is important to note that the number of topics should be selected proportionally because a large number of topics will lead to a significant quantity of small and considerably similar topics [38, 39]. Also, interpretation of topics becomes more challenging due to the dispersion of keywords between topics [40]. Afterward, the topics resulting from the implementation of the LDA algorithm were interpreted using the most important words and publications of each topic. Python and the libraries related to text mining, including Gensim, NLTK, and Spacy, were applied to implement text mining algorithms such as word frequency determination, TF-IDF, and topic modeling algorithms [41]. Python 3.9.5 and the libraries related to text mining, including Gensim, NLTK, and Spacy, were applied to implement text mining algorithms such as word frequency determination, TF-IDF, and topic modeling algorithms [44]. Python is an open-source, compact, and versatile programming language with a simple syntax. It is also straightforward to develop and provide various libraries for working with texts [44].

4. Results

In this section, the collected data were analyzed based on the methodology described in the previous section. The research questions are answered in the following.

4.1. The Percentages of COVID-19 Global and Iranian Publications Based on the Eight-topic Division in Litcovid. Figure 3 indicates the percentage of global and Iranian publications based on the eight-topic division in LitCovid. It should be mentioned that the percentage of Iranian publications in each of the eight categories is measured based on the total number of Iranian publications, and then the percentage of each topic is obtained. The same method has been applied to calculate the percentage of global publications. It was evident that the highest percentage of global and Iranian publications on COVID-19 was in the category of "Prevention"; 39.57% and 32.86% of the publications of Iranian and global researchers in LitCovid, respectively, were in the category of prevention. Moreover, the analysis of the data in Figure 3 revealed that the lowest percentage of Iranian research projects (0.58%) is in the "general" category, and the lowest percentage of global publications (1.94%) is related to the "Forecasting" category.

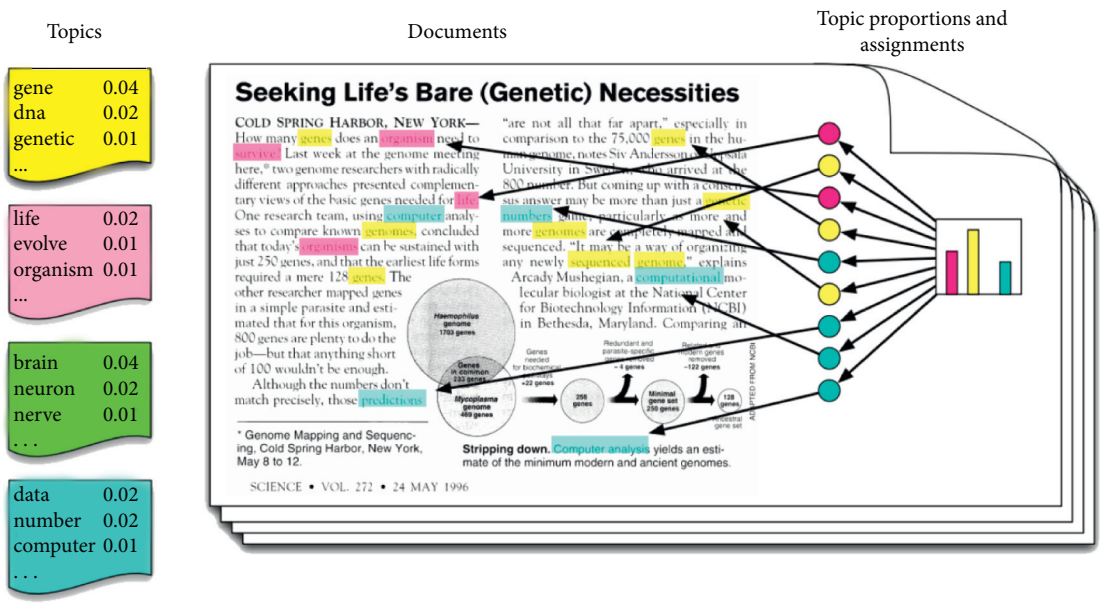


Figure 2: The idea behind the Latent Dirichlet Allocation (LDA) algorithm [17].

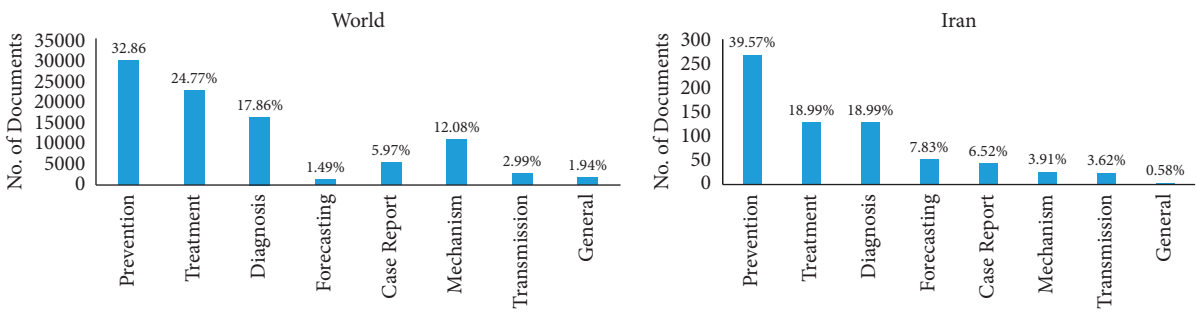


Figure 3: Percentage of COVID-19 global and Iranian publications based on the eight-topic division in LitCovid.

4.2. The Most Important Words Applied Based on the Repetition rate in Iranian COVID-19 Publications in the LitCovid. Figure 4 indicates the 20 words with the highest frequencies in Iranian publications. Figure 5 also illustrated the word cloud of 100 words with the highest frequency in Iranian publications on COVID-19. The data analysis showed that COVID, patient, and Iran had the highest frequencies in Iranian publications on COVID-19 in LitCovid with 2415, 1162, and 887, respectively (Figures 4 and 5).

4.3. The Most Important Words Applied Based on TF-IDF Weighting in Iranian COVID-19 Publications in the LitCovid. The data in Figure 6 included 20 important words based on TF-IDF weighting. Figure 7 also showes the word cloud of 50 top words based on TF-IDF weighting. The analysis of data revealed that "patient," "pandemic," and "outbreak" with the weights of 6.46, 4.57, and 4.29, respectively, are the most important words of Iranian publications in LitCovid based on TF-IDF weighting.

4.4. The Main Topics and Subtopics of Iranian Publications on COVID-19 in the LitCovid Based on the Topic Modeling Algorithm. The data in Table 1 indicates the implementation of the topic modeling algorithm on Iranian publications on COVID-19 based on the eight topic category of the LitCovid. The tags subtopics have been identified based on the most important words and a review of the most significant articles of each subtopic. The number of documents for each of the main topics was also mentioned, e.g., in the topic "general" with four documents, only one topic with the same general title has been identified. The topic "mechanism" involved 18 documents, and three topics named "characteristics," "clinical features," and "genomic sequence" have been identified.

4.5. The Percentages of Iranian COVID-19 Publications in the Eight Topics of LitCovid and the Subtopics Extracted from Text Mining. Figure 8 indicates the percentages of Iranian publications in each of the eight categories of LitCovid and subtopics resulting from text mining of Iranian publications. The analysis of data shown in this figure revealed that in the topic area of "mechanism," the publications percentages of the topics "characteristics," "genomic sequence," and "clinical features" are equal to 38.89%, 38.89%, and 22.22%,

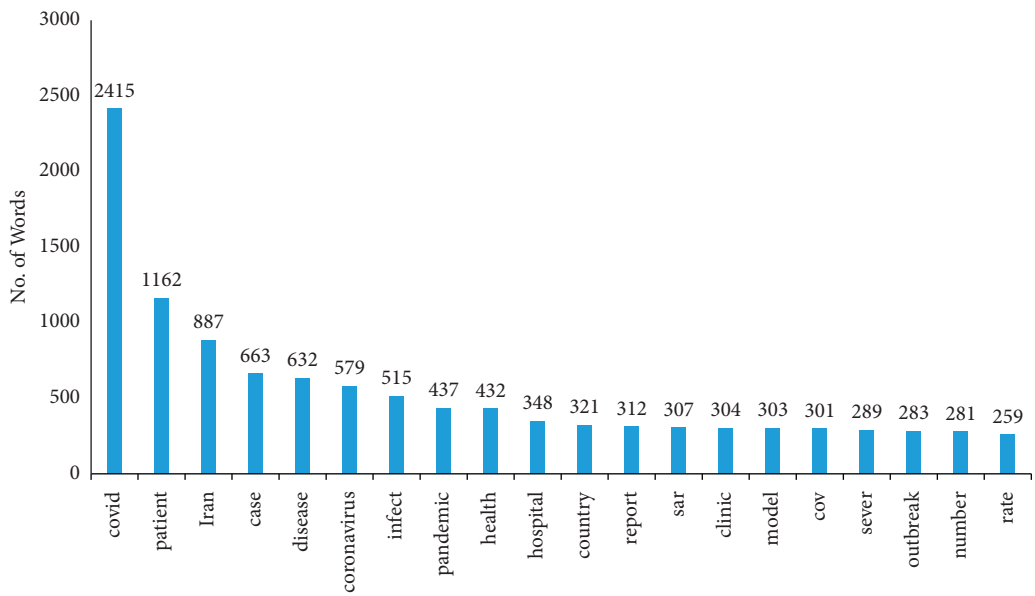


Figure 4: The words with the highest frequencies in Iranian publications on COVID-19 in LitCovid.

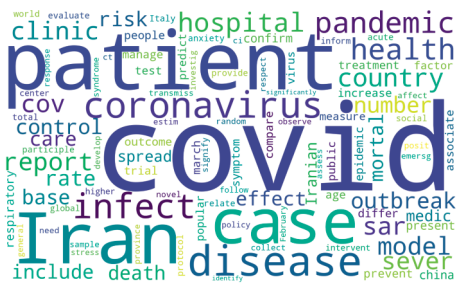


Figure 5: The word cloud with the highest frequency in Iranian publications on COVID-19 in LitCovid.

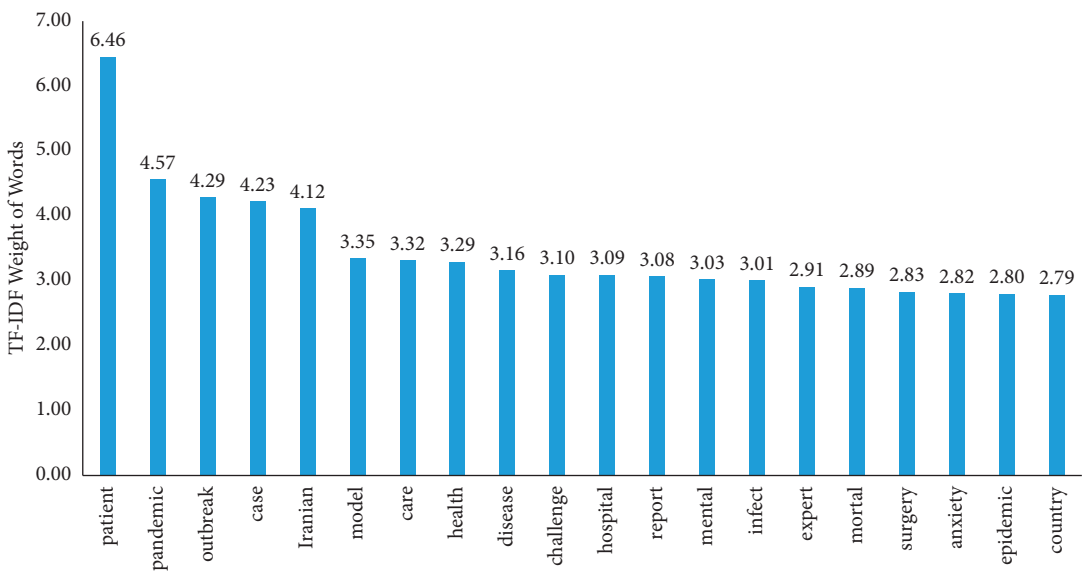


Figure 6: The most important words based on the TF-IDF weighting of Iranian publications on COVID-19 in LitCovid.

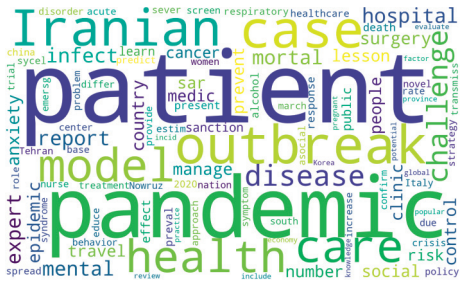


Figure 7: The word cloud of most important words based on the TF-IDF weighting of Iranian publications on COVID-19 in LitCovid.

Table 1: The topics extracted from the implementation of topic modeling algorithm on Iranian publications on Covid-19 in LitCovid.

LitCovid topic	Subtopics based topic modeling	Keywords
General	Topic1: General[1]	—
Mechanism	Topic1: Characteristics	Case, virus, people, smoking, potential, expression, disease, spread, find, level
	Topic2: Clinical features	Cell, infection, clinical, virus, case, acute, respiratory, severe, disease, lymphocyte
	Topic3: Genomic sequence	Sequence, respiratory, virus, phylogenetic, isolate, infection, mutation, strain, cause, acute
Transmission	Topic1: Environment	Rate, case, temperature, fatality, cluster, increase, spread, attack, correlation, size
	Topic2: Different areas	Transmission, disease, infection, rate, region, case, outbreak, report, measure, sample
	Topic3: Modes	Case, transmission, travel, cluster, age, early, global, report, individual, infection
Diagnosis	Topic1: Infection	Infection, case, disease, risk, chest, severe, clinical, symptom, evaluate, report
	Topic2: Risk factors	Mortality, clinical, outcome, risk, disease, diabetes, hospital, age, death, factor
	Topic3: Symptoms	Case, symptom, age, clinical, mortality, disease, positive, infection, test, confirm
Treatment	Topic1: Clinical features of mortality	Clinical, mortality, case, risk, hospital, factor, value, confirm, report, age
	Topic2: Clinical features of the disease	Disease, clinical, case, mortality, outcome, severe, infection, risk, age, rate
	Topic3: Drug	Control, trial, treatment, clinical, cell, drug, vitamin, receive, level, disease
	Topic4: Outcome	Trial, clinical, treatment, outcome, control, intervention, participant, arm, randomization, registration
Prevention	Topic1: Behaviors	Pandemic, health, disease, care, risk, control, outbreak, dental, infection, intervention
	Topic2: Management	Pandemic, hospital, health, stroke, knowledge, disease, management, base, care, infection
	Topic3: Policy	Pandemic, public, health, people, infection, policy, risk, disease, model, case
	Topic4: Control	Case, pandemic, disease, risk, control, spread, outbreak, infection, epidemic, people
	Topic5: Other diseases	Pandemic, outbreak, health, infection, report, care, trauma, increase, case, review
	Topic6: Pandemic status	Case, model, rate, death, spread, pandemic, measure, confirm, mortality, estimate
Case report	Topic1: Children	Case, clinical, infection, acute, test, child, report, syndrome, respiratory, neonate
	Topic2: New symptoms	Case, report, liver, manifestation, infection, symptom, clinical, cutaneous, disease, respiratory
	Topic3: Pregnant	Report, infection, case, old, pregnant woman, chest, treatment, refer, pandemic, hospitalize
	Topic4: Death	Report, case, disease, respiratory, acute, novel, rate, death, fetal, severe
Forecasting	Topic1: Estimate	Case, estimate, model, epidemic, death, predict, confirm, trend, outbreak, base
	Topic2: Modeling	Model, case, spread, rate, confirm, death, predict, pandemic, prediction, trend
	Topic3: Epidemic	Model, case, estimate, trend, base, policy, disease, infection, epidemic, prediction
	Topic4: Spread	Case, base, pandemic, spread, region, weather, mean, humidity, temperature, estimate

respectively. Moreover, the topic area of "transmission" includes three topics, "different areas," "modes," and "environment, with publication percentages of 61.11%, 27.78%, and 11.11%, respectively. The topic area of "diagnosis" also involved three topics, "infection," "risk factors," and "symptoms," with publication percentages of 29.55%, 32.95%, and 37.50%, respectively. The fourth topic area was "treatment," which included four topics "clinical features of mortality," "clinical features of disease," "outcome," and "drug," with publication percentages of 29.79%, 25.53%, 29.79%, and 14.89%, respectively. Moreover, the topic area of "prevention" had six topics: pandemic status (22.47%), management (20.26%), policy (16.74%), control (16.30%), behaviors (12.78%), and other diseases (11.45%). The "case report" was another area of eight topics, which involved four topics: new symptoms (31.58%), children (23.68%), death

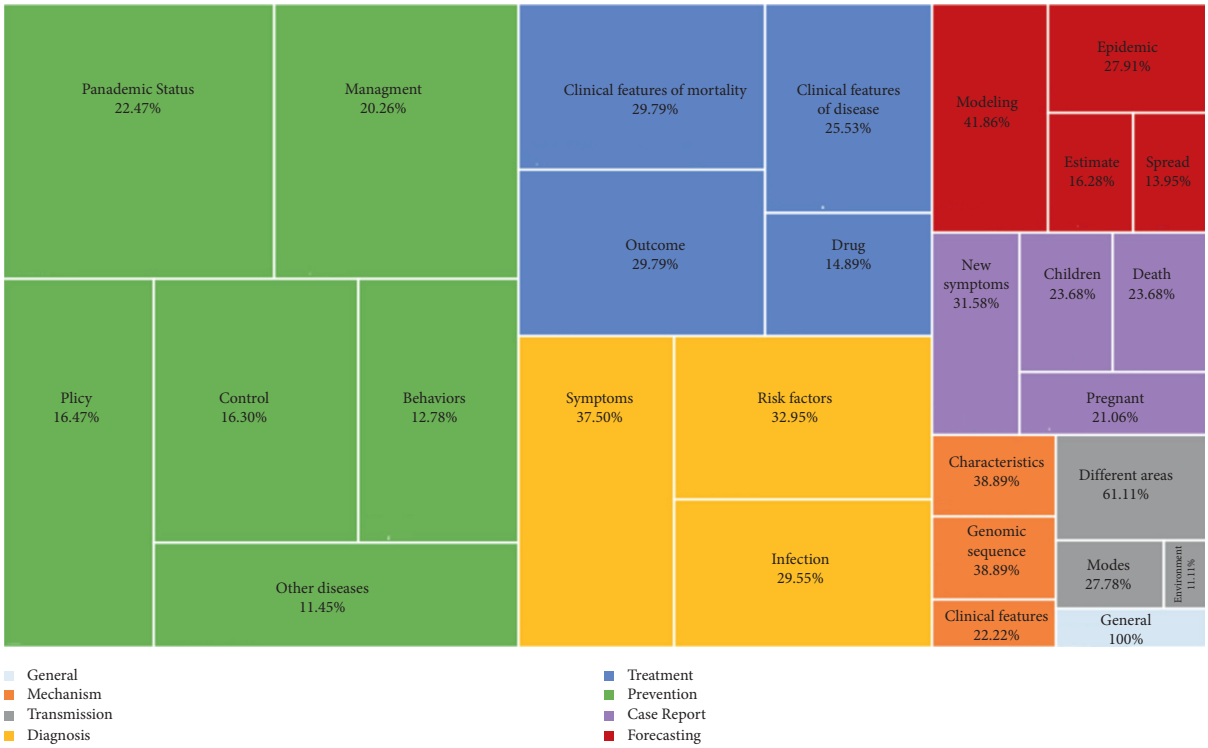


Figure 8: The percentages of Iranian COVID-19 publications in the eight-topic division of LitCovid and the topics extracted from text mining.

(23.68%), and pregnant (21.06%). The topic category of "forecasting" also included four topics: modeling (41.86%), epidemic (27.91%), estimate (16.28%), and spread (13.95%). The topic area "general" also had the lowest percentage of publications and addressed only one topic. The highest number of Iranian publications in LitCovid was related to the topic "pandemic status" in the "prevention" category, and the lowest number of publications was related to the topic "environment" in the "transmission" category.

5. Discussion

The management strategies for the health crisis that resulted from the COVID-19 pandemic are not the same in different countries, and unprecedented efforts have been made to fight the disease in all countries [45]. The global scientific community has responded to this crisis by applying all research resources and capacities to identify virus characteristics, mechanisms of its transmission, clinical aspects of the disease, and prevention and management strategies; so that the number of articles related to this crisis is being increased rapidly [46]. According to an increase in scientific publications, research centers have provided researchers different databases and software to immediately access and follow the trend of these publications; LitCovid is one of the most important and reputable databases [47].

LitCovid is an up-to-date set of all COVID-19 articles indexed in PubMed and categorizes COVID-19 publications into eight topic areas. The results of the present study indicated that the largest percentage of global and Iranian publications in LitCovid is related to the topic categories "prevention," "treatment," and "diagnosis," respectively. The lowest global and Iranian publications correspond to "forecasting" and "general," respectively.

Scientific research also plays an essential role in controlling and preventing diseases, especially epidemics. The research findings and results significantly affect the identification of the virus variants, vaccine production, treatment protocols, preventive measures, and new medicines [48].

The most important words of Iranian publications on COVID-19 in LitCovid have been identified in the present study using text mining techniques. Moreover, the subtopics of these publications have been specified in each of the eight categories of LitCovid.

The results indicated that COVID, patient, Iran, case, disease, coronavirus, infect, pandemic, health, and hospital were the most important words, in terms of frequency, that have been used in Iranian publications in LitCovid. The words, patient, pandemic, outbreak, case, Iranian, model, care, health, coronavirus, and disease were the most significant words used in Iranian publications in LitCovid regarding TF-IDF weighting. In a study conducted by Doanvo et al., COVID, patient, pandemic, coronavirus, and case were the most frequently used words in COVID-19 publications in the CORD-19; this study is in line with the present investigation [35]. Moreover, in an article published by Cheng et al., data analysis indicated that the words, patient, case, number, infection, and study had the most repetitions in COVID-19 publications [36]. Dong et al. reported that infection, cell, protein, diseases, and patient were the most frequently applied words in COVID-19 publications [49]. In addition, the results of Hossain's

research indicated that the keywords used in the COVID-19 publications showed the complexity and extent of this scientific field, which includes various disciplines such as virology, microbiology, infectious diseases, clinical medicine, public health, allied health sciences, social sciences, and other branches of knowledge [50].

Furthermore, in the present paper, the topics of each of the eight-topic categories of Iranian publications in LitCovid were identified using the LDA topic modeling algorithm. "Prevention," including status, management, policy, control, behaviors, and other diseases, involved some studies in the fields of control, response, and management strategies and had the largest percentage of Iranian publications in LitCovid.

Considering the nature of prevention studies, which is an important issue in the current pandemic situation, most studies and research projects are expected to be conducted in this regard. Colavizza concluded that the prevalence of SARS in 2003 had been associated with an increase in coronavirus and epidemic management [51]. Haghani et al. also believes that safety issues, such as physical safety and mental health of patients, constitute a large volume of the published knowledge in the field of COVID-19 [52]. In line with previous studies, Haleem et al. also believed that extensive research was needed to develop a vaccine against coronavirus infection. Besides, there was an urgent need to develop essential items and new therapies to cope with this disease [53].

Moreover, it was shown in some investigations that the countries with lower per capita income have more of a tendency toward the discovery and usage of disease prevention methods to prevent the spread of COVID-19. After "prevention," "treatment" was the next category with the highest percentage of Iranian publications in LitCovid. This topic area included the studies on treatment strategies, therapeutic procedures, and vaccine development, and involved the topics "clinical features of mortality," "outcome," "clinical features of patients," and "drug."

With the emergence of the COVID-19 pandemic, its rapid geographical spread (more than 200 countries), and the announcement of its prevalence as a public health emergency and international concern by the WHO, there is an urgent need for new diagnoses, vaccines, and treatments for this new and serious threat that has been affecting human life for more than one year. However, considering limited knowledge about COVID-19, there is no comprehensive prevention and treatment strategy, and all medical scientists around the world are working on new strategies to reduce virus infections, outbreaks, and the rate of mortality in high-risk environments. However, countries are implementing some measures to prevent the spread of the disease and create more time to produce appropriate vaccines and treatments. Moreover, pharmacotherapy options are consistently used in clinical trials to reduce the severity of infection, morbidity, and mortality [54]. Different drugs and treatments are currently being used to reduce the severity of the disease, and extensive research is underway for COVID-19 treatment [55]. Research on the development of COVID-19 vaccines and discovering new methods to improve human safety against COVID-19 are the most important and cited topics under investigation since the discovery of the vaccine would enhance immunity and protection of people against the COVID-19 virus. They would also be effective in reducing the anxiety of the common populace.

"Diagnosis" was also ranked third of the highest percentage of Iranian publications on COVID-19, which involved studies on disease assessment through symptoms, test results, and radiological features. "Symptoms," "risk factors," and "infection" are the topics extracted from text mining in the present study.

In the diagnosis of COVID-19, different symptoms appear in the patient ranging from asymptomatic and mild symptomatic to severe symptomatic infections. Pneumonia is the most common serious manifestation of infection, characterized by fever, cough, shortness of breath, and bilateral pulmonary infiltrates on a chest X-ray [56]. Moreover, the results regarding the severity of the symptoms of this disease are scattered and incomplete. Besides, unexplained or unusual symptoms make COVID-19 challenging to diagnose and complicate appropriate treatment to patients. Lack of vaccines and effective treatment protocols increases the importance of early and definitive diagnosis of this disease [57].

The "forecasting" category includes some studies related to modeling and forecasting the trend of COVID-19 spread and is in the third rank of the highest percentage of publications. This category involves the topics "modeling," "epidemic," "estimate," and "region," respectively, from the highest to the lowest number of publications. Predicting the prevalence trend of this disease can help health authorities determine the characteristics of the virus transmission and develop appropriate strategies for preventing and controlling the disease. Researchers have applied conventional epidemic models, such as Susceptible-Exposed-Infective-Recovery (SEIR), or machine-learning models, such as logistic regression, to predict these trends in COVID-19 disease [58]. The "case report" category included studies in descriptions of specific patient cases, which is in the next rank in terms of the highest number of publications, and included the topics "new symptoms," "children," "other," and "pregnant."

The "children" play critical role in cases reported in the publications, and they are of particular importance in this disease; the course of this disease in children is mild, and infection caused by this virus in children has a better prognosis. Moreover, due to the mild clinical symptoms in children, many are not diagnosed in the early stages of COVID-19. Since this infection can be transmitted to others, children can play an essential role in transmitting this virus in the family, even from asymptomatic cases. In other words, children can lead to infection clusters in the home environment [59]. Given the predominance of gastrointestinal symptoms in infants and children, the spread of the virus through the feces continues for several weeks after diagnosis in children, and thus COVID-19 can spread in kindergartens or elementary schools [60]. Pregnant women are one of the most vulnerable groups and are prone to infectious diseases due to weakened immune systems, and their infection and

the risk of transmission to the fetus has become a significant concern [61].

The "transmission" category involved studies on characteristics and modes of COVID-19 transmissions, such as human-to-human type. This category included the topics "different areas," "modes," and "environment," respectively, from the highest to the lowest number of publications. The "mechanism" category included studies on the underlying cause(s) of COVID-19 infections and the transmission and possible drug mechanism of action. In this category, "characteristics," "genomic sequence," and "clinical features" are the topics extracted from text mining. The result of the topic modeling indicated that "general" is one of the eight Categories. General information on the web and news on websites are the essential sub-topics. Similar results were obtained from the analysis of the scientific texts and publications on Coronavirus and COVID-19.

Danesh et al. identified the topics of coronavirus scientific publications in the last fifty years using text mining techniques and topic modeling algorithms. The results of their study provided eight topics for the global publications on coronavirus. These topics were "structure and proteomics," "cell signaling and immune response," "clinical presentation and detection," "gene sequence and genomics," "diagnosis tests," "vaccine and immune response and outbreak," "epidemiology and transmission," and "gastrointestinal tissue" [62].

Continuing previous research, Colavizza indicated that the topics in the CORD-19 focus on specific topics such as coronaviruses (SARS, MERS, and COVID-19), public health and epidemics, molecular biology, influenza and the family of viruses, immunology and antiviruses, and methodology (test, diagnosis, and clinical trials) [51]. Another study in line with the previous investigation showed that the publications in the field of coronavirus have initially focused on public health and epidemic control; the chemical structure of the virus; and the studies related to treatment, vaccines, and clinical care [63].

Dehghanbanadaki believed that COVID-19 researchers focus on various aspects of this infection, such as pathogenesis, epidemiology, transmission, diagnosis, treatment, prevention, and complications [64]. Pal also showed that studies on COVID-19 had been published in an extensive range of disciplines such as medicine, biochemistry, molecular biology, immunology, microbiology, social sciences, nursing, pharmacology, neuroscience, environmental sciences, health care, and multidisciplinary [31]. In another article compiled by the topic modeling method, COVID-19 publications were categorized into eight topics: clinical characterization, pathogenesis research, therapeutics research, epidemiological study, virus transmission, vaccine research, virus diagnostics, and viral genomics [49]. In another investigation entitled "Text mining of global COVID-19 publications," the results indicated that the focus in global COVID-19 publications is on clinical management, viral pathogenesis, and public health responses, and little attention has been paid to psychosocial problems or the impacts of COVID-19 on different vulnerable populations [33].

A study carried out by Älgå et al. during the first six months of the COVID-19 epidemic concluded that health care response, clinical manifestations, and psychological impact were the most important published topics [37]. Moreover, Amiri et al. identified three topic clusters of health research, basic science research, and clinical research for COVID-19 publications in Scopus. In the health research cluster, the epidemiological aspects of the disease, public health, and prevention and control of infection were the main focus of the research. The main focus was on the virus's virological, immunological, and genetic aspects in the basic science research cluster. Finally, clinical signs of disease, treatment methods, and diagnostic imaging were the primary focus in the clinical research cluster [8]. In another study conducted "discovering associations in COVID-19 relatedresearch paper," the authors used the text mining method on CORD-19 data. They concluded that, the scientific publications related to various aspects of coronavirus were distinguished with different types of virus (e.g., RNA), clinical manifestations (e.g., pneumonia), consequences (e.g., quarantine), acquaintance (e.g., H7N9), and virus description (e.g., pathogen) [65].

6. Conclusion and Suggestions for Future

The findings of the present article have raised structural viewpoints of COVID-19 documents in Iran to guide and provide practical solutions to researchers, planners, and policymakers, and demonstrated different aspects of Iranian publications on COVID-19. A suggestion is that in an independent study, text mining method to be applied to all LitCovid publications, the topics of global publications in each of the eight topic areas to be obtained using topic modeling algorithms, and the results of text mining of global publications are to be compared with the present paper.

We suggested that researchers identify the most significant countries and institutions in the publication of each category in LitCovid. According to the limitations in LitCovid, it is also suggested that LitCovid designers design a mechanism wherein in the section of bibliographic data download, in addition to article titles, journal names, and PMID, other bibliographic data including the authors' names, partner countries, organizational affiliation, publisher as well as the exact publishing date are downloadable and exportable to the statistical, text mining, and scientometrics software.

Appendix

Data Availability

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

Disclosure

The present article was extracted from a research project with code A-10-1263-5 and research ethics ID IR.GMU.REC.1400.002, approved by the Infectious Disease

Research Center and implemented with the financial support of this research center.

Conflicts of Interest

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

References

[1] L. S. Wang, Y. R. Wang, D. W. Ye, and Q. Q. Liu, "A review of the 2019 Novel Coronavirus (COVID-19) based on current evidence," *International Journal of Antimicrobial Agents*, vol. 55, no. 6, 2020.

[2] M. Cascella, M. Rajnik, A. Cuomo, S. C. Dulebohn, and R. Di Napoli, *Features, evaluation and treatment coronavirus (COVID-19)*StatPearls, treasure island, FL, USA, 2020.

[3] A. Shamsi, M. J. Mansourzadeh, A. Ghazbani, K. Khalagi, N. Fahimfar, and A. Ostovar, "Contribution of Iran in COVID-19 studies: a bibliometrics analysis," *Journal of Diabetes and Metabolic Disorders*, vol. 19, no. 2, pp. 1–10, 2020.

[4] P. Song and T. Karako, "COVID-19: real-time dissemination of scientific information to fight a public health emergency of international concern," *Bioscience trends*, vol. 14, no. 1, pp. 1-2, 2020.

[5] J. A. Teixeira da Silva, P. Tsigaris, and M. Erfanmanesh, "Publishing volumes in major databases related to COVID-19," *Scientometrics*, vol. 126, no. 1, pp. 831–842, 2021.

[6] F. Danesh, S. Ghavidel, and S. Ghavidel, "Coronavirus: scientometrics of 50 Years of global scientific productions," *Iranian Journal of Medical Microbiology*, vol. 14, no. 1, pp. 1–16, 2020.

[7] B. Harsanto, "The first-three-month review of research on COVID-19: a scientometrics analysis," in *Proccedings of the 2020 IEEE International Conference on Engineering, Technology and Innovation (ICE/ITMC)*, pp. 1–6, IEEE, Cardiff, UK, June 2020.

[8] M. Amiri, N. Taha, and P. Mahdizadeh, "Subjects clustering analysis and science mapping on COVID-19 researches in scopus database," *Journal Mil Med*, vol. 22, no. 6, pp. 663–669, 2020.

[9] X. Liu, Yu Shi, F. Janssens, W. Glänzel, Y. Moreau, and B. De Moor, "Weighted hybrid clustering by combining text mining and bibliometrics on a large-scale journal database," *Journal of the American Society for Information Science and Technology*, vol. 61, no. 6, pp. 1105–1119, 2010.

[10] Y. Zhang, A. L. Porter, and D. Chiavetta, "Scientometrics for tech mining: an introduction," *Scientometrics*, vol. 111, no. 3, pp. 1875–1878, 2017.

[11] A. K. Choudhary, P. I. Oluikpe, J. A. Harding, and P. M. Carrillo, "The needs and benefits of text mining applications on post-project reviews," *Computers in Industry*, vol. 60, no. 9, pp. 728–740, 2009.

[12] K. W. Boyack, D. Newman, R. J. Duhon et al., "Clustering more than two million biomedical publications: comparing the accuracies of nine text-based similarity approaches," *PloS One*, vol. 6, no. 3, Article ID 18029, 2011.

[13] J. Feng, X. Mu, W. Wang, and Y. Xu, "A topic analysis method based on a three-dimensional strategic diagram," *Journal of Information Science*, 2020.

[14] B. Nie and S. Sun, "Using text mining techniques to identify research trends: a case study of design research," *Applied Sciences*, vol. 7, no. 4, p. 401, 2017.

[15] J. R. Rogers, H. Mills, L. V. Grossman, A. Goldstein, and C. Weng, "Understanding the nature and scope of clinical research commentaries in PubMed," *Journal of the American Medical Informatics Association*, vol. 27, no. 3, pp. 449–456, 2020.

[16] D. Abramson, M. Lees, V. V. Krzhizhanovskaya, J. Dongarra, and P. M. A. Sloot, "Big Data Meets Computational Science, Preface for ICCS2014," *Procedia Computer Science*, vol. 29, pp. 1–7, 2014.

[17] D. M. Blei, "Probabilistic topic models," *Communications of the ACM*, vol. 55, no. 4, pp. 77–84, 2012.

[18] X. Cheng, C. Shuai, J. Liu et al., "Topic modelling of ecology, environment and poverty nexus: an integrated framework," *Agriculture, Ecosystems & Environment*, vol. 267, pp. 1–14, 2018.

[19] H. Jelodar, Y. Wang, C. Yuan et al., "Latent Dirichlet allocation (LDA) and topic modeling: models, applications, a survey," *Multimedia Tools and Applications*, vol. 78, no. 11, pp. 15169–15211, 2019.

[20] Q. Chen, A. Allot, and Z. Lu, "LitCovid: an open database of COVID-19 literature," *Nucleic Acids Research*, vol. 49, no. D1, pp. D1534–D1540, 2021.

[21] Q. Chen, A. Allot, and Z. Lu, "Keep up with the latest coronavirus research," *Nature*, vol. 579, no. 7798, p. 193, 2020.

[22] M. Coccia, "The impact of first and second wave of the COVID-19 pandemic in society: comparative analysis to support control measures to cope with negative effects of future infectious diseases," *Environmental Research*, vol. 197, Article ID 111099, 2021.

[23] P. Howitt, A. Darzi, G.-Z. Yang et al., "Technologies for global health," *The Lancet*, vol. 380, no. 9840, pp. 507–535, 2012.

[24] R. Sperandeo, G. Messina, D. Iennaco et al., "What does personality mean in the context of mental health? A topic modeling approach based on abstracts published in PubMed over the last 5 years," *Frontiers in Psychiatry*, vol. 10, p. 938, 2020.

[25] X. Barn, L. H. Nguyen, H. C. Turner et al., "Economic evaluation studies in the field of HIV/AIDS: bibliometric analysis on research development and scopes (GAPRESEARCH)," *BMC Health Services Research*, vol. 19, no. 1, pp. 834–912, 2019.

[26] B. X. Tran, H. T. Phan, Q. N. Nguyen et al., "Pre-exposure prophylaxis in HIV research: a latent dirichlet allocation analysis (gapresearch)," *AIDS Reviews*, vol. 22, no. 2, pp. 103–111, 2020.

[27] T. Saheb and M. Saheb, "Analyzing and visualizing knowledge structures of health informatics from 1974 to 2018: a bibliometric and social network analysis," *Healthcare informatics research*, vol. 25, no. 2, p. 61, 2019.

[28] Y.-M. Kim and D. Delen, "Medical informatics research trend analysis: a text mining approach," *Health Informatics Journal*, vol. 24, no. 4, pp. 432–452, 2018.

[29] M. S. Jalali, S. Razak, W. Gordon, E. Perakslis, and S. Madnick, "Health care and cybersecurity: bibliometric analysis of the literature," *Journal of Medical Internet Research*, vol. 21, no. 2, Article ID 12644, 2019.

[30] F. Danesh, S. Ghavidel, and V. Piranfar, "Coronavirus: discover the structure of global knowledge, hidden patterns and emerging events," *Journal of Advances in Medical and Biomedical Research*, vol. 28, no. 130, pp. 253–264, 2020.

[31] J. K. Pal, "Visualizing the knowledge outburst in global research on COVID-19," *Scientometrics*, vol. 126, no. 5, pp. 4173–4193, 2021.

[32] S. Belli, R. Mugnaini, J. Baltà, and E. Abadal, "Coronavirus mapping in scientific publications: when science advances rapidly and collectively, is access to this knowledge open to society?" *Scientometrics*, vol. 124, no. 3, pp. 2661–2685, 2020.

[33] B. X. Tran, G. H. Ha, L. H. Nguyen et al., "Studies of novel coronavirus disease 19 (COVID-19) pandemic: a global analysis of literature," *International Journal of Environmental Research and Public Health*, vol. 17, no. 11, p. 4095, 2020.

[34] P. Radanliev, D. De Roure, and R. Walton, "Data mining and analysis of scientific research data records on COVID-19 mortality, immunity, and vaccine development - in the first wave of the COVID-19 pandemic," *Diabetes & Metabolic Syndrome: Clinical Research Reviews*, vol. 14, no. 5, pp. 1121–1132, 2020.

[35] A. Doanvo, X. Qian, D. Ramjee, H. Piontkivska, A. Desai, and M. Majumder, "Machine learning maps research needs in COVID-19 literature," *Patterns*, vol. 1, no. 9, Article ID 100123, 2020.

[36] X. Cheng, Q. Cao, and S. L. Stephen, "An overview of the literature on cOVID-19, MERS, and SARS: using text mining and latent dirichlet allocation," *Journal of Information Science*, 2020.

[37] A. Älgå, O. Eriksson, and M. Nordberg, "Analysis of scientific publications during the early phase of the COVID-19 pandemic: topic modeling study," *Journal of Medical Internet Research*, vol. 22, no. 11, Article ID 21559, 2020.

[38] J. Žižka, F. Dařena, and A. Svoboda, *Text Mining with Machine Learning: Principles and Techniques*, CRC Press, Boca Raton, Fl, USA, 2019.

[39] W. B. Frakes, *Information Retrieval: Data Structures & Algorithms*, Prentice-Hall, Upper Saddle River, NJ, USA, 2004.

[40] T. M. Abuhay, S. V. Kovalchuk, K. O. Bochenina et al., "Analysis of computational science papers from iccs 2001-2016 using topic modeling and graph theory," *Procedia Computer Science*, vol. 108, pp. 7–17, 2017.

[41] H. Schütze, C. D. Manning, and P. Raghavan, *Introduction to information retrieval*, Vol. 39, Cambridge University Press, Cambridge, UK, 2008.

[42] W. Wang, Y. Feng, and W. Dai, "Topic analysis of online reviews for two competitive products using latent Dirichlet allocation," *Electronic Commerce Research and Applications*, vol. 29, pp. 142–156, 2018.

[43] A. Ebadi, P. Xi, S. Tremblay, B. Spencer, R. Pall, and A. Wong, "Understanding the temporal evolution of COVID-19 research through machine learning and natural language processing," *Scientometrics*, vol. 126, no. 1, pp. 725–739, 2021.

[44] R. Radim and S. Petr, "Software framework for topic modelling with large corpora," in *Proceedings of the LREC 2010 workshop on new challenges for NLP frameworks*, Valletta, Malta, May 2010.

[45] J. Lau, "Coronavirus crisis inspiring unprecedented global research effort," *Higher Education*, vol. 25, 2020.

[46] A. R. A. H. Hamid, "Social responsibility of medical journal: a concern for COVID-19 pandemic," *Medical Journal of Indonesia*, vol. 29, no. 1, pp. 1–3, 2020.

[47] L. Škorić, A. Glasnović, and J. Petrak, "A publishing pandemic during the COVID-19 pandemic: how challenging can it become?" *Croatian Medical Journal*, vol. 61, no. 2, p. 79, 2020.

[48] J. Lou, Si-J. Tian, Sheg-mei niu et al., "Coronavirus disease 2019: a bibliometric analysis and review," *European Review for Medical and Pharmacological Sciences*, vol. 24, no. 6, pp. 3411–3421, 2020.

[49] M. Dong, X. Cao, M. Liang, L. Li, H. Liang, and G. Liu, "Understand research hotspots surrounding COVID-19 and other coronavirus infections using topic modeling," *medRxiv*, vol. 23, 2020.

[50] M. M. Hossain, "Current status of global research on novel coronavirus disease (COVID-19): a bibliometric analysis and knowledge mapping," *A Bibliometric Analysis and Knowledge Mapping*, vol. 9, p. 374, 2020.

[51] G. Colavizza, R. Costas, V. A. Traag, N. J. Van Eck, T. van Leeuwen, and L. Waltman, "A scientometric overview of CORD-19," *PloS One*, vol. 16, no. 1, Article ID 0244839, 2021.

[52] M. Haghani, M. C. J. Bliemer, F. Goerlandt, and J. Li, "The scientific literature on Coronaviruses, COVID-19 and its associated safety-related research dimensions: a scientometric analysis and scoping review," *Safety Science*, vol. 129, Article ID 104806, 2020.

[53] A. Haleem, M. Javaid, R. Vaishya, and S. G. Deshmukh, "Areas of academic research with the impact of COVID-19," *The American Journal of Emergency Medicine*, vol. 38, no. 7, pp. 1524–1526, 2020.

[54] V. K. Singh, A. Mishra, S. Singh et al., "Emerging prevention and treatment strategies to control COVID-19," *Pathogens*, vol. 9, no. 6, p. 501, 2020.

[55] Y.-J. Han, Z.-G. Ren, X.-X. Li et al., "Advances and challenges in the prevention and treatment of COVID-19," *International Journal of Medical Sciences*, vol. 17, no. 12, pp. 1803–1810, 2020.

[56] C. Jiehao, X. Jin, L. Daojiong et al., "A case series of children with 2019 novel coronavirus infection: clinical and epidemiological features," *Clinical Infectious Diseases*, vol. 71, no. 6, pp. 1547–1551, 2020.

[57] J. Jeon, G. Baruah, S. Sarabadani, and A. Palanica, "Identification of risk factors and symptoms of COVID-19: analysis of biomedical literature and social media data," *Journal of Medical Internet Research*, vol. 22, no. 10, Article ID 20509, 2020.

[58] J. Sun, X. Chen, Z. Zhang et al., "Forecasting the long-term trend of COVID-19 epidemic using a dynamic model," *Scientific Reports*, vol. 10, no. 1, Article ID 12112, 2020.

[59] C.-C. Lai, T.-P. Shih, W.-C. Ko, H.-J. Tang, and P.-R. Hsueh, "Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) and coronavirus disease-2019 (COVID-19): the epidemic and the challenges," *International Journal of Antimicrobial Agents*, vol. 55, no. 3, Article ID 105924, 2020.

[60] P. Zimmermann and N. Curtis, "Coronavirus infections in children including COVID-19," *The Pediatric Infectious Disease Journal*, vol. 39, no. 5, pp. 355–368, 2020.

[61] Y. Luo and K. Yin, "Management of pregnant women infected with COVID-19," *The Lancet Infectious Diseases*, vol. 20, no. 5, pp. 513-514, 2020.

[62] F. Danesh, M. Dastani, and M. Ghorbani, *Retrospective and Prospective Approaches of Coronavirus Publications in the Last half-century: A Latent Dirichlet Allocation Analysis*, Library Hi Tech, hongkong, China, 2021.

[63] M. Haghani and M. C. J. Bliemer, "COVID-19 pandemic and the unprecedented mobilisation of scholarly efforts prompted by a health crisis: scientometric comparisons across SARS, MERS and 2019-nCoV literature," *Scientometrics*, vol. 125, no. 3, pp. 2695–2726, 2020.

[64] H. Dehghanbanadaki, F. Seif, Y. Vahidi et al., "Bibliometric analysis of global scientific research on coronavirus (COVID-19)," *Medical Journal of the Islamic Republic of Iran*, vol. 34, p. 51, 2020.

[65] I. Fister, K. Fister, and I. Fister, "Discovering associations in COVID-19 related research papers," 2020, https://arxiv.org/abs/2004.03397.

Hindawi
Scientific Programming
Volume 2021, Article ID 5540024, 10 pages
https://doi.org/10.1155/2021/5540024

Research Article

Cervical Cancer Diagnosis Model Using Extreme Gradient Boosting and Bioinspired Firefly Optimization

Irfan Ullah Khan, Nida Aslam, Rawan Alshehri, Seham Alzahrani, Manal Alghamdi, Atheer Almalki, and Maryam Balabeed

Department of Computer Science, College of Computer Science and Information Technology, Imam Abdulrahman Bin Faisal University, Dammam 31441, Saudi Arabia

Correspondence should be addressed to Nida Aslam; naslam@iau.edu.sa

Received 24 February 2021; Accepted 27 June 2021; Published 19 July 2021

Academic Editor: Antonio J. Peña

Cervical cancer is frequently a deadly disease, common in females. However, early diagnosis of cervical cancer can reduce the mortality rate and other associated complications. Cervical cancer risk factors can aid the early diagnosis. For better diagnosis accuracy, we proposed a study for early diagnosis of cervical cancer using reduced risk feature set and three ensemble-based classification techniques, i.e., extreme Gradient Boosting (XGBoost), AdaBoost, and Random Forest (RF) along with Firefly algorithm for optimization. Synthetic Minority Oversampling Technique (SMOTE) data sampling technique was used to alleviate the data imbalance problem. Cervical cancer Risk Factors data set, containing 32 risks factor and four targets (Hinselmann, Schiller, Cytology, and Biopsy), is used in the study. The four targets are the widely used diagnosis test for cervical cancer. The effectiveness of the proposed study is evaluated in terms of accuracy, sensitivity, specificity, positive predictive accuracy (PPA), and negative predictive accuracy (NPA). Moreover, Firefly features selection technique was used to achieve better results with the reduced number of features. Experimental results reveal the significance of the proposed model and achieved the highest outcome for Hinselmann test when compared with other three diagnostic tests. Furthermore, the reduction in the number of features has enhanced the outcomes. Additionally, the performance of the proposed models is noticeable in terms of accuracy when compared with other benchmark studies for cervical cancer diagnosis using reduced risk factors data set.

1. Introduction

Cervical cancer is one of the commonly occurring types of cancer in females and mostly develops during their midlives (35 years–44 years) [1]. This type of cancer can be fatal as it does not show clear symptoms in its early stages. Symptoms usually appear in late stages, where it could have spread to other organs like bones, liver, lymph nodes, and lungs. One of the early signs of cervical cancer is when the tube that carries urine from the kidney is blocked. Other late symptoms that can appear are vaginal bleeding, pelvic pain, weight loss, and leg pain [2].

The risk factors that lead to the development of cervical cancer are hormones containing medicines, birth control pills, smoking, and the number of pregnancies. However, it is believed that human papilloma virus (HPV) is the major factor in developing cervical cancer [2]. HPV is a common sexually transmitted infection; it is usually harmless, but sometimes it may lead to cancer [3]. HPV infection becomes at a higher risk of getting cervical cancer. Furthermore, the probability of getting cervical cancer increases if one possesses more than one risk factor. As the cancer does not show signs in its early stages, regular checkups are required especially for those who have the risk factors. In the developing countries, lack of medical equipment and the cost of conducting checkups could also be a burden. With the advent and advancement of machine learning, it has become possible to find robust solutions for early diagnosis of cancer cases using data-driven approaches.

Various studies have contributed to the field of cervical cancer diagnosis using several classification techniques by

using different types of data such as clinical-based, image, and genetic-based data. In our study, we used clinical cervical risk factor data. Two similar studies were conducted by Wu and Zhou [4] and Abdoh et al. [5]; they performed the comparative analysis of two feature selection techniques, namely, recursive feature elimination (RFE) and Principal Component Analysis (PCA). The first study used Support Vector Machine (SVM), and the other study used Random Forest (RF). Both studies used the same number of features. Although the data suffered from imbalance, an oversampling was applied to the data in [4] and SMOTE was used in [5]. Both studies identified two risk factors to be removed such as time since the first and last diagnosis of STDs (sexually transmitted diseases), due to a lot of missing entries. Furthermore, the study [4] discovered that less computational cost was an advantage given by both SVM-PCA and SVM-RFE, whereas high computational cost is a limitation to the SVM model. Moreover, STDs, intrauterine device (IUD), hormonal contraceptives, and first sexual intercourse were identified as the highly relevant features [5]. Overall, the outcome of both the studies showed that using 30 features produced highest results. Furthermore, it was found that the SMOTE-RF model performed well for all targets.

Similarly, Lu et al. [6] and Karim and Neehal [7] used ensemble models to estimate the risk of cervical cancer. Both studies performed data cleaning mechanism to replace missing values. The former study used an ensemble classifier with voting strategy using a combination of a private and public data set. The private data set contains 472 records taken from Chinese hospital. The public data set was obtained from the UCI repository; 14 features were used. The private data set was collected using questionnaire. The results revealed that voting ensemble classifier produced better results when compared to Linear Regression, Decision Tree (DT), Multilayer Perceptron (MLP), SVM, and K-NN classifiers. On the other hand, Karim and Neehal study used DT, MLP, and SVM using sequential Minimal Optimization (SMO) and K nearest neighbor (KNN) techniques. Experiments showed that SMO has a better performance in terms of accuracy, precision, recall, and F-measure. Similarly, Ul-Islam et al. [8] used DT, RF, Logistic Model Tree, and ANN for cervical cancer detection. Apriori algorithm was used to identify features that strongly relate to cancer. The study found that age, number of sexual partners, hormonal contraceptives, number of pregnancies, and first sexual intercourse are significant risk factors. Results indicated that RF produced best outcome when compared to the other models.

Al-Wesabi et al. [9] conducted a comparison between different machine learning classifiers such as Gaussian Naïve Bayes (GNB), KNN, DT, LR, and SVM. The outcome of the classifiers was not satisfactory due to the data imbalance. To resolve this problem, undersampling, oversampling, and SMOTETomek were applied. Oversampling had the best result among all three methods. Moreover, a Sequential Feature Selector was applied with both forward and backward versions. Both the Sequential Forward Feature Selector (SFS) and Sequential Backward Feature Selector (SBS) enhanced the performance of the prediction with an accuracy of 95%. After selecting the common features between DT and KNN, the accuracy exceeded 97% for the DT. The results revealed that age, first sexual intercourse, number of pregnancies, smoking, hormonal contraceptives, and STDs: genital herpes were the main predictive features.

Similarly, several studies have been made using deep learning and transfer learning for cervical cancer diagnosis. Fernandes et al. [10] and Adem et al. [11] used deep learning and showed significant outcome in terms of diagnosis accuracy. The study [10] used a loss function that provides a supervised optimization of dimensionality reduction and classification models. The study indicated that it can be useful in examining records of patients if the Biopsy and perhaps other testing results are absent and are capable of classifying successfully whether they have cervical cancer or not. On the other hand, the researchers in [11] used a deep neural network model with softmax function to classify the data sets. The performance of the softmax function with stacked autoencoder was compared with the other machine learning methods (DT, KNN, SVM, Feed Forward NN, and Rotation Forest models). It was found that the softmax function with a stacked autoencoder model produced better outcome classification rate of 97.8%.

Similarly, Fernandes et al. [12] applied transfer learning with partial observability for cancer screenings. The limitation of the study was that several patients were resisting answering some questions for privacy concerns. Challenges were also faced in defining quality as there are multiple readings and it started relying on human preference. Therefore, as an alternative of an ordinal scale, a simple binary scheme was used. Nevertheless, the model performance was considerable.

Conclusively, the finding made after the above-mentioned literature is that the data set found at UCI repository had several missing values; therefore, previous studies have removed at least 2 features. Missing values were due to patient's concerns regarding their privacy. After removing 2 features due to huge missing value, SVM-PCA seemed to provide satisfactory performance. However, SMO and SMOTE-RF were amongst the best performing models. Another approach to deal with the imbalance in UCI cervical risk factor data set was using oversampling. Deep learning proved to be effective, especially where the Biopsy and possibly other screening results are absent. Age, first sexual intercourse, number of pregnancies, smoking, hormonal contraceptives, IUD, STDs, STDs: genital warts, or HPV infections were identified as the top key features. The significant outcomes made by the machine learning classifiers motivate the need for further investigation and enhancement of the outcomes for the prediction of cervical cancer.

In this study, three ensemble-based classifiers extreme Gradient Boosting, Ada Boost, and RF are used to classify cervical cancer. Cervical Cancer Risk factor data set from UCI machine learning repository was collected at "Hospital Universitario de Caracas" in Caracas, Venezuela [13]. In addition to the importance of correctly classifying cancerous and noncancerous cases, it is also essential to identify key risk factors that contribute to developing cancer. Nature-inspired Firefly feature selection and optimization algorithm

was applied. Furthermore, the Synthetic Minority Oversampling Technique (SMOTE) is used to balance the classes of the data as it suffers greatly from imbalanced problem.

The paper is organized as follows: Section 2 presents material and methods. Section 3 contains experimental setup and results. The comparison of the proposed model with the existing studies using the same dataset is discussed in Section 4. Finally, Section 5 contains the conclusion.

2. Material and Method

2.1. Dataset Description. The cervical cancer risk factors data set used in the study was collected at "Hospital Universitario de Caracas" in Caracas, Venezuela and is available on the UCI Machine Learning repository [13]. It consists of 858 records, with some missing values, as several patients did not answer some of the questions due to privacy concerns. The data set contains 32 risk factors and 4 targets, i.e., the diagnosis tests used for cervical cancer. It contains different categories of feature set such as habits, demographic information, history, and Genomic medical records. Features such as age, Dx: Cancer, Dx: CIN, Dx: HPV, and Dx features contains no missing values. Dx: CIN is a change in the walls of cervix and is commonly due to HPV infection; sometimes, it may lead to cancer if it is not treated properly. However, Dx: cancer variable is represented if the patient has other types of cancer or not. Sometimes, a patient may have more than one type of cancer. In the data set, some of the patients do not have cervical cancer, but they had the Dx: cancer value true. Therefore, it is not used as a target variable.

Table 1 presents a brief description of each feature with the type. Cervical cancer diagnosis usually requires several tests; this data contains the widely used diagnosis tests as the target. Hinselmann, Schiller, Cytology, and Biopsy are four widely used diagnosis tests for cervical cancer. Hinselmann or Colposcopy is a test that examines the inside of the vagina and cervix using a tool that magnifies the tissues to detect any anomalies [3]. Schiller is a test in which a chemical substance called iodine is applied to the cervix, where it stains healthy cells into brown color and leaves the abnormal cells uncolored, while cytology is a test that examines body cells from uterine cervix for any cancerous cells or other diseases. And Biopsy refers to the test where a small part of cervical tissue is examined under a microscope. Most Biopsy tests can make significant diagnosis.

2.2. Dataset Preprocessing. The data set suffers from a huge number of missing values; 24 features out of the 32 contained missing values. Initially, the features with the huge percentage of missing values were removed. STDs: Time since first diagnosis and STDs: Time since last diagnosis features were removed since they have 787 missing values (see Table 2), which is more than half of the data. However, the data imputation was performed for the features with fewer numbers of missing values. The most frequent value technique was used to impute the remaining missing values. Additionally, the data set also suffers from huge class imbalance. The data set target labels were imbalanced with 35 for the Hinselmann, 74 for Schiller, 44 for Cytology, and 55 Biopsy out of the 858 records as shown in Figure 1. SMOTE was used to deal with class imbalance. SMOTE works by oversampling the minority class by generating new synthetic data for minority instances based on nearest neighbors using the Euclidean Distance between data points [14]. Figure 1 shows the number of records per class labels in the data set.

2.3. Firefly Feature Selection. Dimensionality reduction is one of the effective ways to select the features that improve the performance of the supervised learning model. In the study, we adopted nature-inspired algorithm Firefly for selecting the features that better formulate the problem. Firefly was proposed by Yang [15] and was initially proposed for the optimization. Metaheuristic Firefly algorithm is inspired by fireflies' and flash lightening capability of a fly. It is a population-based optimization algorithm to find the optimal value or parameter for a target function. In this technique, each fly is pulled out by the glow intensity of the nearby flies. If the intensity of the gleam is extremely low at some point, then the attraction will be declining. Firefly used three rules; that is, (a) all the flies should be of the same gender; (b) the criteria of attractiveness depend upon the intensity of the glow; (c) target function will generate the gleam of the firefly. The flies with less glow will move towards the flies with brighter glow. The brightness can be adjusted using objective function. The same idea is implemented in the algorithm to search the optimal features that can better fit the training model. Firefly is more computationally economical and produced better outcome in feature selection when compared with other metaheuristic techniques like genetic algorithms and particle swarm optimization [16]. The time complexity of firefly is $O(n^2t)$ [17]. It uses the light intensity to select the features. Highly relevant features are represented as the features with high intensity light.

For feature selection, initially, some fireflies will be generated, and each fly will randomly assign the weights to all features. In our study, we generated 50 number of flies ($n = 50$). The dimension of the data set is 30. Furthermore, the lower bound was set to -50, while the upper bound is equal to 50. The maximum generations were 500. Additionally, α (alpha) was initially set to 0.5 and in every subsequent iteration, we used the (1) and (2) to update α (alpha) value.

$$X = \lim_{i=1 \longrightarrow 500} 1 - 10^{-(4/0.9)^{1/i}}, \tag{1}$$

$$\alpha = (1 - X) \times \alpha. \tag{2}$$

However, the gamma (γ) was set to 1. The number of features selected using Firefly for Hinselmann was 15, for Schiller 13 features, for Cytology 11 features, and 11 features for Biopsy, respectively.

2.4. Ensemble-Based Classification Methods. Three ensemble-based classification techniques such as Random Forest, Extreme Gradient Boosting, and Ada Boost were used to

Table 1: Statistical description of the data set.

Feature type	Feature name	Hinselmann Mean (μ) ± Std (σ)		Schiller Mean (μ) ± Std (σ)		Cytology Mean (μ) ± Std (σ)		Biopsy Mean (μ) ± Std (σ)	
		Positive	Negative	Positive	Negative	Positive	Negative	Positive	Negative
Demographic	Age	26.7 ± 7.7	26.8 ± 8.5	29.6 ± 11	26.6 ± 8.2	26.2 ± 8.4	26.9 ± 8.5	28.6 ± 8.9	26.7 ± 8.5
	Smokes	0.2 ± 0.4	0.1 ± 0.4	0.2 ± 0.4	0.2 ± 0.3	0.1 ± 0.3	0.1 ± 0.4	0.2 ± 0.4	0.1 ± 0.3
	Smokes (years)	2.5 ± 7.2	1.1 ± 3.9	2.4 ± 6.2	1.1 ± 3.8	1.1 ± 3.4	1.2 ± 4.1	2.2 ± 6.2	1.1 ± 3.9
	Smokes (packs/year)	0.7 ± 2.6	0.4 ± 2.2	0.6 ± 1.9	0.4 ± 2.2	0.5 ± 2.3	0.4 ± 2.2	0.7 ± 2.3	0.4 ± 2.2
	Number of sexual partners	2.2 ± 0.9	2.5 ± 1.7	2.5 ± 1.2	2.5 ± 1.7	2.7 ± 1.3	2.5 ± 1.7	2.5 ± 1.3	2.5 ± 1.7
	First sexual intercourse (age)	16.8 ± 2.0	16.9 ± 2.8	17 ± 2.5	16.9 ± 2.8	16.9 ± 2.9	16.9 ± 2.8	17.1 ± 2.6	16.9 ± 2.8
	Number of pregnancies	2.4 ± 1.4	2.5 ± 1.7	2.6 ± 1.7	2.2 ± 1.4	2.1 ± 1.4	2.2 ± 1.4	2.3 ± 1.3	2.2 ± 1.4
	Hormonal contraceptives	0.7 ± 4.5	0.7 ± 0.5	0.6 ± 0.5	0.7 ± 0.5	0.7 ± 0.5	0.7 ± 0.5	0.7 ± 0.5	0.7 ± 0.5
Habit	Hormonal contraceptives (years)	2.9 ± 4.8	1.9 ± 3.5	3.2 ± 5.2	1.9 ± 3.4	3.3 ± 6.4	1.9 ± 3.4	3.3 ± 5.4	1.9 ± 3.4
	IUD	0.2 ± 0.4	0.1 ± 0.3	0.2 ± 0.4	0.1 ± 0.3	0.1 ± 0.3	0.1 ± 0.3	0.2 ± 0.4	0.1 ± 0.3
	IUD (years)	0.6 ± 1.5	0.4 ± 1.8	0.9 ± 2.9	0.4 ± 1.7	0.5 ± 1.7	0.4 ± 1.8	0.7 ± 2.0	0.4 ± 1.8
	STDs	0.2 ± 0.4	0.1 ± 0.3	0.2 ± 0.4	0.1 ± 0.3	0.2 ± 0.4	0.2 ± 0.3	0.2 ± 0.4	0.1 ± 0.3
	STDs (number)	0.3 ± 0.9	0.2 ± 0.5	0.4 ± 0.8	0.1 ± 0.5	0.3 ± 0.7	0.1 ± 0.5	0.3 ± 0.8	0.1 ± 0.5
	STDs: condylomatosis	0.1 ± 0.3	0.1 ± 0.2	0.1 ± 0.3	0.0 ± 0.2	0.1 ± 0.3	0.0 ± 0.2	0.1 ± 0.3	0 ± 0.2
	STDs: cervical condylomatosis	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0
	STDs: vaginal condylomatosis	0 ± 0	0.0 ± 0.1	0 ± 0	0 ± 0.1	0 ± 0	0 ± 0.1	0 ± 0	0 ± 0.1
History	STDs: vulvo-perineal condylomatosis	0.1 ± 0.3	0.0 ± 0.2	0.1 ± 0.3	0 ± 0.2	0.1 ± 0.3	0 ± 0.2	0.1 ± 0.4	0 ± 0.1
	STDs: syphilis	0.0 ± 0.2	0.0 ± 0.1	0.0 ± 0.2	0 ± 0.1	0 ± 0	0 ± 0.1	0 ± 0	0 ± 0.1
	STDs: pelvic inflammatory disease	0 ± 0	0.0 ± 0.0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0
	STDs: genital herpes	0 ± 0	0.0 ± 0.0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0.1	0 ± 0
	STDs: molluscum contagiosum	0 ± 0	0.0 ± 0.0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0
	STDs: AIDS	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0
	STDs: HIV	0.1 ± 0.3	0 ± 0.1	0.1 ± 0.3	0 ± 0.1	0.1 ± 0.3	0 ± 0.1	0.1 ± 0.3	0 ± 0.1
	STDs: Hepatitis B	0 ± 0	0.0 ± 0.0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0
	STDs: HPV	0 ± 0	0.0 ± 0.0	0 ± 0	0 ± 0	0 ± 0	0 ± 0.1	0 ± 0	0 ± 0.1
	Dx: CIN	0 ± 0	0.0 ± 0.1	0.0 ± 0.1	0 ± 0.1	0 ± 0	0 ± 0.1	0.1 ± 0.2	0 ± 0.1
Genomics	Dx: HPV	0.1 ± 0.3	0.0 ± 0.1	0.1 ± 0.3	0 ± 0.1	0.1 ± 0.3	0 ± 0.1	0.1 ± 0.3	0 ± 0.1
	Dx	0.1 ± 0.3	0.0 ± 0.2	0.1 ± 0.3	0 ± 0.2	0.1 ± 0.3	0 ± 0.2	0.1 ± 0.3	0 ± 0.1
	Dx: cancer	0.1 ± 0.3	0.0 ± 0.1	0.1 ± 0.3	0 ± 0.1	0.1 ± 0.3	0 ± 0.1	0.1 ± 0.3	0 ± 0.1

train the model. The description of these techniques is discussed in the section below.

2.5. Random Forest. Random Forest (RF) was first proposed by Breiman in 2001 [18]. Random forest is an ensemble model that uses decision tree as individual model and bagging as ensemble method. It improves the performance of decision tree by adding many trees to reduce the overfitting in the decision tree. RF can be used for both classification and regression. RF generates a random forest that contains decision trees and gets a prediction from each one of them and then selects the best solution with the maximum votes [19].

When training a tree, it is important to measure how much each feature decreases the impurity, as the decrease in the impurity indicates the significance of the feature. The tree classification result depends on the impurity measure used. For classification, the measures for impurity are either Gini impurity or information gain and for regression, and the measure for impurity is variance. Training decision tree consists of iteratively splitting the data. Gini impurity decides the best split of the data using the formula.

$$G = 1 - \sum_i p_i^2, \tag{3}$$

where p (i) is the probability of selecting a datapoint with class; i.e., Information gain (IG) is also another measure to decide the best split of the data depending on the gain of each feature. The formula that calculates the information gain is given in the following equation:

$$\begin{gathered}\text{Entropy} = -\sum_i p_i \log_2 p_i, \\ \text{IG}(\text{parent}, \text{child}) = \text{Entropy}(\text{parent}) - [p1(c1) * \text{entropy}(c1) + p(c2) * \text{entropy}(c2) + \cdots].\end{gathered} \tag{4}$$

Table 2: Missing records per attribute in the data set.

Feature name	Missing values
First sexual intercourse (age)	7
Smoking	13
Smokes (years)	13
Smokes (packs/year)	13
Number of sexual partners	26
Number of pregnancies	56
Hormonal contraceptives	108
Hormonal contraceptives (years)	108
IUD	117
IUD (years)	117
STDs	105
STDs (number)	105
STDs: condylomatosis	105
STDs: cervical condylomatosis	105
STDs: vaginal condylomatosis	105
STDs: vulvo-perineal condylomatosis	105
STDs: syphilis	105
STDs: pelvic inflammatory disease	105
STDs: genital herpes	105
STDs: molluscum contagiosum	105
STDs: AIDS	105
STDs: HIV	105
STDs: hepatitis B	105
STDs: HPV	105
STDs: time since first diagnosis	787
STDs: time since last diagnosis	787

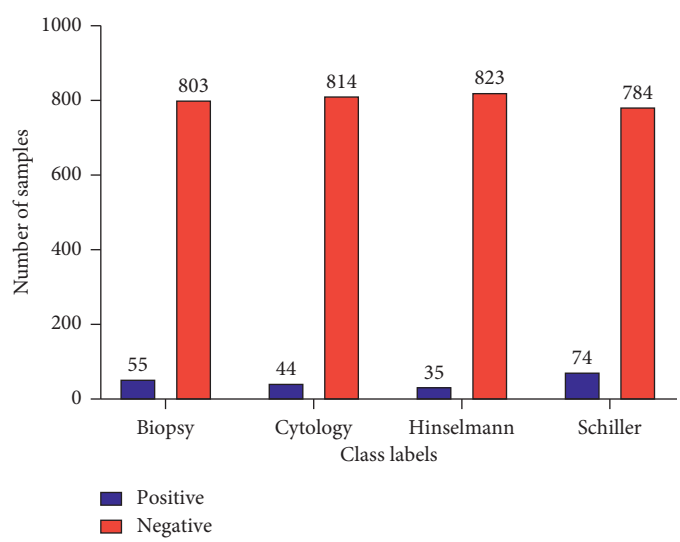


Figure 1: Number of records per class label in the data set.

2.6. Extreme Gradient Boosting. eXtreme Gradient Boosting (XGBoost) is a tree-based ensemble technique [20]. XGBoost can be used for classification, regression, and ranking problems. XG boosting is a type of gradient boosting. Gradient Boosting (GB) is a boosting ensemble technique that makes predicators sequentially instead of individually. GB is a method that produces a strong classifier by combining weak classifiers [21]. The goal of the GB is building an iterative model that optimizes a loss function. It pinpoints the failings of weak learners by using gradients in the loss function [21]:

$$y = ax + b + e, \tag{5}$$

where e denotes the error term. The loss function measures how good is the model at fitting the underlying data. The loss function depends on the optimization goal, for regression is a measure of the error between the true and predicated values, whereas, for classification, it measures the how good is a model at classifying cases correctly [21]. This technique takes less time and less iterations, since predictors are learning from the past mistakes of the other predictors. The

GB works by teaching a model C to predict values of the form

$$\mathcal{Y}' = C(x). \tag{6}$$

By minimizing a loss function, e.g., MSE:

$$\frac{1}{n}\sum_{i}\left(y_i' - y_i\right)^2, \tag{7}$$

where i iterates over a training set of size n of true values of the target variable yy′ = estimated values of C (x) y = true values & n = number of instances in y.

Considering a GB model with M phases and m as a single phase being $(1 \le m \le M)$, to improve some deficient model F_m, a new estimator h_m (x) is added. Therefore,

$$h_m(x) = \mathcal{Y} - F_m(x). \tag{8}$$

Estimator h will be fitted to $\mathcal{Y} - F_m(x)$, which is the difference between the true value and the predicated value, i.e., the residual. Thus, we attempt to adjust the errors of the previous model (F_m) [22].

XGBoost is better than Ada boost in terms of speed and performance. It is highly scalable and runs 10 times faster as compared to the other traditional single machine learning algorithms. XGBoost handles the sparse data and implements several optimization and regularization techniques. Moreover, it also uses the concept of parallel and distributed computing.

2.7. AdaBoost. Adaptive Boosting (AdaBoost) is a meta-learner originally proposed for the binary classification proposed by Freund and Schapire [23]. It is an ensemble technique to build a meta classifier by combining several weak classifiers using progressive learning.

AdaBoost uses the concept of boosting data sampling technique; adaptive sampling was used to assign high weights to the misclassified events. The misclassified samples will be selected in the next iteration to better train the model, and the final prediction was made using weighted voting. AdaBoost has reduced error rate, has a better effect on the prediction as compared to bagging [24], and uses decision tree stumps. Initially, all the samples in the data set have equal weights. Let x be the number of samples in the data set, and let y be the target. The target is a binary class represented by 0 and 1. The first decision tree stump will use some records from the data set, and predictions will be performed. After the initial prediction, the weights to the sample will be updated. More weights will be assigned to the data samples that were misclassified. The samples with the high weights will be selected in the next iteration. The process will be continued, unless the error rate is completely reduced, or a certain target level is achieved.

AdaBoost contains two main steps, combination and step forward using sequential iterative approach. All the instances in the training set have equal weights in the first iteration. However, in subsequent iterations, the weights are changed based on the error rates. The instances with error have increased weights. For the binary class classification problem containing T training samples is represented in the following equation:

$$\{(x_i, y_i)\}_{i=1}^{T}, \text{with,} \quad y_i \in \{0, 1\}. \tag{9}$$

Let C be the linear combination of weak classifiers. The combination of the classifiers is represented as

$$C(x) = \sum_{n=1}^{N} w_n c_n(x), \tag{10}$$

where N is the number of weak classifiers, w represents the weights, and C (x) represents weak classifiers. In every next iteration, the classifier is trained based on the performance of the classifier in previous iteration.

$$C(x)_t = C(x)_{t-1} + w_n c_n(x), \tag{11}$$

where $C(x)_t$ represents the classifier in t iteration. $C(x)_{t-1}$ is the performance of the classifier at $t-1$ iteration.

The weights can be calculated using the following equation:

$$w_n = \frac{1}{2}\ln\left(\frac{1-\epsilon_n}{\epsilon_n}\right), \tag{12}$$

ϵ_n represents the error rate of the weak classifier.

2.8. Optimization Strategy. This section discusses optimization strategy to find the best hyperparameters combination that produces the highest targeted outcomes. Firefly optimization algorithm was used for parameter tunning. The details of Firefly are discussed in Section 2.3. Table 3 presents the hyperparameter values of Random Forest for all the four targets, For RF "gini" index criterion was used. Table 4 represents the parameters used for XGBoost. Gbtree booster was used with the random state of 42 and the learning rate of 0.05. Similarly, Table 5 presents the optimal feature vales for AdaBoost. Furthermore, Figures 2–4 represent the Grid Search optimization graph for Random Forest, Extreme Gradient Boosting, and AdaBoost classifier.

3. Experimental Setup and Results

The model was implemented in Python language 3.8.0 release using Jupyter Notebook environment. Ski-learn library was used for the classifiers along with other needed built-in tools, while separate library (xgboost 1.2.0) was used for XGBoost ensemble. There is K-fold cross validation with $K = 10$ for partitioning the data into training and testing. Five evaluation measures such as accuracy, sensitivity (recall), specificity (precision), positive predictive accuracy (PPA), and negative predictive accuracy (NPA) were used. Sensitivity and specificity are focused more during the study due to the application of the proposed model. Accuracy denotes the percentage of correctly classified cases, sensitivity measures the percentage of positives cases that were classified as positives, and specificity refers to the percentage of negative cases that were classified as negatives. Moreover, the criteria for the selection of the performance evaluation

Table 3: Random Forest optimized parameters set for all four targets using Firefly.

Parameters	Optimal values obtained
n_estimators	100
max_features	Log2
criterion	gini
max_depth	15
min_samples_split	5
min_samples_leaf	1

Table 4: XGBoost optimized parameters set for all four targets using Firefly.

Parameters	Optimal values obtained
Booster	gbtree
Random_state	42
nthread	8
learning_rate	0.05
gamma (i.e. min_split_loss)	0.1
max_depth	3

Table 5: AdaBoost optimized parameters set for all four targets using Firefly.

Parameters	Optimal values obtained
base_estimator	None
n_estimators	600
learning_rate	1.0

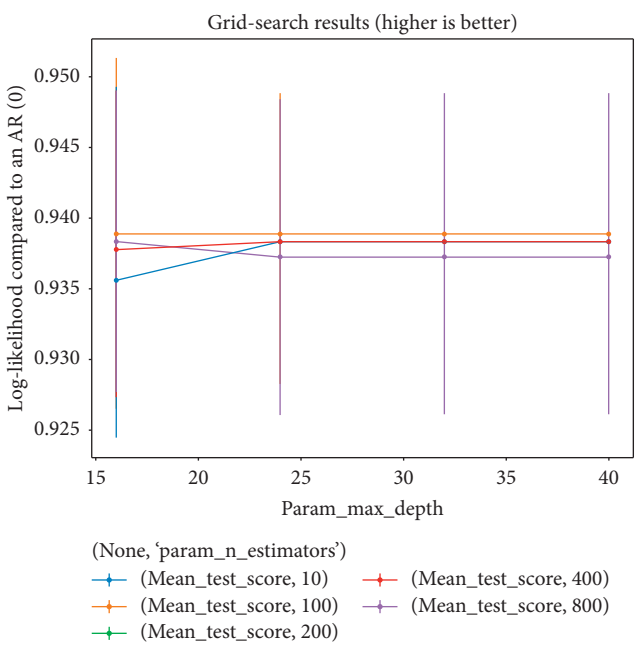


Figure 2: Grid search optimization for random forest.

measures depend upon the measures used in the benchmark studies. Two sets of experiments were conducted for each target using selected features by using Firefly feature selection algorithm and 30 features for four targets. The SMOTE technique was applied to generate synthetic data. The results of model are presented in section below.

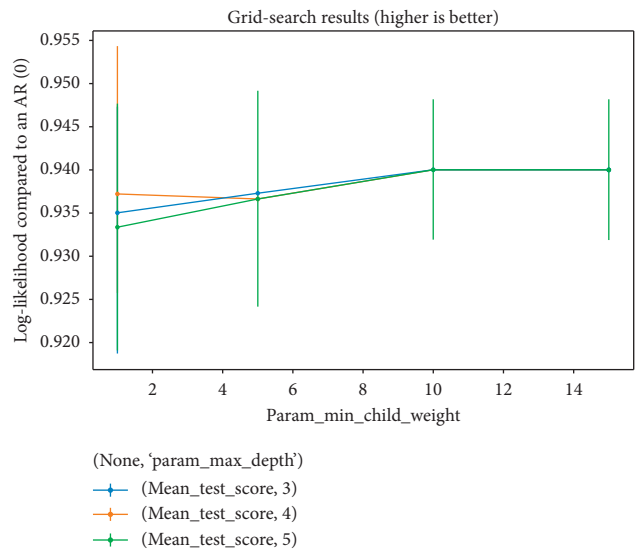


Figure 3: Grid search optimization for extreme gradient boosting.

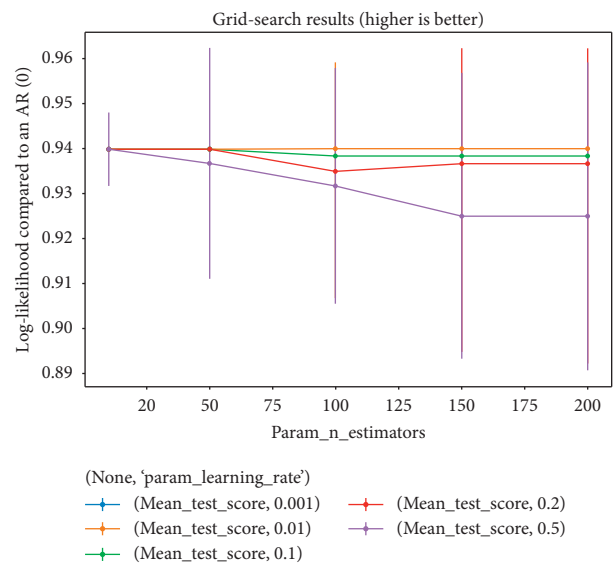


Figure 4: Grid search optimization for AdaBoost.

3.1. Hinselmann. Table 6 presents the accuracy, sensitivity, specificity, PPA, and NPA for the RF, AdaBoost, and XGBoost models, respectively, using SMOTE for Hinselmann test target class. The number of selected features for Hinselmann was 15. XGBoost outperformed the other classifiers for both feature sets. However, the performance of XGBoost with selected feature is better when compared with 30 features. The model produces an accuracy of 98.83, sensitivity of 97.5, specificity of 99.2, PPA of 99.17, and NPA of 97.63, respectively.

3.2. Schiller. Table 7 presents the outcomes for the Schiller test. Like Hinselmann target, XGBoost with selected features outperformed that of Schiller, respectively. However, the outcomes achieved by the model for Schiller are lower when compared with Hinselmann target class. The performance of

Table 6: Performance of ensemble classifiers using 30 and selected feature for Hinselmann target.

Features	Model	Accuracy	Sensitivity	Specificity	PPA	NPA
30	RF	96.36	97.05	95.72	95.44	97.23
	AB	90.08	90.34	89.84	89.21	90.91
	XGB	97.37	96.72	98	97.93	96.84
Selected	RF	97.57	97.12	98.01	97.93	97.23
	AB	94.94	95	94.88	94.61	95.26
	XGB	**98.38**	**97.55**	**99.2**	**99.17**	**97.63**

Table 7: Performance of ensemble classifiers using 30 and selected feature for Schiller target.

Features	Model	Accuracy	Sensitivity	Specificity	PPA	NPA
30	RF	93.84	94.54	93.12	93.36	94.35
	AB	86.62	85.32	88.13	89.21	83.91
	XGB	92.36	93.99	90.76	90.87	93.91
Selected	RF	95.97	95.49	96.48	96.68	95.89
	AB	89.6	86.92	92.89	93.78	85.22
	XGB	96.98	95.9	96.92	97.1	95.65

RF and XGBoost is similar with selected feature for Schiller with a minor difference. The number of features selected by Firefly for Schiller was 13.

3.3. Cytology. Table 8 presents the outcome of all the classifiers for the cytology diagnosis tests. Like Hinselmann and Schiller diagnostic test, XGBoost outperformed Cytology test as well with selected features. For specificity and accuracy, similar outcomes were achieved using 30 and selected features. Similarly, the performance of RF is similar in both 30 and selected features. The number of features selected by Firefly feature selector for Cytology was 11.

Table 8: Performance of ensemble classifiers using 30 and selected feature for Cytology target.

Features	Model	Accuracy	Sensitivity	Specificity	PPA	NPA
30	RF	95.91	94.09	97.87	97.95	93.88
	AB	93.66	91.12	96.52	96.72	90.61
	XGB	96.32	94.49	98.3	98.36	94.29
Selected	RF	95.91	94.44	97.47	97.54	94.29
	AB	92.64	90.94	94.47	94.67	90.61
	XGB	**96.93**	**95.26**	**99.74**	**98.73**	**95.1**

3.4. Biopsy. Similarly, performance was not drastically different, yet using all the features resulted in a higher accuracy than when using SMOTE with selected features for Biopsy as shown in Table 9. XGB obtained the highest accuracy of 97.1 with all features. However, for other measures, the performance of the XGBoost is better with the selected features. Similar performance was achieved for all measures when classified using RF for both feature sets 30 and selected, respectively. The number of selected features used for Biopsy target class was 11.

Table 9: Performance of ensemble classifiers using 30 and selected feature for Biopsy target.

Features	Model	Accuracy	Sensitivity	Specificity	PPA	NPA
30	RF	96.68	95.45	97.32	97.88	95.53
	AB	93.57	93.99	93.17	92.8	94.31
	XGB	97.1	95.49	97.52	97.46	95.53
Selected	RF	96.27	95.8	96.72	96.61	95.93
	AB	89.21	88.02	90.42	90.25	88.21
	XGB	**96.68**	**95.83**	**98.74**	**98.73**	**95.93**

Overall, after comparing all the four-diagnostic tests, Hinselmann test achieved the better outcome and can be used for the diagnosis of cervical cancer as shown in Table 10. As per the outcome achieved in the proposed study, Hinselmann diagnosis test has better performance when compared from other cervical cancer diagnosis tests like Schiller, Biopsy, and Cytology, respectively. Similar findings have been made in Abdoh et al. [5] and Wu and Zhou [4] study.

Table 10: Results of proposed model for 4 diagnosis tests for cervical cancer.

Diagnosis test	Accuracy	Sensitivity	Specificity	PPA	NPA
Hinselmann	**98.38**	**97.55**	**99.2**	**99.17**	**97.63**
Schiller	96.98	95.9	96.92	97.1	95.65
Biopsy	96.68	95.83	98.74	98.73	95.93
Cytology	96.93	95.26	98.73	98.77	95.1

4. Comparison with Existing Studies

The study used three ensemble techniques AdaBoost, extreme Gradient Boosting, and Random Forest. Furthermore, the proposed study is the pioneer in using bioinspired algorithm for feature selection and optimization for cervical cancer diagnosis. To explore the significance of our proposed study, the outcome of the study was compared with the benchmark studies. The criteria for the benchmark studies selection were based on data set used for the diagnosis of cervical cancer. Table 11 contains the comparison of the proposed technique with the benchmark studies in the literature. The best outcomes in the benchmark studies were achieved using 30 features. However, some of the outcomes in the previous studies were achieved with the reduced features. The number in the brackets next to some of the outcomes represents the number of features.

Therefore, based on Table 11, the proposed study outperforms the two studies in the benchmark interms of accuracy with reduced risk factors. However, the achieved sensitivity and NPA are less than those of Wu and Zhou [4] but higher than those of Abdoh et al. [5]. The number of features in Wu et al. study is 30, while the proposed study used reduced risk factors. The specificity and PPA of the proposed study are higher than those of the benchmark studies except for the Schiller diagnosis test.

In nutshell, the main contributions of the current study are applying bioinspired algorithm for feature selection and for model optimization for cervical cancer risk factors. The

Table 11: Comparison of the proposed study with benchmark studies.

Target class	Model	Accuracy	Sensitivity	Specificity	PPA	NPA
Hinselmann (15)	Abdoh et al. [5]	93.97	**100**	89.96	84.97	**100**
	Wu and Zhou [4]	97.6	96.65	98.54	98.48	96.78
	Proposed study	**98.38**	97.55	**99.2**	**99.17**	97.63
Schiller (13)	Abdoh et al. [5]	90.18	**98.73**	84.63	80.75	**99.03**
	Wu and Zhou [4]	95.01	93.24	**97.58 (12)**	**97.29 (12)**	93.81
	Proposed study	**96.39**	95.9	96.92	97.1	95.65
Cytology (11)	Abdoh et al. [5]	92.75	**100**	87.92	83	**100**
	Wu and Zhou [4]	96.94	95.58 (8)	99.01	98.94	95.76 (8)
	Proposed study	**96.98**	95.83	**99.74**	**98.73**	95.93
Biopsy (11)	Abdoh et al. [5]	94.13	**100**	90.21	86.07	**100**
	Wu and Zhou [4]	96.06	94.94 (6)	97.76 (11)	97.58 (11)	94.91
	Proposed study	**96.93**	95.26	**98.73**	**98.77**	95.1

proposed model enhanced the outcomes when compared with the previous studies related with cervical cancer risk factors data set. Despite the above-mentioned advantages, the study suffers from some limitations: the data set suffers from huge imbalance, and augmented data was generated using SMOTE. Moreover, the current study was based on open-source data set, and further testing is required to use other real and open-source data sets.

To alleviate the above-mentioned limitations, there is a need for validating the model on real data set from the hospital.

5. Conclusion

This study presents an investigation of several ensemble techniques such as Random Forest, AdaBoost, and Extreme Gradient Boosting for diagnosing cervical cancer. The data set was obtained from the UCI machine learning repository containing 858 records, 32 features, and 4 target variables. The target variables are the diagnosis test used for cervical cancer. Experiments were conducted for each target class separately. Data preprocessing includes imputing missing values and class balancing using SMOTE. Moreover, bioinspired firefly algorithm was used to optimize the models, and to identify the key features. To compare the performance of the models, the experiments were conducted with 30 features and the selected features using SMOTED data. Extreme Gradient Boosting outperformed the other two models for all four target variables. For future work, the model will be validated on multiple data sets. Also, other models that can handle outliers and unbalanced data differently should be investigated.

Data Availability

The study used open-source data set available at https://archive.ics.uci.edu/ml/datasets/Cervical+cancer+%28Risk+Factors%29.

Conflicts of Interest

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

References

[1] American Cancer Society, *Key Statistics for Cervical Cancer*, American Cancer Society, Atlanta, GA, USA, 2021, https://www.cancer.org/cancer/cervical-cancer/about/key-statistics.html.

[2] "Cervical cancer: risk factors," 2020, https://www.cancer.net/cancer-types/cervical-cancer/risk-factors#:%7E:text=The%20most%20important%20risk%20factor,100%20different%20types%20of%20HPV.

[3] "Cervical cancer," 2020, https://www.nccc-online.org/hpvcervical-cancer/cervical-cancer-overview/.

[4] W. Wu and H. Zhou, "Data-driven diagnosis of cervical cancer with support vector machine-based approaches," *IEEE Access*, vol. 5, 2017.

[5] S. F. Abdoh, M. Abo Rizka, and F. A. Maghraby, "Cervical cancer diagnosis using random forest classifier with SMOTE and feature reduction techniques," *IEEE Access*, vol. 6, 2018.

[6] J. Lu, E. Song, A. Ghoneim, and M. Alrashoud, "Machine learning for assisting cervical cancer diagnosis: an ensemble approach," *Future Generation Computer Systems*, vol. 106, pp. 199–205, 2020.

[7] E. Karim and N. Neehal, "An empirical study of cervical cancer diagnosis using ensemble methods," in *Proceedings of the 2019 1st International Conference on Advances in Science, Engineering and Robotics Technology (ICASERT)*, Dhaka, Bangladesh, December 2019.

[8] A. Ul-Islam, S. H. Ripon, and N. Qaisar Bhuiyan, "Cervical cancer risk factors: classification and mining associations," *APTIKOM Journal on Computer Science and Information Technologies*, vol. 4, no. 1, pp. 8–18, 2019.

[9] Y. M. S. Al-Wesabi, A. Choudhury, and D. Won, "Classification of cervical cancer dataset," in *Proceedings of the 2018 IISE Annual Conference and Expo*, pp. 1456–1461, Orlando, FL, USA, May 2018.

[10] K. Fernandes, D. Chicco, J. S. Cardoso, and J. Fernandes, "Supervised deep learning embeddings for the prediction of cervical cancer diagnosis," *PeerJ Computer Science*, vol. 4, no. 5, pp. e154–21, 2018.

[11] K. Adem, S. Kiliçarslan, and O. Cömert, "Classification and diagnosis of cervical cancer with stacked autoencoder and softmax classification," *Expert Systems with Applications*, vol. 115, pp. 557–564, 2019.

[12] K. Fernandes, J. S. Cardoso, and J. Fernandes, "Transfer learning with partial observability applied to cervical cancer screening," *Pattern Recognition and Image Analysis*, Springer, vol. 10255, pp. 243–250, Berlin, Germany, 2017.

[13] "Cervical cancer (risk factors) data set," 2020, https://archive.ics.uci.edu/ml/datasets/Cervical+cancer+%28Risk+Factors%29.

[14] A. Fernandez, S. Garcia, F. Herrera, and N. V. Chawla, "SMOTE for learning from imbalanced data: progress and challenges, marking the 15-year anniversary," *Journal of Artificial Intelligence Research*, vol. 61, pp. 863–905, 2018.

[15] X. Yang, *Nature-Inspired Metaheuristic Algorithms*, Vol. 4, Luniver Press, Cambridge, UK, 2nd edition, 2010.

[16] B. H. Nguyen, B. Xue, and M. Zhang, "A survey on swarm intelligence approaches to feature selection in data mining," *Swarm and Evolutionary Computation*, vol. 54, 2020.

[17] M. Anbu and G. S. Anandha Mala, "Feature selection using firefly algorithm in software defect prediction," *Cluster Computing*, vol. 22, no. s5, Article ID 10925, 2019.

[18] L. Breiman, "Random forests," in *Hands-On Machine Learning with R*, pp. 203–219, CRC Press, Boca Raton, FL, USA, 2019.

[19] T. M. Oshiro, P. S. Perez, and J. A. Baranauskas, "How many trees in a random forest?" in *Lecture Notes in Computer Science*, vol. 7376, pp. 154–168, Springer, New York, NY, USA, 2012.

[20] T. Chen and C. Guestrin, "XGBoost: a scalable tree boosting system," in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, vol. 13–17, pp. 785–794, San Francisco, CA, USA, August 2016.

[21] A. Natekin and A. Knoll, "Gradient boosting machines, a tutorial," *Frontiers in Neurorobotics*, vol. 7, 2013.

[22] J. H. Friedman and J. J. Meulman, "Multiple additive regression trees with application in epidemiology," *Statistics in Medicine*, vol. 22, no. 9, pp. 1365–1381, 2003.

[23] Y. Freund and R. E. Schapire, "A decision-theoretic generalization of on-line learning and an application to boosting," *Journal of Computer and System Sciences*, vol. 55, no. 1, pp. 119–139, 1997.

[24] S. B. Kotsiantis, "Supervised machine learning: a review of classification techniques," *Informatica*, vol. 31, pp. 249–268, 2007.

Hindawi
Scientific Programming
Volume 2021, Article ID 3370580, 7 pages
https://doi.org/10.1155/2021/3370580

Research Article

Big Data and Deep Learning Model for FMS Score Prediction of Aerobics Athletes

Wenying Xiong, Dongqin Huang, and Wei Xu

College of Physical Education and Health, Jiangxi University of Traditional Chinese Medicine, Nanchang 330000, Jiangxi, China

Correspondence should be addressed to Wei Xu; 20112001@jxutcm.edu.cn

Received 20 April 2021; Revised 18 June 2021; Accepted 3 July 2021; Published 17 July 2021

Academic Editor: Shah Nazir

In recent years, competitive aerobics has developed rapidly in my country, and the corresponding sports injury risks have gradually increased. A number of studies have shown that due to the characteristics of aerobics itself, difficult movement requirements, fast-paced music accompaniment and coherent coordinated movements, athletes will suffer sports injuries if they are not paying attention. Therefore, discovering the causes of athletes' injuries in time and preventing them in time is crucial for improving athletes' skill level and prolonging sports life. Through the functional movement screening (FMS) test, understanding young aerobics athletes' insufficiency in trunk stability, joint flexibility, muscle extension, and core strength can further help athletes reduce the risk of sports injuries. Therefore, this article proposes a novel sports injury risk model based on big data technology and deep learning, which can effectively predict the risk of sports injury and can play a positive role in improving the quality of athletes' movements and prolonging their sports life.

1. Introduction

In recent years, competitive aerobics [1–3] has developed rapidly in my country, and the corresponding sports injury risks [4–6] have gradually increased. A number of studies have shown that due to the characteristics of aerobics itself, strict time requirements, more difficult action requirements, fast-paced music accompaniment, and coherent coordinated actions, athletes will suffer sports injuries if they are not paying attention. The shoulders, elbows, wrists, waists, thighs, knees, calves, and ankles are the most prone to injury during aerobics training [7–9]. Among them, the most prone to injury is the ankle joint. In addition, the type of injury most likely to occur for competitive aerobics athletes is closed injury, most of which are joint strain, sprain and muscle strain, and chronic injuries are the main ones. However, the current scholars' research on aerobics injuries is usually carried out through manual investigation methods.

Aerobics athletes are in the golden stage of physical development [10]. During this period, various physical qualities will be significantly improved. However, in interviews with aerobics athletes and coaches, it is found that the types of injuries of aerobics athletes are higher than those of other athletes. Moreover, this is due to the weakness of the aerobics athletes' own muscles and joints, which limits their skill development ability [11] and long-term training of irregular technical movements and body postures. In the teenage years of aerobics athletes, scientific and reasonable training can not only promote the physical development of adolescents but also improve their athletic ability more effectively. Therefore, discovering the causes of aerobics athletes' injuries in time and preventing them in time is crucial for improving athletes' skill level and prolonging sports life [12].

Functional movement screening [13–15] is a test model proposed by American orthopedic training experts in 1996. The test mainly uses 7 functional movements to detect the stability of the subjects' overall movement, joint flexibility, softness, balance ability, core strength, and proprioception, find out whether the subject's movement pattern exists or is potentially compensatory and noncompetitive symmetry, and then analyze the risk of subjects' sports injuries. As long

as there is a serious asymmetry or defect in the action posture, it is necessary to use accurate and appropriate corrective training to correct the wrong posture in time and, at the same time, make more effective suggestions so that the athlete can better prevent the risk of injury. For more than ten years, the FMS test has been continuously applied and adjusted in functional sports and clinical aspects. The test methods and scoring rules have been very complete and standardized. At present, as an important supplement to traditional training methods, FMS has been widely used in professional sports leagues.

At present, domestic FMS research has a wide range of research objects in the field of sports, but there are few research studies on young aerobics athletes. Young athletes are related to the development of sports in the future, and their physical ability is as important as their sports life span. Therefore, through the FMS test, understanding young aerobics athletes' insufficiency in trunk stability, joint flexibility, muscle extension, and core strength can further help athletes reduce the risk of sports injuries. Therefore, this paper constructs an FMS prediction algorithm based on big data technology and deep learning [16–20], which can further improve the prediction accuracy of motion loss risk. The following are the main innovative points of this paper:

(i) A novel sports injury risk model based on big data technology and deep learning, which can effectively predict the risk of sports injury and can play a positive role in improving the quality of athletes' movements and prolonging sports life

(ii) FMS test and deep neural network to construct a joint algorithm for sports injury prediction, which can further improve the accuracy of sports injury risk prediction

(iii) Simulation and ablation experiments are carried out, and the experimental results prove the effectiveness and superiority of the algorithm in this paper

2. Related Work

2.1. FMS. FMS, functional motor screening, was first proposed by American orthopedic training experts Gray Cook and Lee Burton. Derived from the famous functional movement training, it was first applied in the 1990s. At present, it has been widely used as a testing method in the field of physical therapy, rehabilitation, and physical training and is suitable for all kinds of people. FMS test combines the comprehensive knowledge of sports anatomy, sports physiology, sports biomechanics, and neurology to connect with the basic movements of the human body. After long-term practice and research, seven basic movements' tests and three exclusion tests are finally determined. From the standpoint of human basic sports ability, a comprehensive assessment of the human body is conducted using a simple scientific and intuitive method of grading basic action to detect human body movement coordination, flexibility, stability, and symmetry, and screening out the compensatory actions that are not conducive to the development of the level of movement and harm the body [21].

The test significance of FMS is to evaluate the quality of actions and to score and rank some of the defects and asymmetries of certain action patterns so that specialized technical actions can be better developed. As a result, Cook proposed the best performance pyramid (Figure 1) and pointed out that the first layer is "Movement": the human body's most basic athletic ability, that is, the most fundamental flexibility and stability in sports. The second layer is "Performance": the speed, strength, endurance, and other qualities of human movement. The third layer is "Skill." Cook puts forward that the flexibility of the body and the ability to control stability are the basis for the free movement of the human body during exercise, and it also provides a guarantee for the completion of high-quality movements so as to achieve a higher level of exercise goals.

Lloyd et al. [22] found in the study of juvenile football players that the total FMS score will also be affected by the age of the player and the level of different athletic abilities displayed. The "physical maturity" shown by athletes of different ages is closely related to the quality of movement completion in the FMS test. Marques et al. [23] conducted FMS tests on young football players aged 14–20 and found that young football players have asymmetrical heights on the left and right sides of the body. 91% of the players had 0 or 1 points during the test. 82% of the athletes' test score does not exceed 14 points, indicating that they have a higher potential risk of injury.

2.2. Sports Injuries of Aerobics. Malliou et al. [24] used a questionnaire survey method and conducted statistics and analysis on the injuries of athletes engaged in aerobics training and found that, in the injured population, lower extremity injuries accounted for 97.3%, and the injuries of the stomp joints and knee joints are the most common in aerobics training. At the same time, it is pointed out that the training time, years, and training level will all have an impact on the injury. Bintoudi et al. [25] investigated two aerobics pedal athletes and found that they had knee joint pain and fat pad edema. They also discussed the possible pathogenic factors and mechanisms of aerobics.

3. Methodology

Due to the characteristics of competitive aerobics, athletes are required to complete a series of high-intensity movements in a short time, which requires a higher level of physical fitness and physical flexibility of the athletes. Studies have confirmed that long-term high-intensity repetitive exercise training and asymmetric sports skills and postures will increase the risk of athletes' injuries. At the same time, taking into account the rapid development of big data and neural network technology [26–30], the construction of aerobics sports injury is feasible.

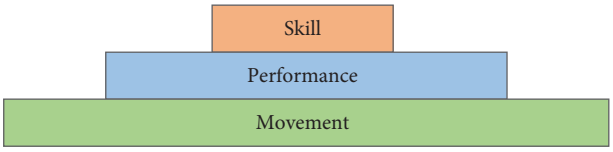


Figure 1: Best performance pyramid.

3.1. The Functional Movement System of Aerobics Athletes. FMS motion screening detects weak points in human movement to correct incorrect posture and reduces the risk of sports injury. The functional movement system (FMS) includes seven test movements, including squat over the top mode, hurdle step, lunge squat, shoulder flexibility, straight leg lift, push-ups, and rotational stability. The functional movement screening test only detects pain or no pain and does not assign a score. A total of 7 functional movements are scored on a scale of 0 to 3, with a total of four scores based on how well the subjects complete the above 7 basic test movements. The subjects will get 3 points if they complete the specified action pattern, with the highest score of 3 points; 2 points will be given to the subjects who are in the mode of completion (having the ability to complete) but have compensatory actions or poor quality of completion when completing the actions; 1 point will be given when the subject cannot complete the functional movements or does not have the ability to complete the movements and fails to maintain body balance during the movements. The injured person has pain in any part of the body at the time of completing the action. A total FMS score of 21, with a functional screening score of less than 14 indicating a greater risk of injury during training, is obtained; the higher the score, the lower the risk of injury.

3.2. Sports Injury Prediction Model

3.2.1. Analysis of Time and Space Correlation of Aerobics. We know that aerobics actions have time and space correlation (as shown in Figure 2), so this article uses the CNN model [31–33] to extract the time and space characteristics of aerobics actions and effectively utilizes the time and space characteristics of aerobics actions. First, we introduce the theory of correlation analysis.

(1) Covariance: in traditional statistical analysis, covariance can measure the joint change of two random variables. The sign of covariance reflects the linear relationship between variables. Suppose the random variables are X and Y; then, the covariance calculation equation between these two random variables is as follows:

$$\operatorname{cov}(X,Y) = E\{[X - E(X)][Y - E(Y)]\}, \tag{1}$$

where $E(X)$ and $E(Y)$ are the expected values of random variables X and Y, respectively. Specifically, the covariance represents the expectation of the overall error of the two variables.

(2) Pearson correlation coefficient: the correlation coefficient was put forward by the statistician Karl Pearson. It is an index to analyze the degree of linear correlation among variables. The Pearson correlation coefficient is widely used. Assuming that there are two sets of data sets X and Y, n is the sample size, and r is used to represent the Poisson correlation coefficient of the two. The calculation equation of the Poisson correlation coefficient is as follows:

$$r = \frac{\operatorname{cov}(X,Y)}{\sqrt{\sigma_X^2}\sqrt{\sigma_Y^2}} = \frac{\sum_{i=1}^{n}(x_i - \overline{x})(y_i - \overline{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \overline{x})^2}\sqrt{\sum_{i=1}^{n}(y_i - \overline{y})^2}}, \tag{2}$$

where $\operatorname{cov}(X,Y)$ is the covariance of X and Y, σ_X^2 and σ_Y^2 are the variances of the variables X and Y, respectively, and $\overline{x}$ and $\overline{y}$ are the mean values of the variables x and y.

By extending the low-dimensional random vector to the high-dimensional random vector, the correlation coefficient matrix can be obtained, and each element in the correlation coefficient matrix is the correlation coefficient of the row vector and column vector where it is located. Assuming θ is a column vector composed of n scalars randomly, $\theta = [X_1, X_2, \ldots, X_n]$, the calculation equation of the correlation coefficient matrix is as follows:

$$r = \begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1n} \\ r_{21} & r_{22} & \cdots & r_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ r_{n1} & r_{n2} & \cdots & r_{nn} \end{bmatrix}. \tag{3}$$

The calculation equation for finding the element value of the ith row and jth column of the correlation coefficient matrix r is as follows:

$$r_{ij} = \frac{\operatorname{cov}(x_i, x_j)}{\sigma_{x_i}\sigma_{x_j}} = \frac{E\{[X_i - E(X_i)][X_j - E(X_j)]\}}{\sqrt{E(X_i^2) - E^2(X_i)}\sqrt{E(X_j^2) - E^2(X_j)}}, \tag{4}$$

where σ_{X_i} and σ_{X_j} are the mean square deviations of random variables X_i and X_j, respectively.

3.2.2. Construction of Spatiotemporal Feature Matrix. The following is the specific approach for the original data refactoring and data integration for the matrix form as a model of the original input data in this paper: to begin, combine the adjacent section 25 testing points with 15 consecutive time nodes of aerobics movement data, with a 5-second interval between time nodes, and then build each testing point of time series data as given in the equation below:

$$P_{i,j} = [t_{i,1}, t_{i,2}, \ldots, t_{i,15}], \tag{5}$$

where $t_{i,j}$ represents the i th detection point and the aerobics action data of the j th time node among the selected 15 time nodes. Finally, the data of each detection point is integrated, and the calculation equation of the 15×15 space-time characteristic matrix is as follows:

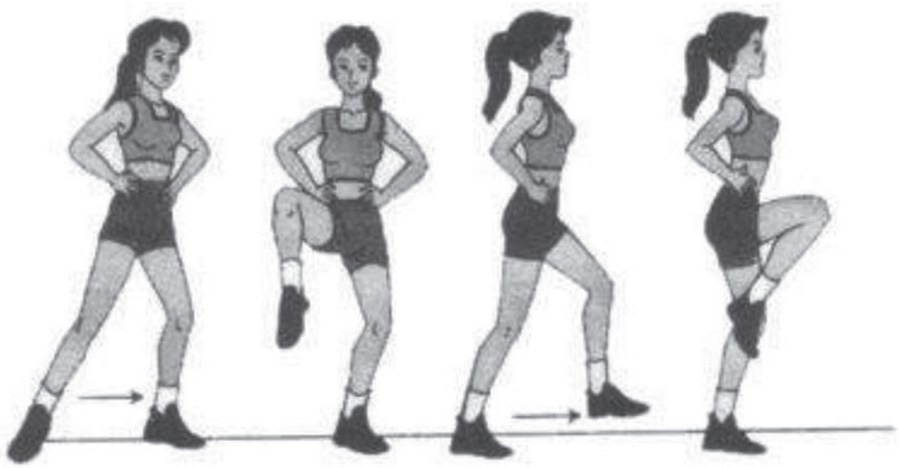


Figure 2: Basic structure of standard recurrent neural network.

$$D = \begin{bmatrix} t_{1,1} & t_{1,2} & \cdots & t_{1,15} \\ t_{2,1} & t_{2,2} & \cdots & t_{2,15} \\ \vdots & \vdots & \ddots & \vdots \\ t_{15,1} & t_{15,2} & \cdots & t_{15,15} \end{bmatrix}. \quad (6)$$

The matrix form of aerobics action data constructed above fully considers the spatiotemporal correlation, but the amount of data is still too large, resulting in time-consuming model training and increased complexity. In response to this problem, we use convolutional neural networks to simplify matrix features. Convolution filtering and downsampling are two operations to extract features.

3.2.3. CNN Model. Multilayer CNN includes convolution layer, pooling layer, full connection layer, and output layer. Neurons in adjacent layers are connected to each other, but neurons in the same layer are not connected to each other. We divide CNN into three parts: input feature matrix, feature extraction, and feature vector output. Feature extraction mainly includes convolution and pooling. The convolutional layer C carries out convolution calculation through multiple convolutional check input features to obtain multiple convolutional feature vectors, and the pooling layer will carry out local feature extraction. The convolution calculation equation is as follows:

$$X_{i,j} = f\left[\sum_{q-1}^{r}\sum_{p-1}^{r}\left(D_{(i+p)(j+q)}C_{pq}\right) + b_c\right]. \quad (7)$$

The constructed feature matrix uses the rxr convolution kernel c to perform a sliding calculation with a step length of 1 plus the bias variable b to obtain a feature vector of $(n-r+1)x(m-r+1)$ dimensions. Input the feature vector into the activation function factory. The calculation equation of the linear rectification (ReLU) activation function used in this paper is as follows:

$$f(x) = \max(0, x). \quad (8)$$

Compared with activation functions such as sigmoid, this function reduces the computational complexity, overcomes the problem of gradient disappearance, and converges faster. The convolution calculation uses partially connected features to simplify some features of the feature matrix. During iteration, the principle of gradient descent is used to continuously adjust the shared weights in the convolution kernel so that the data features can be fundamentally mined.

3.2.4. CNN Extracts Spatiotemporal Features. We use the preprocessed aerobics action data of 15 adjacent detection points, and the data of the first 125 seconds of a single sample of each detection point constructs a 15×15 feature matrix. We set two convolutions and two pooling processing feature matrices, and the ReLU activation function is introduced. The size of the first layer of convolution kernel is set to 3X3, the step size is set to 1, and the size of the feature matrix obtained after one convolution is 23×23; then, the average pooling is performed, the size is set to 2×2, and the step size is set to 2; here, padding is set to "SAME," and the size of the feature matrix after one pooling becomes 12×12; then, after another convolution, the size of the convolution kernel is set to 3×3, and the step size is set to 1, and we get the 10×10 feature matrix which is subjected to mean pooling again. The size is set to 2×2, the step size is 2, and the 5×5 feature matrix is obtained. The CNN model structure is shown in Figure 3.

4. Experiments and Results

4.1. Experimental Subject. In this study, 200 college students from a certain city were selected as the research objects. According to gender and aerobics level, they were divided into male aerobics group, male first-grade aerobics group, male second-grade aerobics group, male general student group, and female athlete group. There was no statistical difference in basic information such as height, weight, and BMO among 10 people in each group, including the first-level female aerobics group, the second-level female aerobics

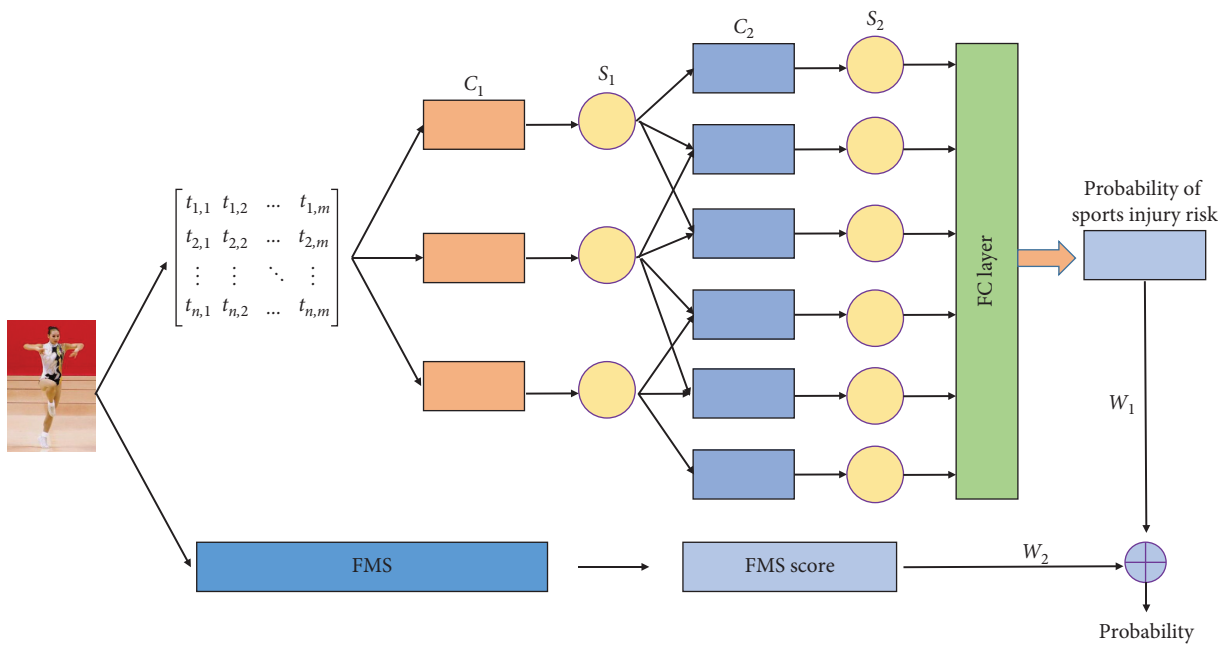
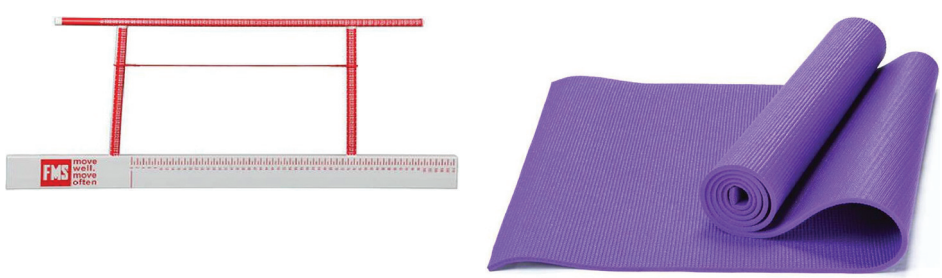


Figure 3: The overall structure of our model.

Table 1: Basic situation of the research object.

Group	Sample size	Age	Height (cm)	Weight (kg)	BMI
Master aerobics group	10	21.5	172.1	62.3	21.1
Level 1 aerobics group	10	22.1	174.3	63.1	20.5
Level 2 aerobics group	10	19.6	171.6	60.3	19.6
Ordinary student group	10	18.9	168.8	58.3	21.4



(a) (b)

Figure 4: FMS test suite. (a) FMS suite. (b) Yoga mat.

Table 2: Experimental hardware platform and software simulation environment.

CPU	Intel (R) core (TM) i5-4200M CPU @ 2.50 GHz
RAM	8.00 GB
Operating system	Windows10
Development environment	VS2015
Development tools	OpenCV + Pycharm

group, and the female general student group. The basic situation is detailed in Table 1, and the FMS test suite is shown in Figure 4.

4.2. Simulation Experiment Environment. The hardware environment and software environment are shown in Table 2.

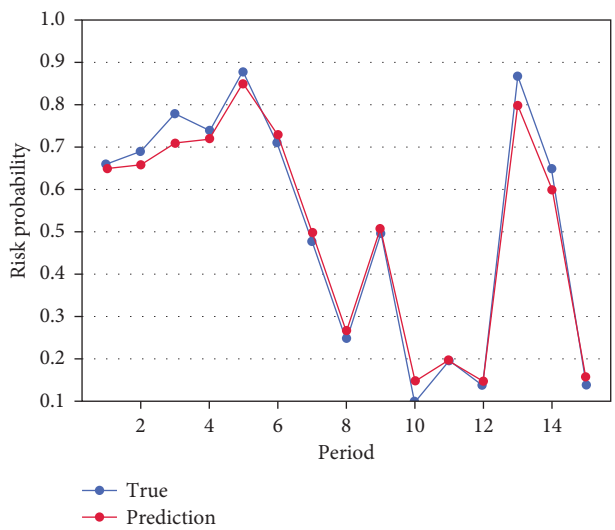


Figure 5: The prediction result of aerobics injury risk.

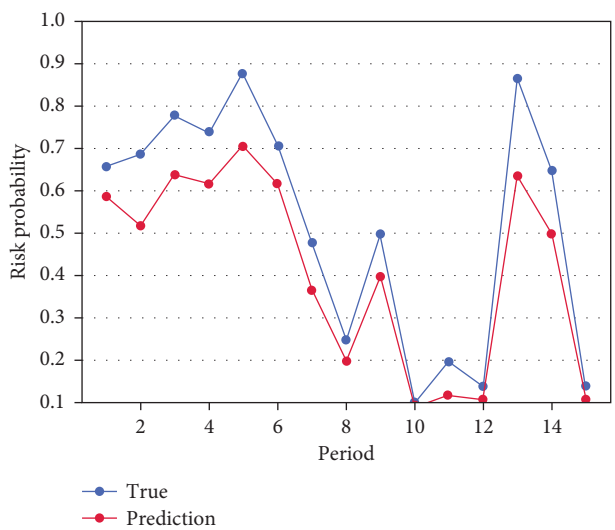


Figure 6: Prediction results of aerobics injury risk without FMS.

4.3. Experimental Results. In this section, experimental results are presented to verify the proposed system. From Figure 5, we can see that the new sports injury risk model proposed in this article based on big data technology and deep learning has achieved excellent performance and accurately predicts the possibility of aerobics sports injury risk. As can be seen from the figure, the prediction curve and the real curve are basically covered, which fully proves the effectiveness of the algorithm in this paper. In order to further demonstrate the advantages of the model in this paper, we also conducted ablation experiments to observe the effect of FMS on the performance of the model.

It can be clearly seen from Figure 6 that if FMS is not used, the prediction accuracy of aerobics injury risk will be greatly reduced. Therefore, this further proves the new sports injury risk model based on big data technology and deep learning proposed in this paper.

5. Conclusion

For the last few years, competitive aerobics have advanced speedily, and the corresponding sports injury risks have gradually increased. A number of studies have shown that due to the characteristics of aerobics itself, difficult movement requirements, fast-paced music accompaniment, and coherent coordinated movements, athletes will suffer sports injuries if they are not paying attention. As due to the rapid development of competitive aerobics, the risk of corresponding sports injuries has gradually increased. In addition, because of the timely detection of the cause of the athlete's injury and timely prevention, it is very important to improve the athlete's technical level and prolong the sports life. Therefore, this paper proposes a new type of sports injury risk model based on big data technology and deep learning. This model can effectively predict the risk of sports injury and has a positive effect on improving the athlete's sports quality and prolonging sports life.

Data Availability

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

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

[1] T. Moshenska and I. Bodrenkova, "Structure and content of competitive group compositions in sports aerobics," *Slobozhanskyi Herald of Science and Sport*, vol. 1, no. 57, pp. 40–42, 2017.

[2] S. Yue, "Image recognition of competitive aerobics movements based on embedded system and digital image processing," *Microprocessors and Microsystems*, vol. 82, Article ID 103925, 2021.

[3] Y. Peng, "Simulation algorithm of difficulty movements in competitive aerobics," in *Proceedings of the International Conference on Cyber Security Intelligence and Analytics*, pp. 614–617, Springer, Haikou, China, February 2020.

[4] H. Dhillon, S. Dhilllon, and M. S. Dhillon, "Current concepts in sports injury rehabilitation," *Indian Journal of Orthopaedics*, vol. 51, pp. 529–536, 2017.

[5] A. Hulme, J. Thompson, R. O. Nielsen, G. J. Read, and P. M. Salmon, "Towards a complex systems approach in sports injury research: simulating running-related injury development with agent-based modelling," *British Journal of Sports Medicine*, vol. 53, no. 9, pp. 560–569, 2019.

[6] International Olympic Committee Injury and Illness Epidemiology Consensus Group, R. Bahr, B. Clarsen et al., "International olympic committee consensus statement: methods for recording and reporting of epidemiological data on injury and illness in sports 2020 (including the STROBE extension for sports injury and illness surveillance (STROBE-SIIS))," *Orthopaedic journal of sports medicine*, vol. 8, no. 2, Article ID 2325967120902908, 2020.

[7] M. Behrens, K. Müller, J. I. Kilb et al., "Modified step aerobics training and neuromuscular function in osteoporotic patients:

a randomized controlled pilot study," *Archives of Orthopaedic and Trauma Surgery*, vol. 137, no. 2, pp. 195–207, 2017.

[8] J. G. Zuzda, M. Pasek, R. Latosiewicz, and M. Bierkus, "Is there any effect of step aerobics training on physical fitness of female students?" in *Proceedings of the International Conference of Computational Methods in Sciences and Engineering*, vol. 2186, no. 1, December 2019, Article ID 080012.

[9] B. Xie, X. Yan, X. Cai, and J. Li, "Effects of high-intensity interval training on aerobic capacity in cardiac patients: a systematic review with meta-analysis," *BioMed Research International*, vol. 2017, Article ID 5420840, 16 pages, 2017.

[10] V. Pasichnyk, M. Pityn, V. Melnyk, I. Karatnyk, A. Hakman, and Y. Galan, "Prerequisites for the physical development of preschool children for the realization of the tasks of physical education," *Physical Activity Review*, vol. 6, pp. 117–126, 2018.

[11] K. Kendellen and M. Camiré, "Examining the life skill development and transfer experiences of former high school athletes," *International Journal of Sport and Exercise Psychology*, vol. 15, no. 4, pp. 395–408, 2017.

[12] R. P. Troiano, E. A. J. Frongillo, J. Sobal, and D. A. Levotsky, "Movements symmetrization–an effective method of injury prevention, health strengthening and prolonged sport careers of athletes," *Movement in Human Life and Health*, vol. 20, p. 35, 2020.

[13] M. Yu, T. Quan, Q. Peng, Yu Xu, and L. Liu, "A model-based collaborate filtering algorithm based on stacked autoencoder," *Neural Computing & Applications*, vol. 1, 2021.

[14] W. D. Chang, L. W. Chou, N. J. Chang, and S. Chen, "Comparison of functional movement screen, star excursion balance test, and physical fitness in junior athletes with different sports injury risk," *BioMed Research International*, vol. 2020, Article ID 8690540, 10 pages, 2020.

[15] R. Arghadeh, A. Letafatkar, and S. S. Shojaeddin, "Relationship between physical fitness and functional movement screening scores in active males: providing preventing model," *Journal of Clinical Physiotherapy Research*, vol. 1, no. 1, pp. 13–20, 2018.

[16] X. Ning, Ke Gong, W. Li, L. Zhang, B. Xiao, and S. Tian, "Feature refinement and filter network for person re-identification," *IEEE Transactions on Circuits and Systems for Video Technology*, 2020.

[17] W. Cai and Z. Wei, "PiiGAN: generative adversarial networks for pluralistic image inpainting," *IEEE Access*, vol. 8, pp. 48451–48463, 2020.

[18] X. Ning, P. Duan, W. Li, and S. Zhang, "Real-time 3D face alignment using an encoder-decoder network with an efficient deconvolution layer," *IEEE Signal Processing Letters*, vol. 27, pp. 1944–1948, 2020.

[19] X. Zhang, Y. Yang, Z. Li, X. Ning, Y. Qin, and W. Cai, "An improved encoder-decoder network based on strip pool method applied to segmentation of farmland vacancy field," *Entropy*, vol. 23, no. 4, p. 435, 2021.

[20] W. Cai and Z. Wei, "Remote sensing image classification based on a cross-attention mechanism and graph convolution," *IEEE Geoscience and Remote Sensing Letters*, vol. 42, 2020.

[21] Y. Xu, D. Zhan, L. Liu, H. Lv, L. Xu, and J. Du, "A privacy-preserving cross-domain healthcare wearables recommendation algorithm based on domain-dependent and domain-independent feature fusion," *IEEE Journal of Biomedical and Health Informatics*, 2021.

[22] R. S. Lloyd, J. L. Oliver, J. M. Radnor, B. C. Rhodes, A. D. Faigenbaum, and G. D. Myer, "Relationships between functional movement screen scores, maturation and physical performance in young soccer players," *Journal of Sports Sciences*, vol. 33, no. 1, pp. 11–19, 2015.

[23] V. B. Marques, T. M. Medeiros, F. de Souza Stigger, F. Y. Nakamura, and B. M. Baroni, "The Functional Movement Screen (FMS) in elite young soccer players between 14 and 20 years: composite score, individual-test scores and asymmetries," *International journal of sports physical therapy*, vol. 12, no. 6, p. 977, 2017.

[24] P. Malliou, A. Gioftsidou, G. Pafis, A. Beneka, and G. Godolias, "Proprioceptive training (balance exercises) reduces lower extremity injuries in young soccer players," *Journal of Back and Musculoskeletal Rehabilitation*, vol. 17, no. 3-4, pp. 101–104, 2004.

[25] A. Bintoudi, M. Goumenakis, and A. Karantanas, "Suprapatellar fat pad inflammation in step aerobics athletes: MR imaging evaluation of two cases," *Open Medicine*, vol. 7, no. 6, pp. 813–816, 2012.

[26] X. Ning, Ke Gong, W. Li, and L. Zhang, "JWSAA: joint weak saliency and attention aware for person re-identification," *Neurocomputing*, vol. 453, 2020.

[27] W. Cai, B. Liu, Z. Wei, M. Li, and J. Kan, "TARDB-Net: triple-attention guided residual dense and BiLSTM networks for hyperspectral image classification," *Multimedia Tools and Applications*, vol. 80, no. 7, pp. 11291–11312, 2021.

[28] Y. Tong, L. Yu, S. Li, J. Liu, H. Qin, and W. Li, "Polynomial fitting algorithm based on neural network," *ASP Transactions on Pattern Recognition and Intelligent Systems*, vol. 1, no. 1, pp. 32–39, 2021.

[29] J. Zhang, X. Jin, J. Sun, J. Wang, and K. Li, "Dual model learning combined with multiple feature selection for accurate visual tracking," *IEEE Access*, vol. 7, pp. 43956–43969, 2019.

[30] H. Khan, M. A. Jan, M. Alam, and W. Dghais, "A channel borrowing approach for cluster-based hierarchical wireless sensor networks," *Mobile Networks and Applications*, vol. 24, no. 4, pp. 1306–1316, 2019.

[31] J. Zhang, Y. Liu, H. Liu, and J. Wang, "Learning local–global multiple correlation filters for robust visual tracking with kalman filter redetection," *Sensors*, vol. 21, no. 4, p. 1129, 2021.

[32] Y. Ding, X. Zhao, Z. Zhang, W. Cai, and N. Yang, "Multiscale graph sample and aggregate network with context-aware learning for hyperspectral image classification," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 14, pp. 4561–4572, 2021.

[33] J. Zhang, W. Wang, C. Lu, J. Wang, and A. K. Sangaiah, "Lightweight deep network for traffic sign classification," *Annals of Telecommunications*, vol. 75, no. 7, pp. 369–379, 2020.

Hindawi
Scientific Programming
Volume 2021, Article ID 9988823, 13 pages
https://doi.org/10.1155/2021/9988823

Research Article

A Novel Adaptive Mutation PSO Optimized SVM Algorithm for sEMG-Based Gesture Recognition

Le Cao,[1] Wenyan Zhang,[1] Xiu Kan,[1,2] and Wei Yao[1]

[1]*School of Electronic and Electrical Engineering, Shanghai University of Engineering Science, Shanghai 201620, China*
[2]*School of Mathematics, Southeast University, Nanjing 210096, China*

Correspondence should be addressed to Xiu Kan; xiu.kan@sues.edu.cn

Received 16 March 2021; Revised 3 June 2021; Accepted 9 July 2021; Published 17 July 2021

Academic Editor: Shah Nazir

In the field of noncontact human-computer interaction, it is of crucial importance to distinguish different surface electromyography (sEMG) gestures accurately for intelligent prosthetic control. Gesture recognition based on low sampling frequency sEMG signal can extend the application of wearable low-cost EMG sensor (for example, MYO bracelet) in motion control. In this paper, a combination of sEMG gesture recognition consisting of feature extraction, genetic algorithm (GA), and support vector machine (SVM) model is proposed. Particularly, a novel adaptive mutation particle swarm optimization (AMPSO) algorithm is proposed to optimize the parameters of SVM; moreover, a new calculation method of mutation probability is also defined. The AMPSO-SVM model based on combination processing is successfully applied to MYO bracelet dataset, and four gesture classifications are carried out. Furthermore, AMPSO-SVM is compared with PSO-SVM, GS-SVM, and BP. The sEMG gesture recognition rate of AMPSO-SVM is 0.975, PSO-SVM is 0.9463, GS-SVM is 0.9093, and BP is 0.9019. The experimental results show that AMPSO-SVM is effective for low-frequency sEMG signals of different gestures.

1. Introduction

Surface electromyography signal is the superposition of action potentials of multiple active motor units in time and space during muscle contraction. Skeletal muscle activity can be measured and predicted by sEMG signal. Different gesture actions have different action potentials. Different gesture actions can be distinguished by analyzing the differences between action potentials [1]. Hand gesture recognition based on sEMG has been widely used in the diagnosis of skeletal muscle system diseases, rehabilitation medicine, biological feedback, and human-computer interaction, especially in the field of sEMG prosthetic control [2–4]. There are many methods and researches on gesture recognition based on sEMG now, which mainly focus on signal feature extraction, feature selection, and classification [5].

Feature extraction is a key step in EMG control. EMG control requires accurate recognition, so every link is designed to minimize the error. The capability of a recognition system is directly related to feature selection [6]. For the EMG control system, how to extract effective feature information and make full use of the characteristics of the signal to characterize the EMG signal is the main problem to be solved. According to the theory of pattern recognition [7–9], the feature extraction methods of EMG signal are mainly as follows: time domain analysis method, frequency domain analysis method, time-frequency analysis method, time series analysis method, and nonlinear dynamics analysis method [10–13]. In motion control, in addition to the accuracy of recognition, the time used for recognition is also essential. The time domain analysis method deals with the EMG signal as a function of time and obtains the statistical characteristics of the signal. Time series refers to a group of data arranged in time order. Time series analysis is a data processing method that uses a parameter model to analyze and process the observed ordered random data. In recent years, the parametric model method has become an important solution for surface electromyography signal analysis because of its high-frequency resolution, among

which AR model method is a typical one. AR model is a linear, second-order moment stationary model, which is suitable for short data analysis and convenient operation, especially for real-time processing of EMG control. The extraction of time domain features and model coefficients is relatively simple and the calculation is relatively fast.

In order to retain more information of EMG signal, the number of channels of EMG signal acquisition instruments reaches 8 or more, which leads to the problem of the high dimension of the signal after feature extraction. In order to reduce the disaster of high dimension feature operation and preserve the global characteristics of the signal, it is necessary to use the dimension reduction algorithm for feature selection. During the past few years, feature selection methods have been under intensive research by using various types of methods, such as principal component analysis (PCA) [14], unsupervised clustering [15], and rough set theory [16]. After the feature information of the signal is obtained, the selection of the classifier is followed to determine the accuracy of signal recognition.

Gesture recognition is one of the algorithms of EMG pattern recognition. EMG pattern recognition is normally divided into three stages [17]. The third stage is classification, which establishes a model for gesture classification. Gesture classification is the switch of motion control. The first step of motion control is to recognize the motion intention in the form of gesture correctly. Gesture recognition is the key step to complete the follow-up work, including force prediction and motion control. On the other hand, one of the key problems in gesture pattern recognition is to identify the motion type of the sEMG signals. Machine learning and deep learning are widely used for data processing. There are many pattern classification methods used in sEMG, including K-nearest neighbor (KNN), artificial neural network (ANN), and support vector machine [18–23].

However, the research on the above features extraction and gesture recognition focuses on the surface muscle electrical signals with high sampling frequency. With the development of Bluetooth communication and low-power embedded technology, wearable sensors begin to play a role in gesture recognition. Gesture recognition based on the wearable sensor is generally robust and accurate and is becoming an important part of the motion control platform [24]. It is very popular to use convenient and comfortable wearable devices for gesture recognition and motion control. For example, MYO bracelet is a noninvasive surface electromyography acquisition method developed by Thalmic laboratory. It provides a more convenient interface for motion control based on human gesture recognition. However, wearable sensors (such as MYO bracelet) have the lower sampling frequency than the widely used 1 kHz EMG acquisition devices. Gesture recognition methods based on high-frequency sEMG signals have been quite mature, but these methods are not suitable for sEMG signals below 1 kHz. Reference [25] uses SVM to classify sEMG signals obtained at 200 Hz and 1 kHz sampling frequency, and the classification accuracy of 1 kHz is obviously higher than 200 Hz. In reference [26], the prosthetic hand of the forearm amputee is controlled by a surface electromyography signal with a frequency of 1 kHz. The motion control based on low sampling rate is still less accurate, and how to maintain high classification accuracy is a key point.

Motivated by the above discussion, an AMPSO optimized SVM algorithm is proposed, which is successfully applied to the recognition of sEMG signals with a low sampling rate. The main contributions of this paper can be summarized from the following three points:

(1) A novel adaptive mutation method is proposed to improve the original PSO algorithm; moreover, a new calculation method of mutation probability is introduced in this paper

(2) A combined gesture recognition method based on a low sampling rate EMG signal is proposed, which has the advantages of simple calculation, fast acquisition speed, and high classification accuracy

(3) A feature selection strategy, which can effectively remove the redundancy between features and improve the accuracy of subsequent sEMG classification algorithms, is designed to solve the problem of high feature dimension

The rest of this paper is arranged as follows. In Section 2, it discusses the feature extraction method of sEMG gesture recognition task. In Section 3, it introduces the AMPSO-SVM algorithm for sEMG pattern classification and demonstrates the process of adaptive PSO mutation in detail. In Section 4, the comparative experiments are developed to demonstrate the effectiveness of the proposed AMPSO-SVM method. Finally, it summarizes the research work in Section 5.

2. Feature Extraction and Selection

As shown in Figure 1, we design a combination method for sEMG gesture classification. This combination is suitable for EMG signals with low sampling rate and high dimension. This proposed algorithm is efficient with few parameters. Moreover, the parameters of SVM are selected automatically by AMPSO to avoid complicated calculations.

Specific time domain and time series features are extracted from the sEMG signals after the denoise through the proposed combination. After feature extraction, the high-dimensional samples are selected by GA and finally sent to AMPSO-SVM classifier for gesture classification.

2.1. Feature Extraction. Feature extraction is an essential step in EMG gesture pattern recognition. The purpose of feature extraction is to distinguish different gesture actions as far as possible, which is directly related to the discrimination ability of the gesture recognition system. Features or attributes of EMG signals representing various upper limb movements are extracted. The feature processing includes three steps: noise reduction, maximum contribution feature determination, and normalization.

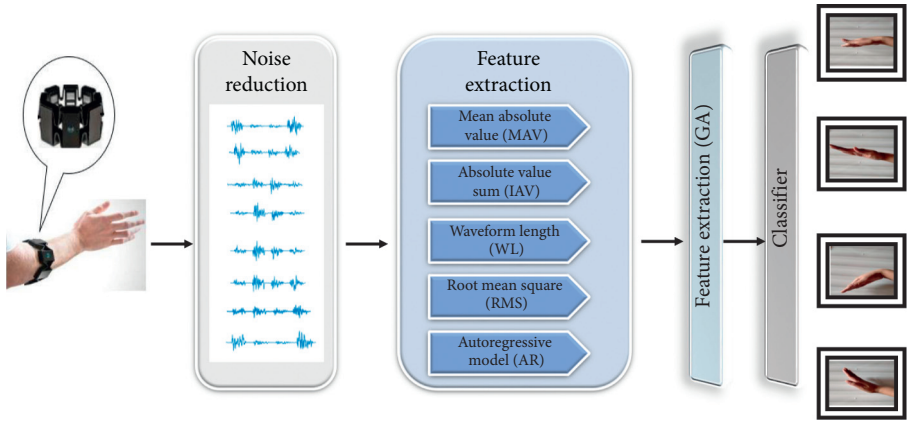


Figure 1: EMG signal processing flow.

2.1.1. Noise Reduction. The practical process of data acquisition will always evolve the noise due to the environmental impact. The existence of noise often conceals the information of the data. Therefore, it is necessary to preprocess the actual signals, and the most important step of preprocessing is noise reduction [27]. DB44 wavelet method is selected to decompose the data into five levels, and the five-layer wavelet coefficients are obtained as the data before feature extraction. Figure 2 shows the waveform of one channel EMG signal after denoising.

2.1.2. Maximum Contribution Feature Determination. The actions in this dataset are hand at rest, hand clenched in a fist, wrist flexion, wrist extension, radial deviations, ulnar deviations, and extended palm (the gesture was not performed by all subjects). The main local structures of the wrist are radial carpal tunnel, ulnar carpal tunnel, and carpal tunnel. The related muscles were flexor carpi ulnaris, flexor carpi radialis, extensor carpi longus radialis, extensor digitorum, and extensor carpi ulnaris. When the muscle contraction force slightly changes, the RMS, MAV, and other time domain characteristics of sEMG signal change greatly, and the AR coefficient can accurately estimate the power spectrum of the signal. The position of the electrode on the muscle surface almost does not affect the coefficient and can be less disturbed by the external environment. Based on the physiological information of muscle, the researchers select the characteristics with strong correlation with muscle contraction and relaxation state in clinical manifestations that include mean absolute value (MAV), integrated absolute value (IAV), waveform length (WL), root mean square (RMS), and autoregressive model (AR) [28–32]. The physiological explanations of the five characteristics which are closely related to muscle contraction are shown in Table 1.

The sliding window is 200 and the step size is 50 (in sampling points). The overlap rate is 75%. The definition formulas for the selected feature are as follows:

$$\begin{aligned}
\mathrm{RMS} &= \sqrt{\frac{1}{M}\sum_{k=1}^{M} x_k^2},\\
\mathrm{MAV} &= \frac{1}{M}\sum_{k=1}^{M} |x_k|,\\
\mathrm{IAV} &= \sum_{k=1}^{M} |x_k|,\\
\mathrm{WL} &= \frac{1}{M}\sum_{k=1}^{M-1} |x_{(k+1)} - x_k|,\\
x_k &= -\sum_{i=1}^{P} a_i x_{k-i} + \omega_k,
\end{aligned} \tag{1}$$

where $x_k\,(k = 1, 2, \ldots, M)$ is the time series of the signals, $P = 7$ is the order of AR model (the best result is the 7th order AR coefficient), a_i is the parameter of AR model, and ω_k is a white noise. In this paper, the Burg method is used to select the parameters of AR model $a_1 - a_7$.

In the process of the experiment, 88-dimensional feature sample data are extracted at once; five features are extracted from each channel of MYO bracelet. In other words, one dimension is extracted from each channel of MAV, IAV, WL, and RMS, and seven dimensions are extracted from each channel of AR model.

2.1.3. Data Normalization. Data normalization is a primary work of data processing. Different evaluation indexes often mean different dimensions and dimension units, which will affect the results of data analysis. To eliminate the dimensional influence between indexes, data normalization processing is needed to solve the comparability between data indexes. In this paper, the [0, 1] interval normalization is set for the EMG dataset. The comparison of classification accuracy with and without normalization is shown in Table 2,

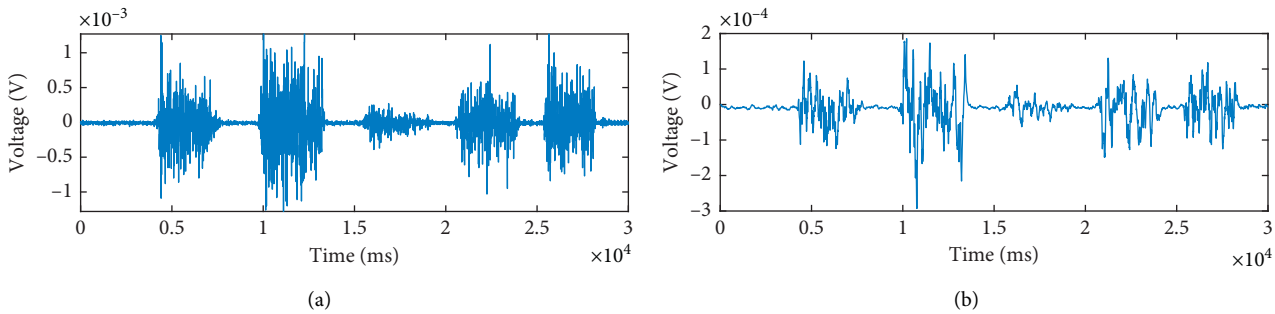


Figure 2: Surface electromyography signal after noise reduction. (a) Original signal. (b) Signal after denoising.

Table 1: Physiological explanation of characteristics.

Features	Physiological explanation
RMS	RMS mainly reflects the type of motor units and the degree of synchronization during muscle activity
MAV	MAV is well matched with the change of muscle tension and also reflects the change of muscle contraction
IAV	IAV is the characteristic after all the amplitude of the signal is converted into the positive value; it is the most intuitive response of muscle contraction force; the larger the IAV is, the greater the muscle contraction force is
WL	WL can reveal the change of amplitude, frequency, and duration of EMG waveform
AR	EMG spectrum changes with the change of muscle contraction state, which will cause the change of AR model coefficient; by monitoring the AR model coefficient, the muscle contraction state can be estimated

Table 2: Normalized comparison.

Selected normalization method	Classification accuracy of test set (%)
[0, 1] interval normalization	94.8148
No normalization	58.9815

where $C = 31$ and $\sigma = 0.01$ are set as in the reference. Table 2 illustrates that the classification accuracy is significantly improved from 58.9815% without the normalization to 98.8148% with [0, 1] interval normalization.

2.2. Feature Selection. If the dimension of the feature matrix is too large, it will put forward higher requirements for subsequent gesture recognition. Therefore, the computation of the feature matrix must be within the computational load of the computer. In this paper, we use a genetic algorithm as the feature selection method, and the recognition rate of k-nearest neighbor classification is used as the fitness function. A genetic algorithm needs to map the solution space to the coding space, and each binary code corresponds to a chromosome. The detailed steps of feature selection are shown as follows.

Initialization: in this paper, the coding length is designed to be 88, each bit in the chromosome represents a channel feature, and the gene of the chromosome is only available for the numbers of "0" or "1." If the corresponding position is "1," then the channel feature is selected; otherwise, the channel feature is discarded.

Calculate fitness function: the input feature matrix is divided into the training set and test set. For each feature combination, the training samples and test samples are reconstructed using the selected features. We use KNN to identify the test samples and obtain the recognition rate R. Let $f(X)$ be the fitness value, count be the number of predicted correct samples, and P be the number of test samples. The fitness value formula can be defined as follows:

$$f(X) = \frac{\text{count}}{P}. \tag{2}$$

Selection operation: the selection operation adopts the proportional selection operator. Set n as the number of individuals and M the sum of fitness. The fitness of all individuals in the population is summed as follows:

$$M = \sum_{k=1}^{n} f(X_k). \tag{3}$$

Meanwhile, the relative fitness value of each individual in the population can be calculated as follows:

$$m_k = \frac{f(X_k)}{M}, \quad k = 1, 2, \ldots, n, \tag{4}$$

where m_k represents the probability that the individual is selected and successfully inherited to the next generation. Roulette function is used to generate a random number between (0, 1) that is used to determine the number of times for each individual. Individuals with higher fitness values are more likely to be selected to pass on their genes to the next generation.

Cross operation: the single point crossover operator is used in the crossover operation process, randomly selects the same position of two individuals among the selected individuals, and exchanges positions according to a certain probability.

Mutation operation: mutation operation adopts the simplest single point mutation operator. Firstly, the mutation points are randomly generated, and then the gene value on the locus is changed according to the position of the mutation point. The binary coding is carried out in this scheme and the corresponding result of mutation operation can only change between "0" and "1."

Let the maximum evolutionary algebra be 40, the maximum population size M be 20, the chromosome length be 88, crossover probability pc = 0.3, and mutation probability pm = 0.03. The pseudocode of the genetic algorithm is shown in Algorithm 1.

The most representative combination of input variables can be selected after the proposed interactive evolution. Using the genetic algorithm to select 88-dimensional feature components, whose feature is retained, finally, only 44-dimensional feature components are retained. After the feature samples are determined, the method of machine learning is used to recognize the gesture.

3. Recognition Model

A 44-dimensional feature vector is obtained after the feature selection process, which is sent to the feature classifier for feature recognition and classification. The dataset used in this study uses MYO bracelet to measure. MYO bracelet has eight channels, and its dimension is relatively high. The sampling frequency of MYO bracelet is 200 Hz, and the number of data samples is small. Therefore, MYO EMG data contains characteristics of few samples and high dimensions. SVM is an effective classifier in solving small sample, nonlinear, and high-dimensional pattern recognition problems [33, 34]. Therefore, SVM and the obtained feature sample have a good combination. In this section, adaptive particle mutation with the design scheme of AMPSO-SVM model is used for sEMG gesture recognition. The flowchart of AMPSO-SVM gesture recognition algorithm is shown in Figure 3.

3.1. SVM. The existing methods for action pattern recognition of sEMG signals are basically based on the nonlinearity and nonstationarity of sEMG signals; however, these characteristics are not combined with the sample size, feature dimension, and the points that may fall into local minima to construct classifiers. SVM is capable of solving nonlinear, small sample, high dimension, and local minimum problems. In this paper, SVM theory is applied to the action pattern recognition of sEMG signal; moreover, a more effective action pattern recognition method of sEMG signal is discussed. SVM algorithm was originally designed for binary classification problems. When dealing with multiclass problems, it is necessary to construct appropriate multiclass classifiers. The method is proposed to design an SVM between any two samples; therefore, the samples of k categories require $k(k-1)/2$ times SVM. When an unknown sample is classified, the category with the most votes is the class of the unknown sample. The original spatial sample data are mapped to the high dimension through a nonlinear transformation with the kernel function for the linear nonseparable problem of SVM. In this paper, the researchers select the following radial basis function (RBF) [35–37] as the SVM kernel function:

$$K(x_i, x) = \exp\left(-\frac{1}{\sigma}\|x_i - x\|^2\right),$$
$$\lambda = \sum_{i=1}^{n} \alpha_i y_i K(x_i, x_j). \tag{5}$$

Therefore, SVM can be transformed into the following optimization problems:

$$\min \frac{1}{2}\sum_{i=1}^{j}\sum_{j=1}^{l} y_j \alpha_j - \sum_{j=1}^{l} \alpha_j, \quad 0 \le \alpha_i \le C, \tag{6}$$

where σ is the kernel function parameter, which implicitly determines the distribution of data after mapping to the new feature space; C is the penalty parameter, which is used to adjust the weight of the margin size and classification accuracy preference in the optimization direction. Penalty parameters C and kernel function parameters σ affect the classification accuracy and the stability of the SVM classifier. Since traditional parameter optimization methods, such as grid search method and gradient descent method, are easy to fall into local optimum in the process of optimization, and manual parameter adjustment causes a large error. Therefore, it is necessary to introduce a suitable parameter search method to find the sum of parameters in SVM. The problem has also been studied by previous researchers [38, 39]. The particle swarm optimization (PSO) algorithm is used in this paper to find the optimal parameters that are capable of improving the accuracy of the model classification and anti-interference ability.

3.2. AMPSO-SVM. In the past several years, biological evolutionary algorithms have been used to optimize the parameters of classifiers including the drosophila algorithm, genetic algorithm, gravity search algorithm, and particle swarm optimization algorithm [40–43]. Particularly, PSO as a global search evolutionary algorithm is proposed by Kennedy and Eberhart in 1995 [44]. Considering that SVM classifier is sensitive to penalty parameter and kernel radius parameter and it is difficult to find the appropriate parameters artificially, particle swarm optimization algorithm is introduced to optimize SVM classifier to find the appropriate parameters. In each iteration, particles update their positions and velocities according to the following equations:

$$v_{id}^{(k+1)} = \omega v_{id}^{k} + c_1 \mathrm{rand}\left(p\mathrm{best}_{id}^{k} - x_{id}^{k}\right) + c_2 \mathrm{rand}\left(g\mathrm{best}_{id}^{k} - x_{id}^{k}\right), \tag{7}$$

$$x_{id}^{(k+1)} = x_{id}^{k} + v_{id}^{(k+1)}, \tag{8}$$

where c_1 and c_2 are acceleration coefficients, adjusting the maximum step size to the global best particle and the

```
Initialization;
k = 0;
  while k to maximum iterations do
    while k = 1 to M do
      Calculate the fitness of population f(X) according to formula (6);
    end while
  while k = 1 to M do
    Select operation to population;
  end while
  while k = 1 to M/2 do
    Crossover operation to population according to pc = 0.3;
  end while
  while k = 1 to M do
    Mutation operation to population according to pm = 0.03;
  end while
  k = k + 1;
end while
Output optimal results.
```

Algorithm 1: The algorithm for feature selection.

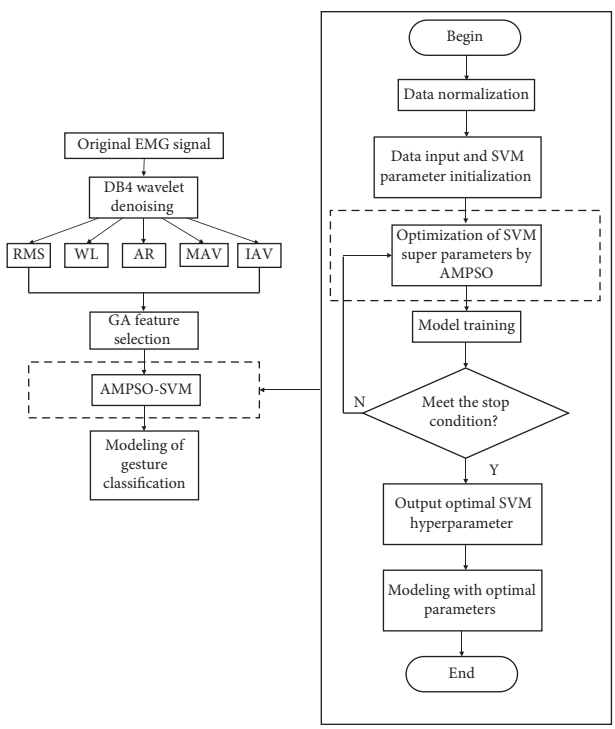


Figure 3: Flowchart of AMPSO-SVM gesture recognition model algorithm.

individual best particle respectively; rand is the random number uniformly distributed in [0, 1]; ω is the inertia weight, maintaining the balance of global and local search capabilities.

Aiming at the premature convergence problem of PSO algorithm, in order to obtain the better SVM parameter selection using PSO algorithm, this paper proposes an adaptive mutation operation after particle position update and defines a new adaptive mutation probability calculation method, so as to obtain an adaptive mutation PSO algorithm.

In the process of iteration, particles continuously fly in the optimal direction. If a particle finds a current optimal position, then other particles quickly approach it, and the speed of all particles drops to zero in a short time. If the optimal particle is a local optimum, then the particle swarm cannot search for the optimal point in the solution space, even though the premature convergence phenomenon occurs, and it is difficult to jump out of the local extremum. In order to deal with the premature convergence phenomenon in the process of PSO, a simple mutation operator is introduced on the basis of the ordinary PSO algorithm. The basic idea is to reinitialize the particles with a certain probability after each update and set the probability as R. In reference [45], a mutation method has been proposed for binary PSO algorithm, $R = 1/N$ is the mutation probability, and N is the data dimension, but, in the process of algorithm iteration, R is a fixed value. Hence, this method cannot adapt to the characteristics of PSO algorithm properly, while PSO is a kind of heuristic algorithm, which needs strong global searchability in the early stage and local searchability in the later stage. Based on the above characteristics, the mutation probability of PSO in the early stage is larger than that in the later stage.

In this paper, we define a novel mutation probability formula as follows:

$$R = 1 - \frac{i}{(N + i)}, \tag{9}$$

where N is the dimension of the data and i is the algebra of the current algorithm iteration. When a new particle is generated, each particle produces a certain mutation probability. According to the analysis, the range of defined mutation probability R gradually decreases from 1 to 0; that is, almost all mutation occurs in the early stage, and few

```
Initialization;
while rand < R do
k = ceil(2 * rand);
end while
   while k == 1 do
      pop(j, k) = (20 − 1) * rand + 1;
end while
   while k == 2 do
      pop(j, k) = (σ_max − σ_min) * rand + σ_min;
   end while
```

Algorithm 2: The algorithm for mutation operation rules.

Initialize:
Set the position parameter v_{id}^{k} according to formula (7);
Set the speed parameter x_{id}^{k} according to formula (8);
while $k = 1$ to maximum iterations do
  Calculate the initial *p*best and *g*best, according to the classification accuracy of SVM cross validation;
  Calculate the mutation probability of the current generation of particles according to (9);
  Update the position x_{id}^{k} and velocity v_{id}^{k} of each particle in the population according to (7)-(8);
  Perform particle mutation according to the rules;
  Update the individual optimal position *pbest* and the population optimal position *gbest* according to the fitness of particles;
  $k = k + 1$
end while
Output optimal structure C and σ.

Algorithm 3: The algorithm for searching parameters.

mutations occur in the late stage, which can better adapt to the characteristics of PSO algorithm.

The adaptive mutation rules are defined in Algorithm 2:

The implementation steps of the proposed AMPSO-SVM algorithm are in Algorithm 3:

Taking the accuracy of the training set in the sense of CV as the fitness function value in the AMPSO, the overall algorithm process of optimizing SVM parameters by AMPSO is shown in Figure 4.

In this paper, adaptive PSO is used to determine the parameters C and σ that are more suitable for SVM, and the proposed AMPSO-SVM will be used to classify four EMG gestures in Section 4.

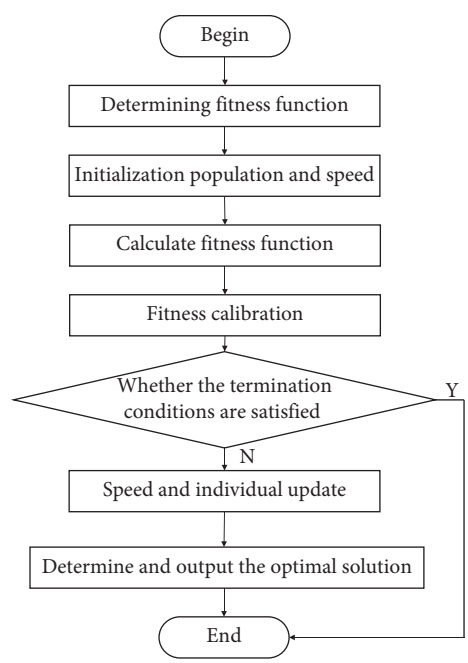


Figure 4: Flowchart of SVM parameters optimized by particle swarm optimization.

4. Experiment and Analysis

4.1. Data Sources. Data utilized in this study were acquired from UCI dataset [46], which adopts a noninvasive method; the dataset is measured by the MYO bracelet worn on the user's forearm. The subjects measured the surface EMG signals of upper limbs by MYO bracelet, viewed the EMG signals by PC software, and then processed the original data by feature extraction, feature selection, and pattern classification.

Now, most of the published datasets use American Delsys electrode and German Otto bock electrode, such as Ninapro [47]. It can be worn above the elbow joint of different people's arms. MYO bracelet has the advantages of stronger mobility, nonrestriction of space and site. The dataset used in this study was collected from 36 healthy volunteers, aged 18 to 41. The bracelet involving eight

sensors is set equidistant around the forearm; moreover, it can simultaneously obtain muscle image signals. Each gesture is executed for 3 seconds, and the pause time between two gestures is 3 seconds.

MYO has eight channels with high dimensions. The sampling frequency of MYO is 200 Hz, and the number of data samples is small. Therefore, MYO has the characteristics of small samples and high dimension. In view of the characteristics of the dataset collected by MYO, it is necessary to study this kind of data.

The experiment evaluates the EMG of these four gestures from two perspectives. The first method is to take all four gesture categories into account to calculate the effectiveness of category prediction. The other method is to consider each category separately and calculate the effectiveness of each category separately. The proposed AMPSO-SVM algorithm is validated and compared with four popular methods including adaptive PSO, ordinary PSO, grid search for SVM parameters (GS-SVM), and BP algorithm.

4.2. Multiclassification. In this section, we randomly select one person from the dataset (http://archive.ics.uci.edu/ml/datasets/EMG+data+for+gestures) to do a classifier verification experiment. Four kinds of gestures were selected from the dataset, such as twist flexion, twist extension, radial deviation, and ulnar deviation. The error classification charts of BP [48], GS-SVM [49], PSO-SVM [50], and AMPSO-SVM are drawn respectively to intuitively understand the four algorithms' classification performance. By observing the sEMG data of 8 channels from the selected person, it seems that the data range of the selected person's ulnar deviation is quite different from the other three actions.

4.2.1. Test Set Classification Result. The classification accuracy of the four recognition methods for the selected person is shown in Figure 5. The results of BP algorithm, GS-SVM algorithm, PSO-SVM algorithm, and AMPSO-SVM algorithm are shown in Figures 5(a)–5(d), respectively.

Figure 5(a) illustrates that BP algorithm is capable of classifying the ulnar deviation and radial deviation; however, the twist flex and twist extension are not identified successfully. As can be seen from Figure 5(a), the recognition rate of BP algorithm is 0.9019. Moreover, the BP algorithm is too complex for inexperienced users due to a large number of layers and parameters.

Then the twist extension can be classified using GS-SVM as shown in Figure 5(b). However, the twist flex has still not been identified in this case. Although the accuracy is improved in this case, the process is expensive for computation. The recognition rate is 0.9093.

In Figure 5(c), it can be seen that the PSO-SVM algorithm is slightly higher than GS-SVM, and the recognition rate is 0.9463. PSO and GA are commonly used as heuristic algorithms. Compared with GA algorithm, PSO has no operation of selection, crossover, and mutation, so its convergence speed is relatively fast. In this paper, PSO algorithm is used to optimize the parameters of SVM. The idea of mutation in GA is also introduced into PSO, which can integrate the advantages of both.

Figure 5(d) illustrates that the classification performance of the proposed AMPSO-SVM has greatly improved compared with the other three recognition methods. The recognition rate of AMPSO-SVM is 0.975. As can be seen from Figure 5, all features have been accurately identified and classified using the proposed AMPSO-SVM algorithm. Moreover, the test set sample labels are predicted as well.

The performance comparison of the four algorithms can be indicated by the recognition rate. The recognition rate of AMPSO-SVM is 0.975, which is the highest from the 0.9019 of BP algorithm, 0.9093 of GS-SVM algorithm, and 0.9463 of PSO-SVM algorithm.

The parameters of the proposed AMPSO-SVM algorithm are set as follows: the maximum number of iterations is 100, the population size is 25 generations, the learning factors $c_1 = 1.6, c_2 = 1.9$, and the mutation probability is $R = 1 - i/N + i$. The curve of the best fitness value with the iteration times in AMPSO-SVM algorithm is shown in Figure 6.

Figure 6 indicates that the adaptive mutation PSO is optimized iteratively and converges in the 13th generation, the SVM parameters can be found as $C = 4.4792$, $\sigma = 0.17226$; moreover, the classification accuracy of four postures reaches 97.5%. Through the comparison of the above results, it can be clearly seen that the parameters found by the adaptive mutation PSO can obtain higher classification accuracy.

4.2.2. Evaluation Indexes. In this section, we will evaluate the effectiveness of the four algorithms in multiclassification tasks based on three evaluation indicators, including accuracy, Kappa coefficient, andMacroF1. Classification accuracy is the most commonly used standard indicator to evaluate the quality of the model. The calculation formula is as follows:

$$\text{accuracy} = \frac{m}{n}, \tag{10}$$

where m is the correct number of samples in the test set and $n$is the total number of samples in the test set.

Kappa coefficient is used to evaluate the reduction of the classification error rate compared with complete random classification. The calculation formula is defined as follows:

$$\text{Kappa} = \frac{\text{accuracy} - p_e}{1 - p_e}, \tag{11}$$

where c is the data category, $t_1, t_2, \ldots, t_c$ are the real sample numbers of each class, $h_1, h_2, \ldots, h_c$ are the predicted sample number of each class, and the calculation formula of p_e can be defined as follows:

$$p_e = \frac{t_1 \times h_1 + t_2 \times h_2 + \cdots + t_c \times h_c}{n \times n}. \tag{12}$$

Kappa coefficient can be divided into five groups to represent different levels of consistency: 0.0 to 0.20 for very low consistency, 0.21 to 0.40 for fair, 0.41 to 0.60 for

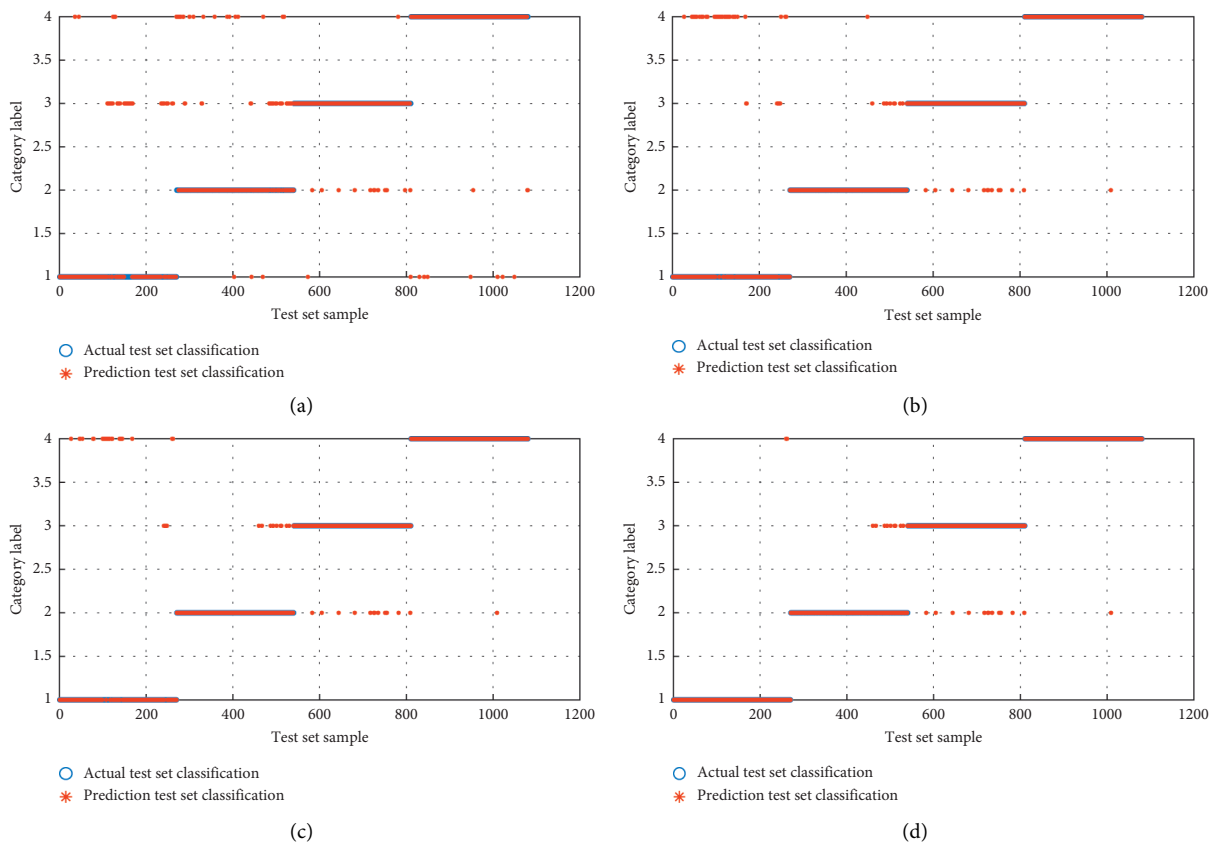


Figure 5: Test set classification graph of the four methods. (a) BP. (b) GS-SVM. (c) PSO-SVM. (d) AMPSO-SVM.

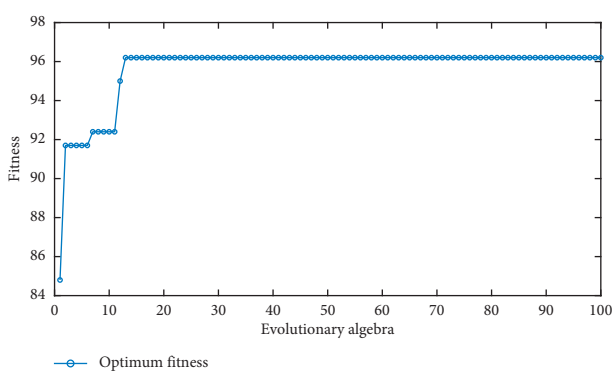


Figure 6: Optimal fitness of adaptive particle mutation PSO.

moderate, 0.61 to 0.80 for substantial, and 0.81 to 1 for almost perfect.

Macro F_1 is the arithmetic mean of each sEMG performance (precision, recall and F_1); G_1, G_2, G_3, and G_4 are used to represent wrist flexion, twist extension, radial deviation, and lar deviation, respectively, and the formula is defined as follows:

$$F_1 = \frac{2\,(\text{precision} \times \text{recall})}{(\text{precision} + \text{recall})}, \tag{13}$$

where F_1 measures the combination of recall and precision of categories G_1, G_2, G_3, and G_4, precision usually is used to measure the precision of categories G_1, G_2, G_3, and G_4, and recall is used to measure recall rate of categories G_1, G_2, G_3, and G_4, respectively.

$$\text{Macro } F_1 = \frac{F_1 + \cdots + F_4}{4}, \tag{14}$$

twhere $F_1 + \cdots + F_4$ is the F_1 from category 1 to category 4.

According to the above three evaluation indicators, it can be seen from Figure 7 that the multiclassification performance of PSO-SVM algorithm is slightly better than GS-SVM algorithm and BP algorithm, while AMPSO-SVM algorithm is obviously the best of all algorithms.

4.3. Classification Performance of Each Action

4.3.1. Different People. In this section, we randomly select five people from the dataset to repeat the experiments five times. Only GS-SVM, PSO-SVM, and AMPSO-SVM are considered in this section to test the classification performance since they are capable of classifying at least three kinds of sEMG. Figure 8 shows the recognition rate of the four actions of these five people. We can see that the recognition rates of the three classifiers for different people are quite different. The results indicate that even though the collected data are significantly affected by different people, the recognition rate of all selected postures using AMPSO-SVM algorithm is still the highest. Moreover, the recognition rates of twist extension and radial deviation using AMPSO-

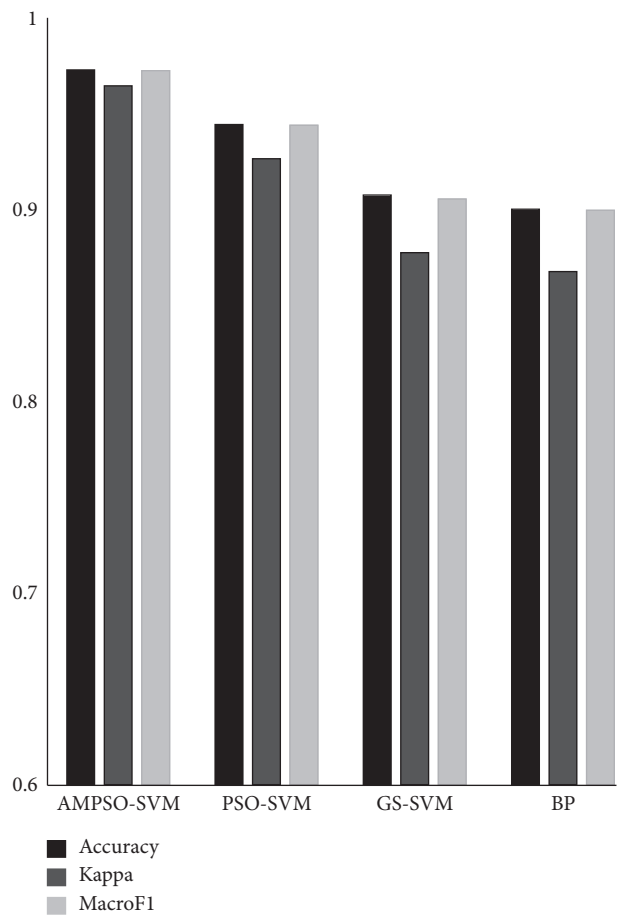


Figure 7: Comparison of four evaluation methods.

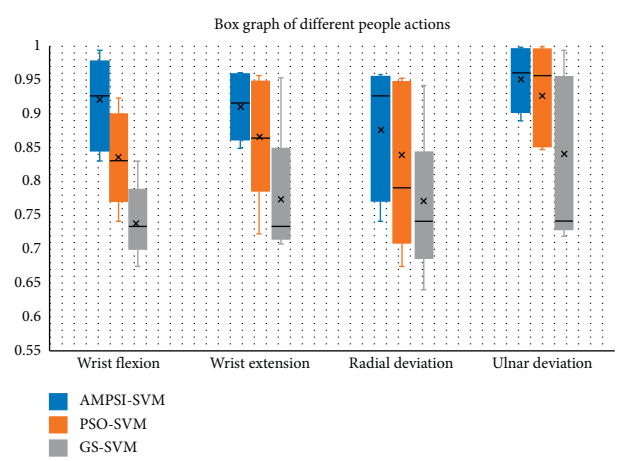


Figure 8: Box graphs of different people.

SVM algorithm are slightly higher than those using PSO-SVM algorithm. Furthermore, the accuracy of the AMPSO-SVM algorithm and PSO-SVM algorithm is closed on recognizing the ulnar deviation.

4.3.2. Average Recognition Rate. The recognition rates of different postures are different due to the different posture performances of subjects. Thus, we calculate the average recognition rate of the five randomly selected people to see the difference among the three classifiers. Figure 9 shows the average recognition rate of the four actions.

The average recognition rate of wrist flexion using AMPSO-SVM algorithm shown in Figure 9(a) is the highest of the three algorithms. The result of AMPSO-SVM is 0.08 higher than that of PSO-SVM. The result in Figure 9(b) is

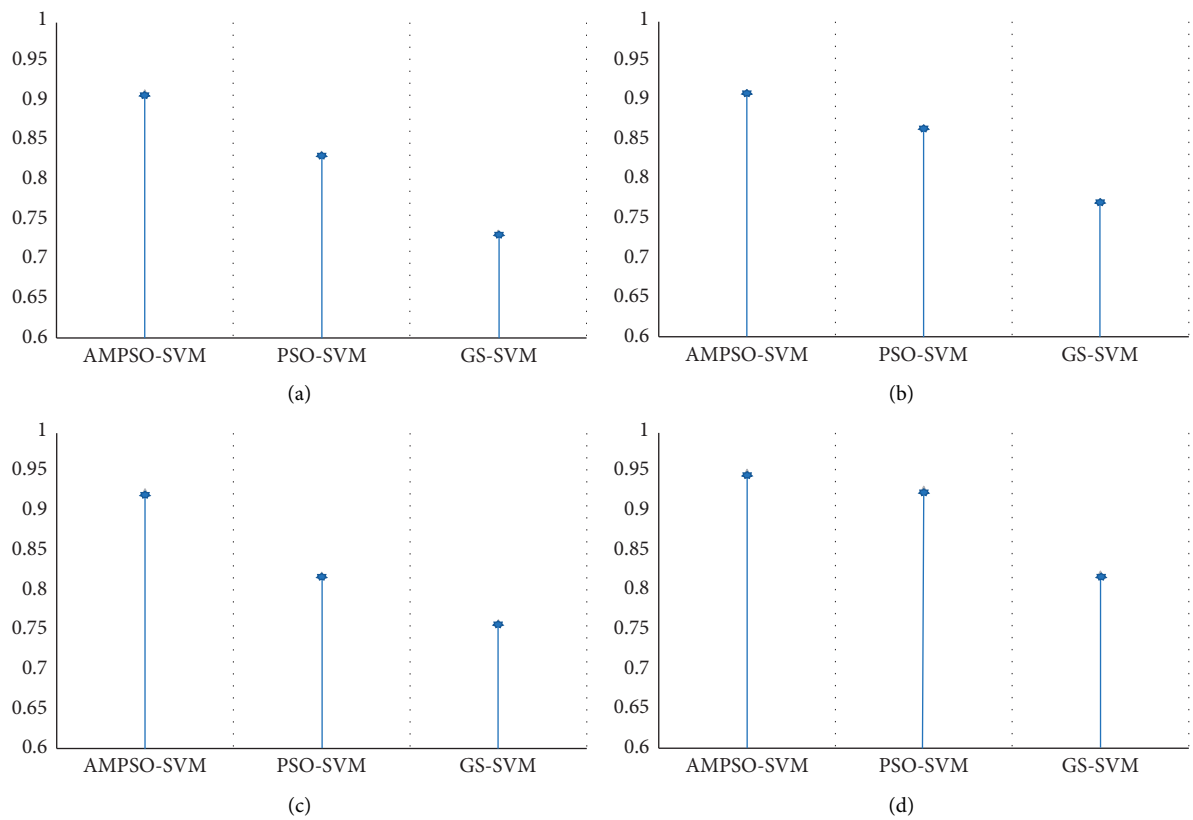


Figure 9: Average recognition rate of four actions. (a) Wrist flexion. (b) Wrist extension. (c) Radial deviation. (d) Ulnar deviation.

similar to the result of wrist extension. The average recognition rate of radial deviation using AMPSO-SVM algorithm is much higher than the value using PSO-SVM algorithm. Moreover, the difference of results between AMPSO-SVM and the other two algorithms used for radial deviation recognition is relatively large as shown in Figure 9(c). Figure 9(d) shows the average recognition rate of ulnar deviation. The result using AMPSO-SVM algorithm is similar to the value using PSO-SVM algorithm, which is higher than the result using GS-SVM. Figure 9 illustrates that the AMPSO-SVM algorithm has the advantage of recognizing the wrist flexion, wrist extension, and radial deviation. However, the capability of PSO-SVM algorithm is similar to the AMPSO-SVM algorithm for recognizing the ulnar deviation.

5. Conclusion

In this paper, a new AMPSO-SVM algorithm is proposed and successfully applied to low sampling rate sEMG gesture recognition. In order to improve the adaptability of gesture recognition technique, a new mutation probability calculation method is defined in particle mutation, which can effectively solve the premature problem of PSO. In the aspect of feature extraction, five features with a high correlation to muscle contraction have been selected. This combination of calculations is simple and fast and can effectively obtain the sEMG signal information. The genetic algorithm is used to solve the problem of high dimension and redundancy of multichannel sEMG signals, which effectively reduces the complexity of subsequent classification. The comparisons of results show that the algorithm is capable of recognizing the sEMG signals with sampling rate accurately. It provides an effective method for gesture recognition of low sampling rate sEMG signal.

Data Availability

The data used in this study can be found at http://archive.ics.uci.edu/ml/datasets/EMG+data+for+gestures.

Conflicts of Interest

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

Acknowledgments

The authors would gratefully acknowledge the support by the National Natural Science Foundation of China (NSFC) as the research program (Grant no. 61703270).

References

[1] J. Ben-Arie, Z. Zhiqian Wang, and S. Rajaram, "Human activity recognition using multidimensional indexing," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 8, pp. 1091–1104, 2002.

[2] F. Riillo, L. R. Quitadamo, F. Cavrini et al., "Optimization of EMG-based hand gesture recognition: supervised vs. unsupervised data preprocessing on healthy subjects and transradial amputees," *Biomedical Signal Processing and Control*, vol. 14, no. 1, pp. 117–125, 2014.

[3] D. A. Zlotolow and S. H. Kozin, "Advances in upper extremity prosthetics," *Hand Clinics*, vol. 28, no. 4, pp. 587–593, 2012.

[4] L. Zhang, W. Qi, Y. Hu, and Y. Chen, "Disturbance-observer-based fuzzy control for a robot manipulator using an EMG-driven neuromusculoskeletal model," *Complexity*, vol. 2020, Article ID 8814460, 10 pages, 2020.

[5] X. Kan, D. Yang, L. Cao et al., "A novel PSO-based optimized lightweight convolution neural network for movements recognizing from multichannel surface electromyogram," *Complexity*, vol. 2020, no. 4, 15 pages, Article ID 6642463, 2020.

[6] A. Phinyomark, P. Phukpattaranont, and C. Limsakul, "Feature reduction and selection for EMG signal classification," *Expert Systems with Applications*, vol. 39, no. 8, pp. 7420–7431, 2012.

[7] M. Nebojša, M. Dimitrije, K. Gunter et al., "Vector autoregressive hierarchical hidden Markov models for extracting finger movements using multichannel surface EMG signals," *Complexity*, vol. 2018, no. 1, 12 pages, Article ID 9728264, 2018.

[8] T. Matsubara and J. Morimoto, "Bilinear modeling of EMG signals to extract user-independent features for multiuser myoelectric interface," *IEEE Transactions on Biomedical Engineering*, vol. 60, no. 8, pp. 2205–2213, 2013.

[9] J. Liu and P. Zhou, "A novel myoelectric pattern recognition strategy for hand function restoration after incomplete cervical spinal cord injury," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 21, no. 1, pp. 96–103, 2013.

[10] B. Hudgins, P. Parker, and R. N. Scott, "A new strategy for multifunction myoelectric control," *IEEE Transactions on Biomedical Engineering*, vol. 40, no. 1, pp. 82–94, 1993.

[11] D. Graupe and W. K. Cline, "Functional separation of EMG Signals via ARMA identification methods for prosthesis control purposes," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. SMC-5, no. 2, pp. 252–259, 1975.

[12] P. C. Doerschuk, D. E. Gustafon, and A. S. Willsky, "Upper extremity limb function discrimination using EMG signal analysis," *IEEE Transactions on Biomedical Engineering*, vol. 30, no. 1, pp. 18–29, 2007.

[13] R. Chowdhury, M. Reaz, M. Ali, A. Bakar, K. Chellappan, and T. Chang, "Surface electromyography signal processing and classification techniques," *Sensors*, vol. 13, no. 9, pp. 12431–12466, 2013.

[14] Y. Deng, F. Gao, and H. Chen, "Angle estimation for knee joint movement based on PCA-RELM algorithm," *Symmetry*, vol. 12, no. 1, p. 130, 2020.

[15] F. Jahanmiri-Nezhad, P. E. Barkhaus, W. Z. Rymer, and P. Zhou, "Spike sorting paradigm for classification of multichannel recorded fasciculation potentials," *Computers in Biology and Medicine*, vol. 55, pp. 26–35, 2014.

[16] Y. Qian, J. Liang, and W. Pedrycz, "Positive approximation: an accelerator for attribute reduction in rough set theory," *Artificial Intelligence*, vol. 174, no. 9, pp. 597–618, 2010.

[17] M. Simão, N. Mendes, O. Gibaru, and P. Neto, "A review on electromyography decoding and pattern recognition for Human-Machine interaction," *IEEE Access*, vol. 7, pp. 39564–39582, 2019.

[18] Y. Wu, S. Liang, L. Zhang, Z. Chai, C. Cao, and S. Wang, "Gesture recognition method based on a single-channel sEMG envelope signal," *EURASIP Journal on Wireless Communications and Networking*, vol. 35, no. 2018, 2018.

[19] Y. M. Aung and A. Al-Jumaily, "sEMG based ANN for shoulder angle prediction," *Procedia Engineering*, vol. 41, pp. 1009–1015, 2012.

[20] L. Bruzzone and D. F. Prieto, "A technique for the selection of kernel-function parameters in RBF neural networks for classification of remote-sensing images," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 37, no. 2, pp. 1179–1184, 2002.

[21] Y. Rong, D. Hao, X. Han, Y. Zhang, J. Zhang, and Y. Zeng, "Classification of surface EMGs using wavelet packet energy analysis and a genetic algorithm-based support vector machine," *Neurophysiology*, vol. 45, no. 1, pp. 39–48, 2013.

[22] R. M. Singh, V. Ahlawat, S. Chatterji, and A. Kumar, "Comparative analysis of SVM and ANN classifier based on surface EMG signals for elbow movement classification," *Journal of Interdisciplinary Mathematics*, vol. 23, no. 1, pp. 153–161, 2020.

[23] S. Ranaldi, C. De Marchis, and S. Conforto, "An automatic, adaptive, information-based algorithm for the extraction of the sEMG envelope," *Journal of Electromyography and Kinesiology*, vol. 42, pp. 1–9, 2018.

[24] D. Jiang, G. Li, Y. Sun, J. Kong, and B. Tao, "Gesture recognition based on skeletonization algorithm and CNN with ASL database," *Multimedia Tools and Applications*, vol. 78, no. 21, pp. 29953–29970, 2019.

[25] A. Phinyomark, R. N. Khushaba, and E. Scheme, "Feature extraction and selection for myoelectric control based on wearable EMG sensors," *Sensors*, vol. 18, no. 5, p. 1615, 2018.

[26] H. Kawasaki, M. Kayukawa, H. Sakaeda, and T. Mouri, "Learning system for myoelectric prosthetic hand control by forearm amputees," in *Proceedings of the 23rd IEEE International Symposium on Robot and Human Interactive Communication*, pp. 899–904, Edinburgh, UK, August 2014.

[27] R. Merletti, G. L. Cerone, and T. Tutorial, "Tutorial. Surface EMG detection, conditioning and pre-processing: best practices," *Journal of Electromyography and Kinesiology*, vol. 54, Article ID 102440, 2020.

[28] G. Y. Yang, S. X. Wang, and Y. Chen, "SEMG analysis basing on AR model and Bayes taxonomy," *Applied Mechanics and Materials*, vol. 44–47, pp. 3355–3359, 2010.

[29] A. Shakeel, T. Tanaka, and K. Kitajo, "Time-series prediction of the oscillatory phase of EEG signals using the least mean square algorithm-based AR model," *Applied Sciences*, vol. 10, no. 10, Article ID 3616, 2020.

[30] D. C. Tkach, A. J. Young, L. H. Smith, E. J. Rouse, and L. J. Hargrove, "Real-time and offline performance of pattern recognition myoelectric control using a generic electrode grid with targeted muscle RP," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 22, no. 4, pp. 727–734, 2014.

[31] A. H. Al-Timemy, G. Bugmann, J. Escudero, and N. Outram, "Classification of finger movements for the dexterous hand prosthesis control with surface electromyography," *IEEE Journal of Biomedical and Health Informatics*, vol. 17, no. 3, pp. 608–618, 2013.

[32] G. R. Naik and H. T. Nguyen, "Nonnegative matrix factorization for the identification of EMG finger movements: evaluation using matrix analysis," *IEEE Journal of Biomedical and Health Informatics*, vol. 19, no. 2, pp. 478–485, 2015.

[33] Y. Li, W. Zhang, Q. Zhang, and N. Zheng, "Transfer learning-based muscle activity decoding scheme by low-frequency sEMG for wearable low-cost application," *IEEE Access*, vol. 9, pp. 22804–22815, 2021.

[34] D. Xiong, D. Zhang, X. Zhao, and Y. Zhao, "Deep learning for EMG-based human-machine interaction: a review," *IEEE/CAA Journal of Automatica Sinica*, vol. 8, no. 3, pp. 512–533, 2021.

[35] S. Guo, M. Pang, B. Gao, H. Hirata, and H. Ishihara, "Comparison of sEMG-Based feature extraction and motion classification Methods for upper-limb movement," *Sensors*, vol. 15, no. 4, pp. 9022–9038, 2015.

[36] A. Hekmatmanesh, H. Wu, and F. Jamaloo, "A combination of CSP-based method with soft margin SVM classifier and generalized RBF kernel for imagery-based brain computer interface applications," *Multimedia Tools and Applications*, vol. 79, no. 25, pp. 17521–17549, 2020.

[37] H. A. Jaber, T. Rashid, and L. Fortuna, "Online myoelectric pattern recognition based on hybrid spatial features," *Biomedical Signal Processing and Control*, vol. 66, no. 5, pp. 1–11, 2021.

[38] Y. Song, D. Wu, W. Deng et al., "MPPCEDE: multi-population parallel co-evolutionary differential evolution for parameter optimization," *Energy Conversion and Management*, vol. 228, no. 2, Article ID 113661, 2021.

[39] D. Wu, H. Liu, J. Xu et al., "An improved quantum-inspired differential evolution algorithm for deep belief network," *IEEE Transactions on Instrumentation and Measurement*, vol. 69, no. 10, pp. 7319–7326, 2020.

[40] L. Shen, H. Chen, Z. Yu et al., "Evolving support vector machines using fruit fly optimization for medical data classification," *Knowledge-Based Systems*, vol. 96, pp. 61–75, 2016.

[41] D. Zeng, S. Wang, Y. Shen, and C. Shi, "A GA-based feature selection and parameter optimization for support tucker machine," *Procedia Computer Science*, vol. 111, pp. 17–23, 2017.

[42] D. Wu, K. Warwick, Z. Ma et al., "Prediction of Parkinson's disease tremor onset using a radial basis function neural network based on particle swarm optimization," *International Journal of Neural Systems*, vol. 20, no. 2, pp. 109–116, 2010.

[43] B. Iason, N. E. Krausz, A. M. Simon et al., "Decoding the grasping intention from electromyography during reaching motions," *Journal of NeuroEngineering and Rehabilitation*, vol. 15, no. 1, pp. 1–13, 2018.

[44] J. Kennedy and R. C. Eberhart, "Particle swarm optimization," in *Proceedings of the IEEE International Joint Conference on Neural Networks*, vol. 4, pp. 1942–1948, Perth, Australia, November 1995.

[45] Y. Zhang, S. Wang, P. Phillips, and G. Ji, "Binary PSO with mutation operator for feature selection using decision tree applied to spam detection," *Knowledge-Based Systems*, vol. 64, pp. 22–31, 2014.

[46] L. Sergey, K. Nadia, and K. Innokentiy, "Latent factors limiting the performance of sEMG-interfaces," *Sensors*, vol. 18, no. 4, Article ID 1122, 2018.

[47] M. Atzori and H. Müller, "The Ninapro database: a resource for sEMG naturally controlled robotic hand prosthetics," *IEEE Engineering in Medicine & Biology Society*, vol. 2015, Article ID 7320041, 7154 pages, 2015.

[48] R. Ma, L. Zhang, G. Li, D. Jiang, S. Xu, and D. Chen, "Grasping force prediction based on sEMG signals," *Alexandria Engineering Journal*, vol. 59, no. 3, pp. 1135–1147, 2020.

[49] M. A. Oskoei and H. Huosheng Hu, "Support vector machine-based classification scheme for myoelectric control applied to upper limb," *IEEE Transactions on Biomedical Engineering*, vol. 55, no. 8, pp. 1956–1965, 2008.

[50] A. Subasi, "Classification of EMG signals using PSO optimized SVM for diagnosis of neuromuscular disorders," *Computers in Biology and Medicine*, vol. 43, no. 5, pp. 576–586, 2013.

Hindawi
Scientific Programming
Volume 2021, Article ID 6971906, 8 pages
https://doi.org/10.1155/2021/6971906

Research Article

Flipped Classroom Design of College Ideological and Political Courses Based on Long Short-Term Memory Networks

Fei Su[1] and Zhe Fan[2]

[1]*College of Sport and Health Science, Xi'an Physical Education University, Xi'an 710000, Shaanxi, China*
[2]*Shijiazhuang Vocational College of Scientific and Technical Engineering, Zhengding, Shijiazhuang 050800, Hebei, China*

Correspondence should be addressed to Zhe Fan; fanzhezhengding@163.com

Received 9 April 2021; Accepted 17 June 2021; Published 13 July 2021

Academic Editor: Shah Nazir

The advancement and rising of information technology have promoted the flipped classroom in an effective way. It flips knowledge transfer and knowledge internalization from two levels of teaching structure and teaching process, reversing the traditional teaching knowledge transfer in class and knowledge deepening after class from time and space. Although the use of flipped classrooms in ideological and political theory courses is relatively uncommon in colleges and universities, realistic teaching and related study findings in some colleges and universities provide some reference value for the use of flipped classrooms in ideological and political theory courses. As a result, the short- and long-time memory network-based flipped classroom design algorithm for ideological and political courses in colleges and universities has a wide range of applications. A neural network prediction model based on a hybrid genetic algorithm is developed in this paper. The hybrid genetic algorithm is used in this model to determine the optimal dropout probability and the number of cells in the hidden layer of the neural network. The hybrid genetic algorithm will lengthen the memory neural network to predict the teaching quality of root mean square error between real value and predictive value as a fitness function, in the process of optimization, genetic algorithm convergence to the local optimal solution of the area.

1. Introduction

Flipped classroom [1–3] benefits from the expansion of information technology. It reverses the transfer of knowledge in the classroom and the deepening of knowledge under the classroom in time and space and realizes the transfer of knowledge and knowledge from the two levels of teaching structure and teaching process. Before class, students complete the cognition of basic knowledge through online self-learning [4, 5]. Teachers in the classroom organize classroom activities to further consolidate and sublimate knowledge and achieve a teaching form of deep learning. Some universities' realistic teaching and related study findings provide some reference value for the use of flipped classrooms in ideological and political theory courses. The interpretation and perception of basic theoretical knowledge is not the only way to enhance the teaching impact of ideological and political theory courses [6, 7], but more importantly, the contradiction transformation of internalization and externalization in behavior. This is a realistic problem faced by teachers of ideological and political theory. It is also a frequently discussed and important topic in the field of ideological and political theory teaching and research [8, 9].

Take online teaching [10, 11] as a useful supplement to traditional teaching methods, continuously innovate online teaching methods and combine online teaching with traditional teaching methods. The characteristic of flipped classroom teaching is to transfer knowledge in the classroom in the traditional teaching model and to reverse the time and space of knowledge deepening in the two parts outside the class, to realize the flip of the structure and process of knowledge transfer and knowledge internalization. In recent years, flipped classroom teaching has gradually become a focus of education and teaching reforms at all stages. Research on flipped classrooms in ideological and political theory

courses in colleges and universities has increased, but there is not much research on teaching reform in this field [12].

This paper begins by examining the empirical connotation of flipped classrooms and then examines the relevance and viability of using flipped classrooms in college ideological and political theory classes [13], as well as the theoretical foundation and core concepts of their use, before using "basic" courses as an example to present ideological and political theory in a concrete manner. A long short-term memory neural network [14–17] prediction model optimized by hybrid genetic algorithm is built in the design and implementation of the flipped classroom teaching process of theoretical courses and is based on neural network technology [18–21]. It can predict and evaluate the teaching quality of flipped classrooms in college ideological and political courses. The key offerings of the proposed study are as follows:

(1) To employ neural network technology to develop a long- and short-term memory neural network prediction model focused on hybrid genetic algorithm optimization that can predict and assess the teaching content of flipped college ideological and political courses

(2) The root mean square error between the true value and the expected value of the teaching output predicted by the long- and short-term memory neural network is used as the fitness function in this paper to automatically find the optimal dropout probability and the number of hidden layer units of the neural network

(3) To use a sequential quadratic programming algorithm to advance local search, rapidly and precisely optimize dropout probability and the number of hidden layer units, and input the obtained optimal parameters into a long- and short-term memory network to predict the teaching output of college ideological and political courses flipped classroom and evaluation

2. Background

2.1. The Connotation of Flipped Classroom. The partnership between education and teaching and information technology is becoming closer as network information technology develops at a rapid pace. The flipped classroom was born in the development of multimedia technology. Before the term "flipped classroom" officially appeared, the similar term "Inverted Classroom" existed, but the two were not completely equivalent. The flipped classroom really appeared in 2007 when two chemistry teachers at Woodland Park High School in the United States reorganized the teaching classroom to help students who were absent. Teachers make teaching videos and teaching materials by themselves and arrange for students to watch them at home. The assignments will be completed under the guidance of teachers within the specified time in class. The period from 2008 to 2012 is the stage of practice and development and innovation of flipped classrooms. In 13 years, it has moved from primary and secondary schools to colleges and universities. Teachers in American colleges and universities began to use flipped courses for teaching [22] and achieved some good results.

With the application and development of flipped classrooms in university teaching, related research on flipped classrooms in ideological and political theory courses continues to emerge. The analysis of its meaning still needs to be traced from the existing research results of the academic community. In the existing research results, there are different understandings and interpretations of flipped classrooms:

(1) A flipped classroom is a teaching model that consists of a structure with set procedures and predictable classroom teaching activities implemented using modern information technology and driven by the "student-centered" educational theory. Lai Huiming believes that "flipped classroom is an innovative teaching model held by recent information technology that is, the conventional teaching method is inverted in the classroom. Independent learning, a teaching style of collaborative learning between teachers and students in the classroom, is carried out using the provided teaching tools.

(2) Flipped classroom is a teaching method. It emphasizes deep learning through classroom communication and learning, not just on the cognitive level.

(3) Flipped classroom is a teaching structure. Scholars regard the operation process as the core foothold of a flipped classroom and emphasize the external form as the biggest feature to highlight the characteristics of the flipped classroom.

(4) A flipped classroom is a type of instruction. It focuses on a state of interactive contact created during the teaching process by teachers and students. The flipped classroom is a method of instruction in which students use interactive learning materials developed and generated by teachers for self-directed learning prior to class, then actively participate in interactive activities between classmates and teachers, and complete classroom exercises in the classroom.

2.2. Features of Flipped Classroom. The flipped classroom is characterized by the reversal of teacher and student positions, as well as the transformation of teaching concepts, the alteration of classroom layout, and the creation of new classroom activities:

(1) A shift in teaching philosophy from "teacher-centered" to "student-centered" is needed before flipped classroom teaching can be implemented

(2) The guarantee for the implementation of flipped classroom teaching is the transformation of the teaching system from "teaching first and then learning" to "learning first and then teaching, teaching through learning"

(3) The trick to using flipped classroom teaching is to focus on the material and create a new form of teaching classroom activity called "leading-the-main body mix"

2.3. The Dilemma of Flipped Classroom in Ideological and Political Theory Course. Educators, instructional objects, and teaching facilities are the three main components of the teaching process, and they all play a role in ensuring that teaching methods are used effectively. The use of any teaching method must consider the coordination between them. Whether the three are coordinated directly affects the teaching effect of the teaching method. The ideological and political theory class flipped classroom has many influences during the implementation process. The use of classrooms in ideological and political theory courses has not been widely recognized and promoted. The most important influencing factor is the insufficient ability to use information technology.

3. Methodology

This section explains the hybrid genetic algorithm-optimized LSTM [23] prediction model, as well as the concepts of the LSTM and the flipped classroom assessment model of college ideological and political courses.

3.1. LSTM

3.1.1. Basic Principles of LSTM. The hidden layer of Simple Recurrent Neural Network has only one state h, which is suitable for short-term input, but it does not work well for long-term input. To improve this situation, LSTM RNN is used to save a longer-term state by adding a state c. The newly added state c is called the unit state. At time t, LSTM RNN has three inputs: the input value x_t of the network at the current time, the output value h_{t-1} of the LSTM RNN at the previous time, and the unit state c_{t-1} at the previous time. The LSTM RNN has two outputs: the output value of the LSTM RNN at the current time h_t and the current unit state c_t. The time dimension diagram of LSTM RNN is shown in Figure 1.

The calculation equations of the input gate, forget gate, and output gate of LSTM are as follows:

$$\begin{aligned} i_t &= \sigma_g\left(W_i x_t + U_i h_{t-1} + b_i\right), \\ f_t &= \sigma_g\left(W_f x_t + U_f h_{t-1} + b_f\right), \\ o_t &= \sigma_g\left(W_o x_t + U_o h_{t-1} + b_o\right), \end{aligned} \tag{1}$$

where $x_t \in R^d$ is the input vector of the LSTM unit; $i_t \in R^h$, $f_t \in R^h$, and $o_t \in R^h$ are the activation vectors of the input gate, forget gate, and output gate, respectively; and $h_t \in R^h$ represents the hidden state vector, which is also called the output vector of the LSTM unit. $W \in R^{h\times d}$, $U \in R^{h\times h}$, and $b \in R^h$ are the weight matrix and bias vector that need to be learned during training.

Calculate the activation vector c_t' used to describe the current unit input based on the previous output and this input. The calculation equation is as follows:

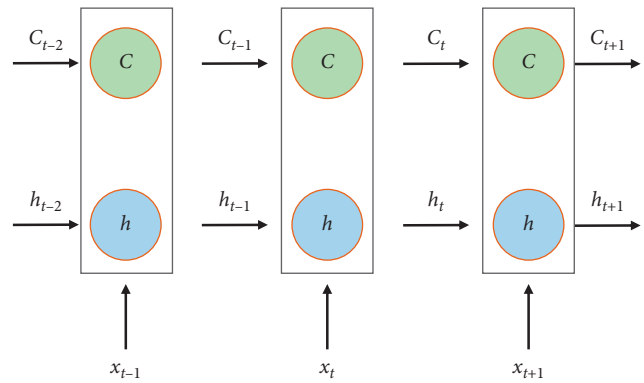


Figure 1: Time dimension of RNN.

$$c_t' = \sigma_h\left(W_c x_t + U_c h_{t-1} + b_c\right), \tag{2}$$

where σ_h represents the hyperbolic tangent tanh activation function. The current cell state c_t can be obtained by inputting the activation vector c_t' of the current cell and the cell state c_{t-1} at the previous moment:

$$c_t = f_t \circ c_{t-1} + i_t \circ c_t'. \tag{3}$$

The operator o represents the Hadamard product (elementwise product). Finally, the output of the LSTM unit can be obtained as

$$h_t = o_t \circ \sigma_h\left(c_t\right). \tag{4}$$

LSTM can be trained on a set of sequence data in a supervised manner, through time backpropagation to calculate the gradient required in the optimization process, to change each weight of the LSTM network to make the error of the corresponding weight (in the output layer of the LSTM network); the derivative is proportional.

The sigmoid activation function and its derivative need to be used in the backpropagation process, and the hyperbolic tangent activation function and its derivative are used at the same time, which are defined as follows:

$$\begin{aligned} \sigma_g'(z) &= y(1-y), \\ \sigma_h(z) &= y = \frac{e^z - e^{-z}}{e^z + e^{-z}}, \\ \sigma_h'(z) &= 1 - y^2. \end{aligned} \tag{5}$$

3.1.2. Dropout. Generally, if the algorithm wants to have a good performance on the training set and the test set, overfitting should be avoided. For this reason, a regularization method needs to be added to the algorithm. Dropout is a regularization method with a better effect. Dropout refers to a regularization method that randomly eliminates or retains a unit in the neural network during neural network training.

When using a trained model to make predictions [24], it is necessary to input test data and perform forward propagation.

At this time, the output value of the layer that has undergone dropout processing needs to be multiplied by the dropout probability during training on the basis of the original output. Although the network discards units with a certain probability through dropout during training, all units are still used during recognition. Therefore, the number of units with output values will increase the inverse times of the dropout probability. Since the units are connected by weight coefficients, they need to be multiplied by the probability. The comparison chart before and after the neural network application dropout is shown in Figures 2 and 3.

Standard backpropagation learning establishes a fragile collaborative adaptation mechanism. This mechanism is suitable for training datasets, but not for data that is not involved in training. Dropout breaks this collaborative adaptation mechanism by randomly hiding any hidden layer unit. It makes the neural network not excessively dependent on a specific input feature, to achieve the purpose of suppressing overfitting.

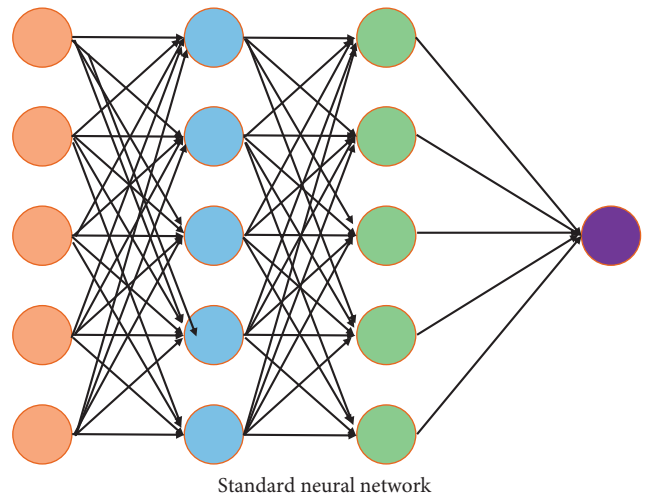


Figure 2: Standard neural network.

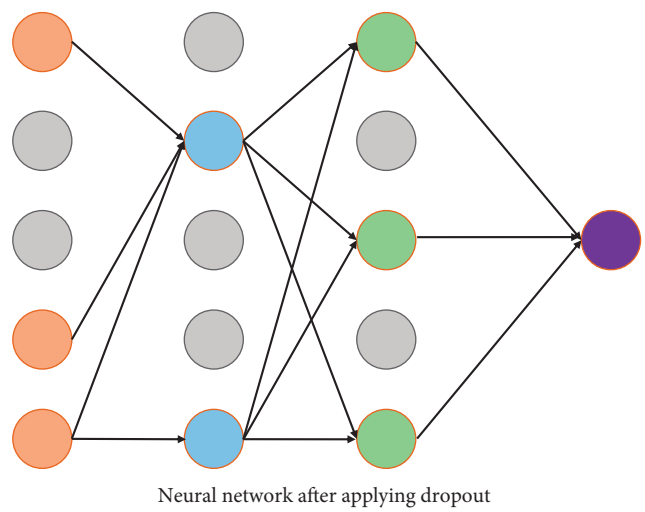


Figure 3: Neural network after applying dropout.

3.1.3. Adam. The optimization algorithm of deep learning is crucial to the training of the model. It adjusts the learning rate of each weight in the neural network through the moment estimation of the gradient. The algorithm uses the powerful function of the adaptive learning rate method to design a separate learning rate for different parameters.

In order to estimate the momentum, Adam uses the exponential moving average to calculate the gradient of the current minibatch:

$$\begin{aligned} m_t &= \beta_1 m_{t-1} + (1-\beta_1) g_t, \\ v_t &= \beta_2 v_{t-1} + (1-\beta_2) g_t^2, \end{aligned} \tag{6}$$

where m and v are moving averages, g represents the gradient of the current minibatch, β_1 and β_2 are the newly introduced hyperparameters of the algorithm, and their default values are 0.9 and 0.999, respectively.

Due to the influence of the moving average, the estimate obtained above is biased, and the revised estimate is

$$\begin{aligned} m_t' &= \frac{m_t}{1-\beta_1} n, \\ v_t' &= \frac{v_t}{1-\beta_2} n, \\ w_t &= w_{t-1} - \eta \frac{m_t'}{\sqrt{v_t'} + \varepsilon}. \end{aligned} \tag{7}$$

3.2. Hybrid Genetic Algorithm

3.2.1. Genetic Algorithm. Choosing the optimal parameters for deep learning tasks is very challenging. Improper selection of the initial values of the learning parameters may result in noise in the data or weak learning ability of the learning algorithm used. Therefore, genetic algorithms can be used to automatically find the optimal learning parameters. The genetic algorithm diagram is shown in Figure 4.

To select the best individual, a fitness function needs to be used. The result of the fitness function represents the quality of the solution, that is, the fitness of the individual. The higher the fitness, the better the quality of the solution.

3.2.2. Sequential Quadratic Programming. SQP combines two basic algorithms for solving nonlinear optimization problems: the active set method and the Newton method. It has a solid theoretical foundation and is designed to solve large-scale technical-related problems. It provides powerful algorithm tools. The constrained nonlinear optimization problem can be written as

$$\begin{cases} \min \quad f(x), \\ h(x) = 0, \\ g(x) \le 0. \end{cases} \tag{8}$$

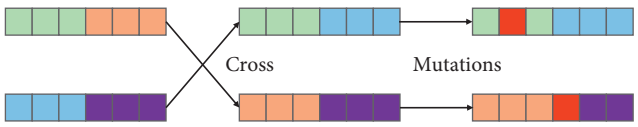


Figure 4: Genetic algorithm.

The Lagrangian equation for nonlinear optimization is as follows:

$$L(x,\lambda,\mu) = f(x) + \sum_{i=1}^{m} \lambda_i h_i(x) + \sum_{j=1}^{p} \mu_j g_j(x). \quad (9)$$

3.2.3. Hybrid Genetic. The genetic algorithm tends to converge to the local optimal solution rather than the global optimal solution. It can be improved by using a larger population, but if a larger initial population is used, the calculation amount of the algorithm will be greatly increased; if a smaller number is used population, the algorithm may not be able to find the optimal solution. Although the genetic algorithm can quickly converge to the region near the optimal solution, it still requires a huge amount of calculation to achieve the final convergence in the nearby region. The sequential quadratic programming algorithm has the advantages of faster calculation speed and good boundary search. Therefore, a hybrid genetic algorithm that combines genetic algorithm and sequential quadratic programming algorithm is selected. The genetic algorithm is first used to run near the optimal solution, and then the SQP optimization algorithm is used to perform a more efficient and fast local search to find the global optimal solution.

3.3. Our Algorithm. To convert the teaching quality prediction problem of the flipped classroom of college ideological and political courses into regression modeling problems, this research proposes a hybrid genetic algorithm-optimized LSTM network method HGA-LSTM to find the best hidden layer unit for teaching quality prediction quantity and dropout probability. Generally, the larger the dataset is, the more the hidden layers and neurons can be used for modeling without overfitting. Since the experimental dataset is large enough, higher accuracy can be achieved. If there are too few neurons in each layer, the predictive model will be difficult to adapt during training. Using more hidden layer units can better update the weights, but it also means more calculations and longer training time, so the number of hidden layer units is not as good as possible but has to be related to the number of datasets adapted. Since LSTM is easier to learn long-term dependence, in order to prevent the gradient from disappearing and exploding, choosing an appropriate dropout probability can effectively avoid the problem of data overfitting. The principle of the algorithm in this paper is shown in Figure 5.

4. Experiments and Results

4.1. Experimental Environment. In the process of neural network training, the setting of hyperparameters will greatly affect the performance of the network model. This article uses *a da m* as the optimizer of our algorithm. We use GTX 1060 to accelerate the entire training process, the running platform is win 10, and the running memory is 8 GB. The entire model is built using the PyTorch framework. In the data processing stage, we extract 10% of the data from each category as the test set, 63% as the training set, and 27% as the validation set. The data in the test set is stored separately. During the training and verification process, the model cannot touch the test set data. Only in the testing phase, the test set data is read.

4.2. Dataset. This paper selects the historical data of the ideological and political courses in colleges and universities in a certain city and produces a training set and a test set.

4.3. Evaluation Index. To characterize the accuracy of the prediction model, error evaluation indicators are needed. Commonly used error evaluation indicators include average absolute error (MAE), root mean square error, and average absolute percentage error (MAPE). The calculation equation is as follows:

$$\begin{aligned} \mathrm{RMSE} &= \sqrt{\frac{\left(\sum_{n=1}^{N} \left|x_{(n)} - \hat{x}_{(n)}\right|^2\right)}{N-1}}, \\ \mathrm{MAE} &= \frac{\sum_{n=1}^{N} \left|x_{(n)} - \hat{x}_{(n)}\right|}{N}, \\ \mathrm{MAPE} &= \frac{\left(\left(\sum_{n=1}^{N} \left|x_{(n)} - \hat{x}_{(n)}\right|\right)/x_{(n)}\right) \times 100\%}{N}, \end{aligned} \quad (10)$$

where $x_{(n)}$ represents the actual value and $\hat{x}_{(n)}$ represents the predicted value.

4.4. Comparative Experiment. The performance comparisons of the four approaches are represented in Table 1. In the training phase, the RMSE errors of the method in this paper and GRU are 0.0194 and 0.0366, respectively. The RUL prediction based on HGA-LSTM is 0.0172 lower than the RMSE error of GRU, indicating that the prediction accuracy of HGA-LSTM in the training set is higher. In the test phase, the RMSE error and RZ based on HGA-LSTM are 0.0270 and 0.99_51, respectively. The RMSE error is 0.0193 less than using GRU and RZ is 0.0300 higher than using GRU, indicating that LSTM RNN has high prediction accuracy and good stability. These results show that the method based on HGA-LSTM has better performance than GRU on both the training set and the test set. SIM RNN cannot learn long-term dependence. The training set RMSE

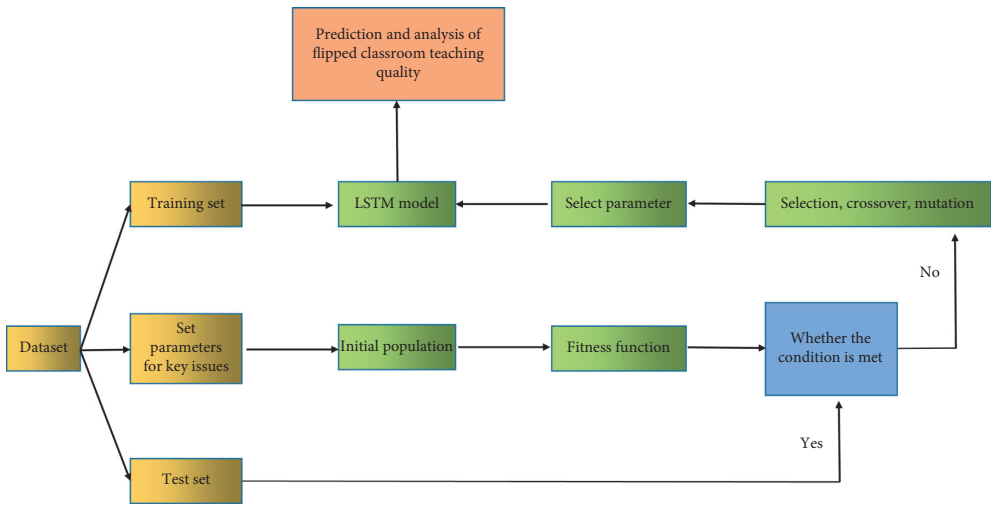


Figure 5: Prediction and analysis of flipped classroom teaching quality.

Table 1: Comparative experiment results.

Method	RMSE	MAPE	MAE
SIM RNN	0.074	0.619	0.057
GRU	0.388	0.350	0.033
LSTM	0.026	0.195	0.018
Ours	**0.021**	**0.192**	**0.016**

Bold values represent the smallest values, which means the method achieves better performance than other methods.

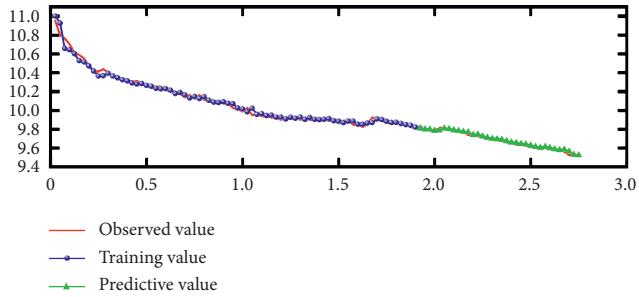


Figure 6: Experimental results of our method.

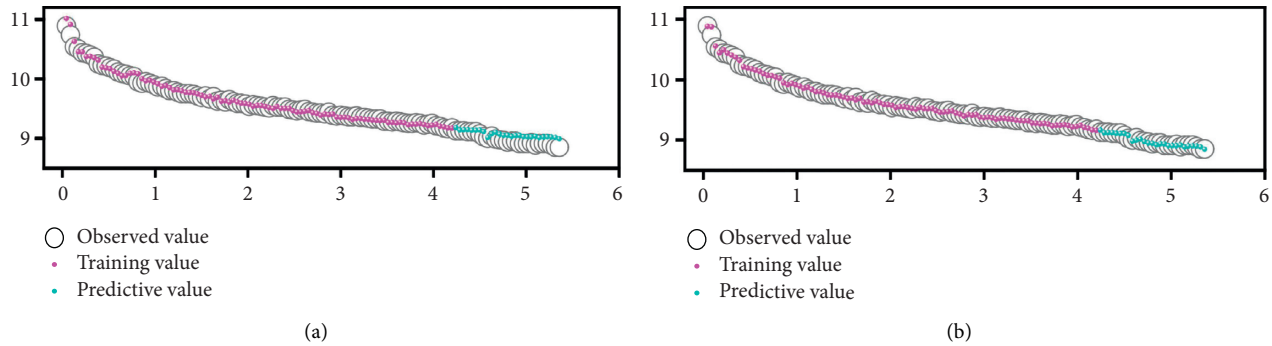


Figure 7: Comparative experiment results of dropout. (a) No dropout. (b) Dropout.

is 0.0261 higher than HGA-LSTM, and the test set RMSE is 0.1099 higher. From the perspective of various errors, the method in this paper provides the best performance for predicting the quality of flipped classroom teaching in college ideological and political courses (as shown in Figure 6).

Table 2: Predicted results of teaching quality.

Method	RMSE	MAPE	MAE
No dropout	0.05	0.44	0.04
Dropout	0.02	0.19	0.01

4.5. Comparison Prediction for Dropout. From Figure 7, it is shown that the prediction effect of LSTM RNN using dropout is significantly better. The specific error is shown in Table 2. The RMSE of the training set and the test set without any antioverfitting algorithm are 0.0402 and 0.0992, respectively, and the error of the test set is 2.47 times of the RMSE of the training set, indicating that there is an obvious overfitting phenomenon. At the same time, we found that adding dropout can increase the model accuracy.

5. Conclusion

The rising of information technology and advancements have encouraged the flipped classroom in an operative and effective way. It flips knowledge transfer and knowledge internalization from two levels of teaching structure and teaching process, reversing the traditional teaching knowledge transfer in class and knowledge deepening after class from time and space. In this paper, we proposed a novel recurrent neural network method with a channel attention mechanism for music feature classification. Because the music classification method based on the convolutional neural network ignores the timing characteristics of the audio itself. In this regard, this article combines the proposed convolution structure with the bidirectional recurrent neural network, proposes a music classification model based on the convolution recurrent neural network, and uses the attention mechanism to assign different attention to the output of the recurrent neural network at different times. The classification accuracy of the model on the GTZAN dataset has increased to 93.1%, the AUC on the multilabel labeling dataset MagnaTagATune has reached 92.3%, surpassing other comparison methods, and the labeling of different music labels has been analyzed. This method has good labeling ability for most of the labels of music genres and also has good performance on some labels of musical instruments, singing, and emotion categories. The effectiveness of the proposed approach is shown from the experimental results.

Data Availability

The data used to support the findings of this study are included within the supplementary information file(s).

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

[1] G. Akçayır and M. Akçayır, "The flipped classroom: a review of its advantages and challenges," *Computers & Education*, vol. 126, pp. 334–345, 2018.

[2] P. Strelan, A. Osborn, and E. Palmer, "The flipped classroom: a meta-analysis of effects on student performance across disciplines and education levels," *Educational Research Review*, vol. 30, Article ID 100314, 2020.

[3] E. Cabı, "The impact of the flipped classroom model on students' academic achievement," *International Review of Research in Open and Distance Learning*, vol. 19, no. 3, 2018.

[4] É. Ádámkó, "Online self-learning," *International Journal of Engineering and Management Sciences*, vol. 5, no. 1, pp. 542–553, 2020.

[5] K. Liu, "Design and application of an online english self-learning platform," *International Journal of Emerging Technologies in Learning*, vol. 12, no. 8, p. 4, 2017.

[6] X. Wu, "Theory and practice of multimedia courseware design for ideological and political theory courses in colleges and universities," in *Proceedings of the 2020 International Conference on Computers, Information Processing and Advanced Education*, pp. 347–350, Ottawa Canada, October 2020.

[7] Y. Rong-bo, "The historical commitment in the ideological and political theory courses in colleges and universities in the new era," *Journal of Hunan First Normal University*, 2019.

[8] Y. N. Zhang, "Research on the innovation of college students' ideological and political education in big data era," *DEStech Transactions on Computer Science and Engineering*, DEStech Publications, Lancaster, PA, USA, 2017.

[9] X. Li and T. Soobaroyen, "Accounting, ideological and political work and Chinese multinational operations: a neo-Gramscian perspective," *Critical Perspectives on Accounting*, vol. 74, Article ID 102160, 2021.

[10] F. Martin, A. Ritzhaupt, S. Kumar, and K. Budhrani, "Award-winning faculty online teaching practices: course design, assessment and evaluation, and facilitation," *The Internet and Higher Education*, vol. 42, pp. 34–43, 2019.

[11] C. S. González-González, A. Infante-Moro, and J. C. Infante-Moro, "Implementation of e-proctoring in online teaching: a study about motivational factors," *Sustainability*, vol. 12, no. 8, p. 3488, 2020.

[12] Y. Lu, ""Microcourse" based flipped classroom teaching mode for ideological and political education of colleges and universities," in *Proceedings of the 2020 6th International Conference on Social Science and Higher Education (ICSSHE 2020)*, pp. 693–697, Atlantis Press, Xiamen, China, January 2020.

[13] P. Zheng, X. Wang, and J. Li, "Exploration and practice of curriculum ideological and political construction reform -take "information security" course as an example," *ASP Transactions on Computers*, vol. 1, no. 1, pp. 1–5, 2021.

[14] X. Ning, K. Gong, W. Li, L. Zhang, X. Bai, and S. Tian, "Feature refinement and filter network for person re-identification," *IEEE Transactions on Circuits and Systems for Video Technology*, p. 1, 2020.

[15] Y. Ding, X. Zhao, Z. Zhang, W. Cai, and N. Yang, "Multiscale graph sample and aggregate network with context-aware learning for hyperspectral image classification," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 14, pp. 4561–4572, 2021.

[16] Y. Tong, L. Yu, S. Li, J. Liu, H. Qin, and W. Li, "Polynomial fitting algorithm based on neural network," *ASP Transactions on Pattern Recognition and Intelligent Systems*, vol. 1, no. 1, pp. 32–39, 2021.

[17] Z. Chu, M. Hu, and X. Chen, "Robotic grasp detection using a novel two-stage approach," *ASP Transactions on Internet of Things*, vol. 1, no. 1, pp. 19–29, 2021.

[18] Y. Zhang, W. Li, . Zhang, X. Ning, L. Sun, and Y. Lu, "AGCNN: adaptive gabor convolutional neural networks with receptive fields for vein biometric recognition," *Concurrency and Computation: Practice and Experience*, John Wiley & Sons, Hoboken, NJ, USA, 2020.

[19] J. Zhang, Y. Liu, H. Liu, and J. Wang, "Learning local-global multiple cfrvtkfr," *Sensors*, vol. 21, no. 4, p. 1129, 2021.

[20] W. Cai and Z. Wei, "PiiGAN: generative adversarial networks for pluralistic image inpainting," *IEEE Access*, vol. 8, pp. 48451–48463, 2020.

[21] R. Liu, X. Ning, W. Cai, and G. Li, "Multiscale dense cross-attention mechanism with covariance pooling for hyperspectral image scene classification," *Mobile Information Systems*, vol. 2021, Article ID 9962057, 15 pages, 2021.

[22] L. Liang, Q. Yin, and C. Shi, "Exploring proper names online and its application in english teaching in university," *ASP Transactions on Computers*, vol. 1, no. 1, pp. 24–29, 2021.

[23] W. Cai, B. Liu, Z. Wei, M. Li, and J. Kan, "TARDB-Net: triple-attention guided residual dense and BiLSTM networks for hyperspectral image classification," *Multimedia Tools and Applications*, vol. 80, no. 7, pp. 11291–11312, 2021.

[24] W. Sun, P. Zhang, Z. Wang, and D. Li, "Prediction of cardiovascular diseases based on machine learning," *ASP Transactions on Internet of Things*, vol. 1, no. 1, pp. 30–35, 2021.

Hindawi
Scientific Programming
Volume 2021, Article ID 5795884, 10 pages
https://doi.org/10.1155/2021/5795884

Research Article

Student Physical Health Information Management Model under Big Data Environment

Hui Wang,[1,2] Ning Wang,[2] MeiJie Li,[3,4] Simeng Mi,[4] and YaYa Shi[4]

[1]*Sports Institute, Henan Agricultural University, ZhengZhou 450002, China*
[2]*Faculty of Physical Culture, Gdańsk University of Physical Education and Sport, 80-336 Gdansk, Poland*
[3]*Humour Division, Huanghe University of Science and Technology, Zhengzhou 450063, China*
[4]*Sports Institute, Korea Gangneung-Wonju National University, Gangneung 25457, Republic of Korea*

Correspondence should be addressed to MeiJie Li; 201508113@hhstu.edu.cn

Received 20 April 2021; Accepted 7 June 2021; Published 10 July 2021

Academic Editor: Shah Nazir

Health is considered an important foundation for students' success. However, with the accelerated pace of life, rising pressure from various parties, weak health awareness, lack of exercise time, and other reasons, students' physical quality is generally declining, the incidence of health diseases is increasing, and the onset age tends to be younger. With the development of the concept of "health first," health management continues to expand and extend and students' health management has attracted more attention from many aspects. Due to the late and low starting point of health management research and the lack of professional theoretical support, a complete, mature, and effective health management service system has not been established to deal with the students' health. In order to make student health management more scientific, normative, and effective, this article has proposed big data technology to build the student health information management model. The first step of the approach is to store and analyze the data of students' physical health. It is necessary to combine the data collection, supervision, data analysis, and data application of students' physical health and gradually improve the national monitoring and evaluation system of students' physical health. Student health check-up management platform is mainly used in realizing the school student information management and student health information relationship between system, science, standardization, and automation, and its main task is to use a computer to perform daily management of all previous medical information of students, such as query, modify, add, delete, and enhance the physical health of students information management ability given the large data analysis of useful information. In addition, we have built a doctor recommendation model based on online questions and answers to give specific health recommendations for students of different physiques.

1. Introduction

In recent years, the physical fitness [1–3] of domestic students has generally been poor. Although the concept of "green health" continues to take root in the hearts of the people, the rate of adolescent obesity, myopia rate, depression rate, and prevalence rates are still increasing year by year. At the same time, the physical health literacy of young people, especially college students, has declined overall, their social adaptability has continued to decline, and their psychological endurance has become increasingly weak. After entering society, it is difficult to cope with work intensity and interpersonal relationships, psychological pressure cannot be effectively resolved, and depression and suicide incidents of college students are endless. The physical health of college students is not only an issue of concern to their parents but also an issue that has received attention from all walks of life [4, 5].

However, through the investigation of college students' life and study, we found that the reasons for the decline of college students' physical quality [6] year by year are mainly in the following two aspects: one is that the students' self-restraint is weak so that they have no time to eat and sleep and for indulgence. For example, irregular work and rest, often staying up late, unhealthy diet and eating junk food, and the arrangement between learning and sports being not

appropriate all constantly push the physical health level of college students to the second (sub-)health state, and even "senile disease" body is not rare [7]. Nowadays, the vast majority of people rely on their youth, often ignore their health, and wantonly squander their physical fitness. Secondly, in terms of teaching in colleges and universities, most schools no longer offer PE classes after the junior year, leaving only the routine morning run. The time for physical exercise is constantly compressed, and even the standards of physical examination are changed again and again. Therefore, the weak sports consciousness, lack of exercise time, and other problems have led to the fall of our country's education field into a vicious circle of focusing on intellectual education and high scores and ignoring physical education, low ability, and weak physic, which to a certain extent explains the urgency, necessity, and difficulty of establishing health management service system in college physical education teaching [8, 9].

The significance of this study is to build an effective health management service system for colleges and universities with the help of the era background of information big data and provide a theoretical reference for promoting the development of student physical health management services. At present, the assessment index system of college physical education is mainly composed of "scores" and "attitudes," and the evaluation methods have not undergone substantial changes. Lack of "knowledge" and "physical fitness" and other important indicators that can show physical health, body shape, physical function, and mental state cannot fully reflect the guiding concept of "health" and cannot fully reflect the physical and health status of students, so there are obvious shortcomings. Students only complete the physical exercise required by the school and related tasks assigned by the physical education teacher in order to obtain high scores and good grades. This inherent thinking leads to insufficient health awareness of students, low subjective and conscious initiative, and low efficiency of physical education courses. The young people in our country are gradually getting younger disease. In the era of highly developed information technology, modern management technology [10] has not been fully developed and used. The current university student health management service system lacks good management and monitoring, and problems such as monotonous management methods, insufficient theoretical innovation, and incomplete monitoring mechanisms are prominent. Therefore, it is particularly important to further improve the index system of health management and establish a comprehensive, scientific, and practical health management service system [11].

The main contributions of this article are as follows:

(1) It builds a student health information management model based on big data technology. The first step is to store and analyze the data of the students' physical health and then combine the data collection, supervision, data analysis, and data application of the students' physical health to improve the student's physical health monitoring and evaluation system.

(2) It proposes to build a student health check management platform, which is mainly used for school student information management, system, science, standardization, and automation of student health information relationship, and can use computers in daily management to manage all previous medical information of students.

(3) This article conducts experiments, and the experimental results proved the effectiveness of constructing a student health information management model based on big data technology. The methodology of the article is briefly given in section 3.

The article is structured as follows: Section 2 discusses the background of the study with the existing related work. Experiments and results of the proposed study are discussed briefly in section 4. The article is concluded in section 5.

2. Background

The exploration of health management [12] has always existed. From the beginning of the use of fire to ancient medicine and then to modern medicine, people have constantly been fighting against various diseases in order to better maintain the health of life. With the development of history, human knowledge continues to accumulate, and modern health management has been gradually developed and improved under the impetus of the market economy. In the 1950s, the US economy developed rapidly and people's living standards improved rapidly. People's attention and importance to health were gradually put on the agenda. Under such a social background, health management as a new term came into being in the United States, and in just a few years, it has rapidly developed into an independent discipline and a special industry.

The importance of health management has been widely recognized in American society, so it is extremely urgent to construct a reasonable and effective health management mechanism. The American government divides the student health management system into three parts: physical health, health education, and health promotion. The first two levels are easily disturbed by various factors, the management cannot be carried out in a comprehensive way, and the results are relatively poor. Under such circumstances, the term "health promotion system" has slowly emerged [13]. In the 1980s, the concept and basic connotation of the health promotion system were gradually clarified. Some comprehensive schools further proposed the transformation of the school health system into a health promotion system in the health promotion plan for college students and gradually refined the school health system from the initial three parts to eight parts.

In the context of the prevailing concept of European health management [14], Finland also launched a large-scale exploration of health management models in the 1960s and 1970s based on its own national conditions. The Finnish government tried a new health management model to make full use of community health services. A pilot study was conducted in 1972, based on mutual cooperation among

communities, to guide people to change their original bad behaviors and basic lifestyles and choose healthy lifestyles. After 25 years of practice, the smoking rate, cholesterol, lung cancer, coronary heart disease, and other cardiovascular disease mortality rates in men aged 25-26 have decreased by 50%, 20%, 71%, 73%, and 71%, respectively. It can be seen that the pilot study of the health management model has a significant effect, and it can be promoted and applied. Finland's health management model based on community services has had a significant impact not only in the country but also in countries around the world, providing them with a wealth of valuable experience.

With the rapid development of Internet technology, various management tasks of student health examinations have gradually entered the era of information management. Practice has also proved that the intervention of computers and information-based management can indeed make the school's various tasks more scientific, more systematic, more complete, and more humane, with high efficiency, good quality, and easy preservation and communication. Now, schools have basically entered the age of computer office, and computers have quickly entered ordinary households. Therefore, the management of student health examination will change from paper management to computer network management. The student health examination management platform is mainly used for school student information management to realize the systematization, scientification, standardization, and automation of student health information relations. Its main task is to use computers for daily management of students' previous physical examination information, such as inquiries, modifications, and additions, deleting and analyzing useful information for big data, and designing a student health information management system for these requirements. The implementation of the application of the school's student health management system is an important measure to further promote the standardization and electronic management of student information and to improve the level of education and student health [15].

3. Methodology

3.1. Analysis of the Functional Requirements of the System. This article uses the BS structure to achieve the related requirements to ensure the security, reliability, and scalability of the management system; the relative independence of each module makes the system convenient according to the need for adding new functional modules. Based on the demand of care on the health information management system, the system is divided into eight function modules (as shown in Table 1), respectively, for the students' basic information collection, physical examination, physical examination data storage, information search, information maintenance, statistical data analysis, comprehensive health assessment, data export, system settings module, and log management module.

The health management system for primary and middle school students is designed with four user roles: students, teachers, medical staff, and website administrators. Student users have the authority to register and log in, query student status information, query physical examination information, and view physical examination notification operations. The specific user functions are shown in Figure 1.

After the teacher role logs into the system, it has the authority to view the medical examination notice, maintain student status information, view student medical examination data, and view statistical report information. After physical examination, medical staff logs in to the system, which contains the functions of viewing physical examination notifications, maintaining student physical examination data and viewing statistical report information. After the website administrator logs into the system, he has the right to add users, assign permissions, issue medical notices, and maintain website data.

3.2. Systematic Business Process Analysis. The health management system [16] of primary and middle school students in this article is a set of complete website applications, which has experienced a complete project cycle in the development process. Software development model uses rapid prototyping models and allows the requirements analysis phase of the demand for software to carry out the preliminary analysis and definition of not completely rapid design developed prototype software system (showing all or part of the function and performance of software to be developed, user testing the prototype evaluation, specific improvement opinions, and plenty of detailed software requirements for developers to modify). The development program of design patterns uses MVC design pattern; MVC is the abbreviation of model-view-controller, a software design model, using a method of separation of business logic, data, and interface display organization code, gathered in a business logic component, which tries to improve and personalize custom interface and user interaction at the same time and does not need to write the business logic. MVC enables the same program to use different representations to achieve code separation. The hierarchical pattern is shown in Figure 2.

The advantage of the layered architecture is that it achieves a certain degree of separation of concerns, which is conducive to the reuse of the logic of each layer; it normalizes the calling relationship between layers and can reduce the dependency between layers; if the interface design between layers is reasonable, then it is not difficult to replace the original level of realization with a new one.

We also use tokens to prevent duplicate submissions of data. There are two situations in which information is repeatedly submitted while information is being entered. (1) After the user completes the operation, click Back to return to the page and submit again. (2) Refresh the current page after the user completes the operation. How it works: before processing the client's request, the server side compares the value of the token contained in the request to the value saved in the current session to see if it matches. After the request is processed, and before the information reaches the client, a new token is generated. The token value replaces the token value in the current session and is passed to the client. This way, if the user goes back to the submit page and commits

Table 1: System function module division.

Module	Introduction
User registration	The new user completes the registration operation
User login	The user completes the login operation on the login page
User password modification	Users can modify their own password after logging in
User information view	Basic user information view
Student basic information maintenance	Complete entry, modification, and deletion of basic student information
Student basic information query	Complete the conditional query operation for student information
Physical examination information maintenance	Complete the entry, modification, and deletion of the student's physical examination data
Physical examination information query	Complete the conditional query operation on the student's physical examination data
Big data analysis of physical examination information	Complete the intelligent scoring operation of the medical examination data
Maintenance of medical notice announcement information	Complete the input, modification, and deletion operations of the student's physical examination notification information
Announcement information query of medical examination notice	Complete the conditional query operation on the notification information of the student's physical examination
Health check notification information reminder	Complete the rest of the operation for the notification information of the student's physical examination
User authority information management	Complete the assignment and management of user authority information
User operation log record	Record the user's operation information in detail, and generate a log
User login record	Record the user's login information in detail and generate a log

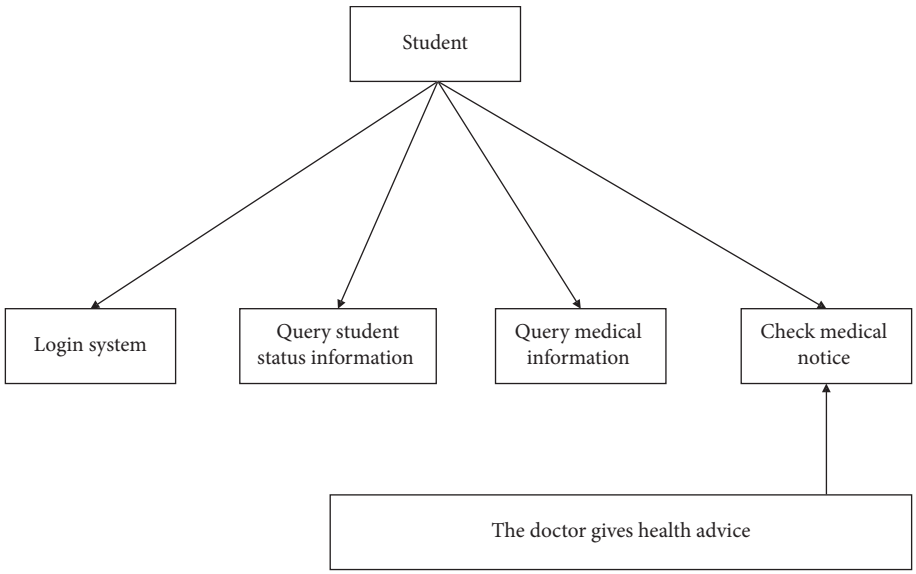


Figure 1: Student behavior use case diagram.

again, the token passed by the client does not match the value of the token in the service, thus effectively preventing the commit.

3.3. Feasibility Study of the System. Systems Development Life Cycle (SDLC) is the life cycle of software until it is scrapped or discontinued. There are problem definition, feasibility analysis, overall description, system design, coding, debugging and testing, acceptance and operation, and maintenance and upgrading to obsolescence in the cycle. In other stages, this time-based method of thinking is a principle of thinking in software engineering, that is, step by step. Each stage must have definition, work, review, and document formation for communication or reference to improve the quality of the software. The feasibility analysis stage is divided into technical feasibility, operational feasibility, and economic feasibility. This project focuses on user needs and feasibility analysis to illustrate the necessity of the design and production of this project.

In the preliminary investigation of the student health management system, it is found that people are more likely to accept the BS structure website system rather than the CS structure software system. It is becoming increasingly mature based on the current Internet technology and deeply penetrated into People's Daily life, so people are happy to accept the website system. This project is based on the decision to adopt the BS architecture of the website system. To build a

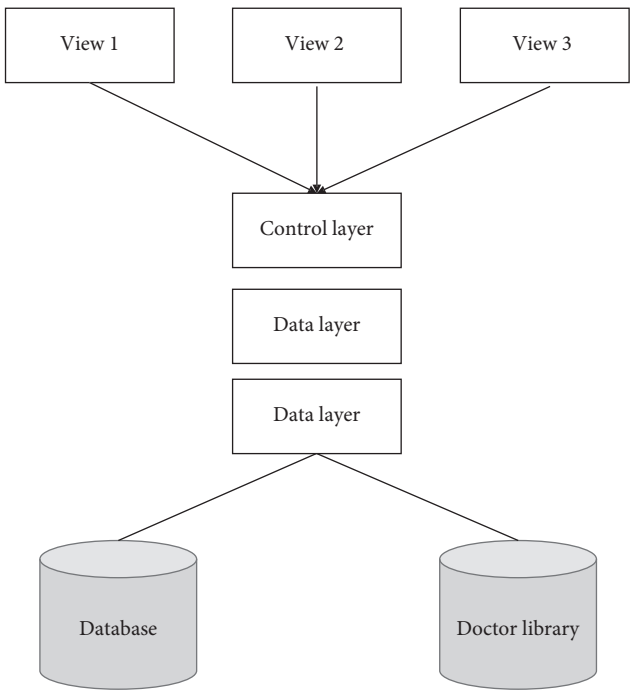


Figure 2: System hierarchical structure.

website, we should consider all aspects of hardware, software, and technology. Set in the technical feasibility, development process operating system uses the familiar Windows 7 ultimate operating systems; web applications run software adopts tomcat7 open-source software; web load balancing uses Nginx 1.8 open-source application software; development program uses the Java programming language; web project framework uses *J* final speed development framework, using SQL database using the MYSQL open-source software structured development language. The front-end display page uses HTML5 technology with JavaScript script verification language and CSS style language to write code and the front-end dynamic display page uses JSP technology to display. The development software uses MyEclipse10 to build the website framework and write the code. The database uses SQLYOG Enterprise to create the database and write and execute the SQL language.

3.4. Doctor Recommendation Model. Doctor recommendation framework consists of four stages: data collection, data preprocessing, model training, and doctor recommendation. The recommendation model is called DRGAN (GAN-based Doctor Recommendation).

3.4.1. Data Preprocessing. The data source of this article is question and answer data between students and doctors. Data preprocessing mainly prepares feature vectors as input for the model training phase of the framework. Data preprocessing includes two steps: word segmentation and feature extraction.

For numerical features, this section directly splices all numerical features of students (or doctors) together. For personal attributes, this article uses One-Hot coding. For example, the gender of the questioner is represented by a vector with three dimensions ["male", "female", "unknown"] (that is, female can be represented as [0, 1, 0]). Because there are not many categories in the dataset in this article, such as age, region, and doctor's title, the impact of the sparseness of One-Hot coding is not significant. However, if the text content data also use On-Hot coding (i.e., the bag-of-words method), a very large and sparse feature vector will be generated, which will significantly affect the performance of the deep learning model. In addition, because the content of the question and answer text is usually shorter and fewer sentences as shown in the example, unlike the data in the disease text topic classification chapter, which is mainly long text, the use of the TFIDF method in this section will also produce very sparse feature vectors. Although the original word embedding representation method (such as Word2-vector) requires a lot of training data as described in the introduction, the idea of representing words can be used for reference. In this section, each word after the text content segmentation is expressed as k dimensions (such as 100 dimensions) random initialization vector and then through the GAN-based model [17, 18] to vector the content of the question and answer text.

3.4.2. Model Training. This article proposes a model based on the GAN method to train doctors recommendations. For a given query q_i, there is the best response (doctor's answer) r_{true} among all the response results R. Compared to making an absolute correlation judgment on a certain response, the model is easier to judge the relative correlation between a pair of responses through comparison. Therefore, this section uses a pair-by-pair setting to sort the responses corresponding to each query.

For each query q_i, this article sets a pair of responses $P = \{ < r_{\text{true}}, r_i > | r_{\text{true}} > r_i \}$, where $r_{\text{true}} > r_i$ indicates that r_i is not as relevant to q_i as r_{true}. In this section, let $g_\theta(r_i|q, r_{\text{true}})$ and $d_\varphi(P|q)$ denote the generative model and discriminant model of GAN, respectively.

Generator: for a given query q and real response r_{true}, it tries to select the response most relevant to the real response from the candidate response list to generate a response pair P. This article uses the soft max function to describe the process of generating the model as follows:

$$\begin{aligned} p(r_i|q, r_{\text{true}}) &= g_\theta(r_i|q, r_{\text{true}}) \\ &= \exp\left(-\left|f_\theta(r_{\text{true}}, q) - f_\theta(r, q)\right|\right). \end{aligned} \tag{1}$$

Discriminator: try to distinguish the correlation between the response generated by its corresponding generative model and the real response to P. This article uses the sigmoid function to estimate the probability that the response to P is correctly distinguished by the model is expressed as follows:

$$\begin{aligned} &p(< r_{true}, r_i > | q) \\ &= d_\varphi(P|q) = \frac{1}{1 + \exp\left(-\left(f_\phi(r_{\text{true}}, q) - f_\phi(r_i, q)\right)\right)}, \end{aligned} \tag{2}$$

where $f_\theta(r,q)$ and $f_\varphi(r,q)$ are scoring functions that measure the relevance of query and response, which will be described in the following part. In the final form, this article uses the standard cross-function to unify the two functions to obtain the objective function as follows:

$$O^{G,D} = \min_\theta \max_\varphi \sum_{i=1}^{n} \left\{ \begin{array}{c} r_{true}\left[\log\left(d_\phi(P|q)\right)\right] \\ +r_{g_\theta}(r_i|q, r_{\text{true}})\left[\log\left(1 - d_\varphi(P|q)\right)\right] \end{array} \right\}, \tag{3}$$

where n is the number of queries. In order to optimize the discriminant model that is actually a binary classifier, this article can use the stochastic gradient descent method to update and obtain the optimal parameter φ in the case of a fixed generative model.

In the part of data preprocessing, this chapter mentions that the dataset contains numerical features, personal attribute features, and text content features. Because the dimensions of the first two features are less, in this section, numerical features and personal attribute features are directly combined and collectively referred to as (profile) features. However, in the feature extraction of the text content, a large number of words with k-dimensional vectors are obtained, and the dimensionality of the vectors obtained by directly concatenating the word vectors is too high. Therefore, it is necessary to apply more effective methods to represent the text content.

The relevance score of a doctor's response to a health question consultation is defined by the use of cosine similarity as follows:

$$S_{qa} = \text{cosine}\left(v_q, v_a\right) = \frac{v_q^T v_a}{\left|v_q\right|\left|v_a\right|}. \tag{4}$$

Scoring function A: the representation process of scoring function 1 is shown in Figure 3. First, the profile feature of the query and the profile feature of the corresponding response are combined into a joint feature vector v and passed into the input layer of a two-layer neural network [19–22]. The process of this neural network can be expressed as follows:

$$\begin{aligned} S_{\text{profile}} &= w_2^T \tan\ h\left(W_1 v_j + b_1\right) + b_2, \\ \tan\ h(x) &= \frac{e^x - e^{-x}}{e^x + e^{-x}}, \end{aligned} \tag{5}$$

where W_1 is the fully connected weight matrix of the input layer and the hidden layer, b_1 is the bias value of the hidden layer, w and b are the weight and bias value of the second hidden layer (i.e., output layer) node, and $\tan\ h(x)$ is one type of activation function. The text content of the query and response can be expressed as Stext according to the QA-CNN method on the right side of Figure 3. Finally, this article combines S_{profile} and S_{text} into a two-dimensional vector and again implements a two-layer neural network to output score S_{nn} that measures the correlation between the query and the response.

Scoring function 2: in scoring function 1, this article firstly models the same type of features and finally outputs the result through a neural network [23–26]. The scoring function 2 (as shown in Figure 4) is another idea; that is, first use a neural network to model the same input source (query source vs. response source) and then use cosine similarity to output the result. Scoring function 2 no longer uses the QA-CNN model for the last step of text content modeling for health consultation (cosine similarity calculation) and only uses the hidden layer representation vector of the text content obtained, such as v_{qt} and v_{rt} in the image. The query and response profile vectors v_{qp}, v_{rp} and their corresponding v_{qp}, v_{rp} are combined to form a joint feature vector, which will be passed into the input layer of another single hidden layer neural network. Finally, this article obtains the joint feature vector v_q and v_r of the hidden layer from the neural network and calculates the cosine similarity $S_{\cos}$ of these two vectors as the correlation score between the query and the response. The process can be expressed mathematically as follows:

$$\begin{aligned} v_q &= \tan\ h\left(W_1'\left[v_{qp}, v_{qt}\right] + b_1'\right), \\ v_r &= \tan\ h\left(W_2'\left[v_{rp}, v_{rt}\right] + b_2'\right) \\ S_{\cos} &= \text{cosine}\left(v_q, v_r\right) = \frac{v_q^T v_r}{\left|v_q\right|\left|v_r\right|}. \end{aligned} \tag{6}$$

3.4.3. Doctor Recommendation. After the generative model and the discriminant model were trained against each other, all the parameters optimized in the training process were used to evaluate the model performance in the experiment. For each questioner (e.g., students), the trained model based on GAN (choose to generate models or discriminant model corresponds to the parameter set depending on which model has better assess performance) output questioner query and response focusing all response points, scored again through the response to sort, finally recommended response of the first N a doctor.

4. Experiments and Results

4.1. Experimental Environment. The experiments with all the algorithms were performed on a computer equipped with a single NVIDIA GTX1080TI GPU (16 GB). We have implemented the model construction through the PyTorch deep learning library, the programming language we use is Python, and we batch processed 100 samples each time.

4.2. Dataset. The datasets in this article are from the websites Ask120.com and Xywy.com. After data cleaning, a limited amount of data is obtained, and the ratio of 4 : 1 is used to randomly divide the data according to the number of queries. Since each query in the real dataset has one optimal response and several fewer other responses, for the convenience of testing, this article constructs a candidate response pool containing 10 responses by randomly selecting some responses corresponding to the original query from all response sets.

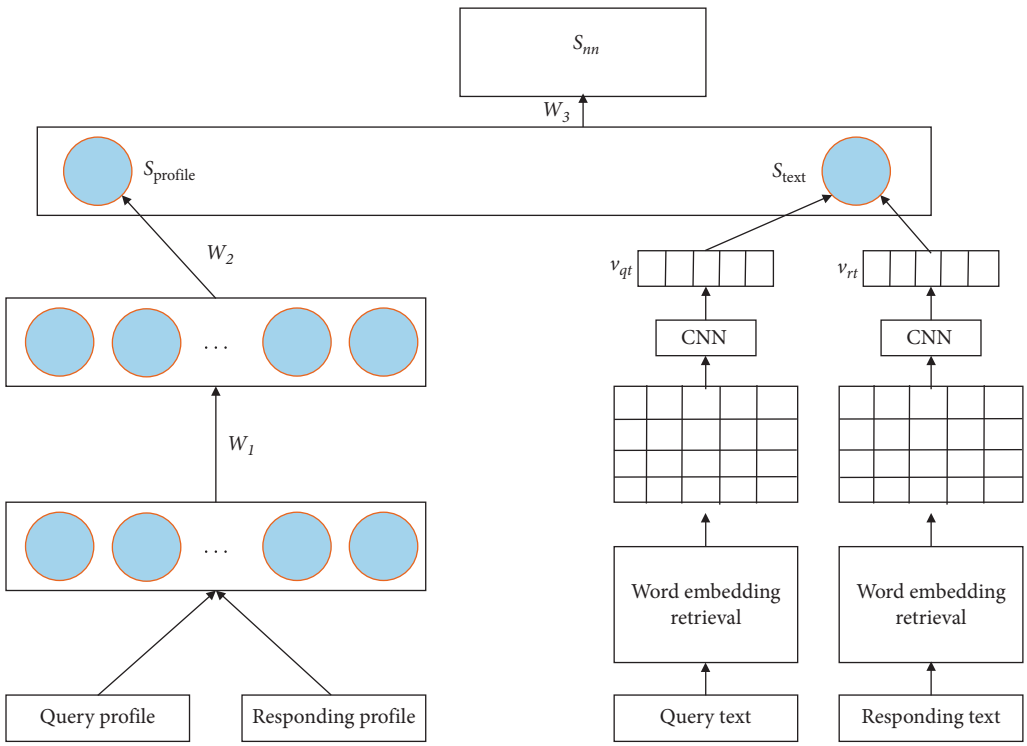


Figure 3: Structure diagram of scoring function A corresponding to query and response.

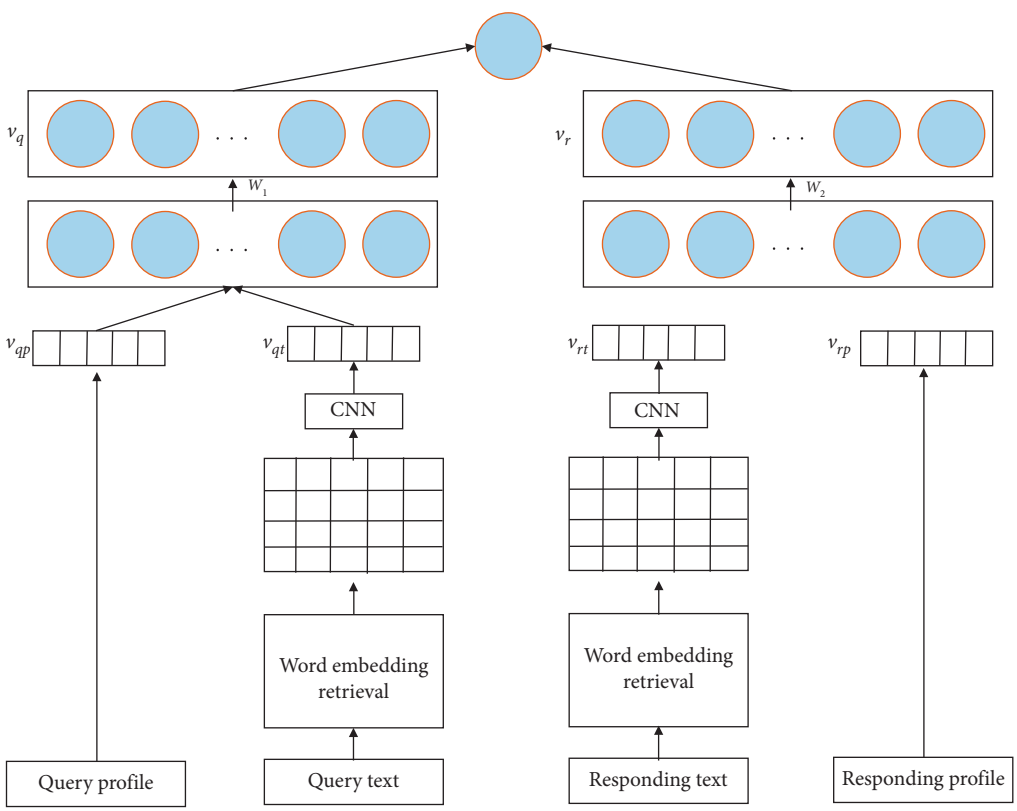


Figure 4: Structure diagram of scoring function 2 corresponding to query and response.

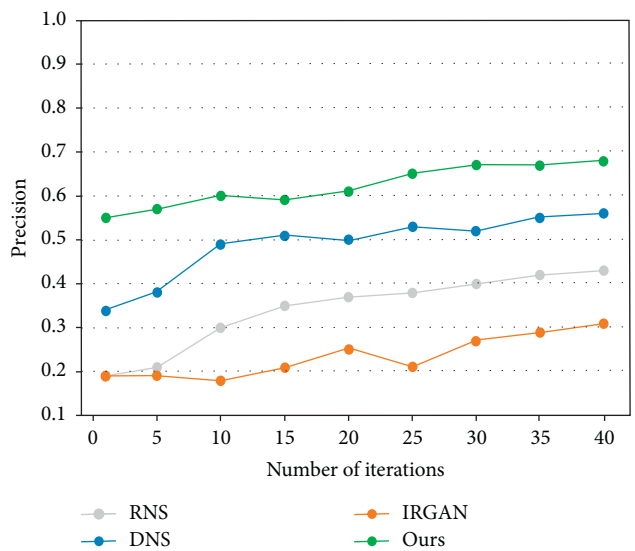


Figure 5: Results of comparative experiments using precision evaluation.

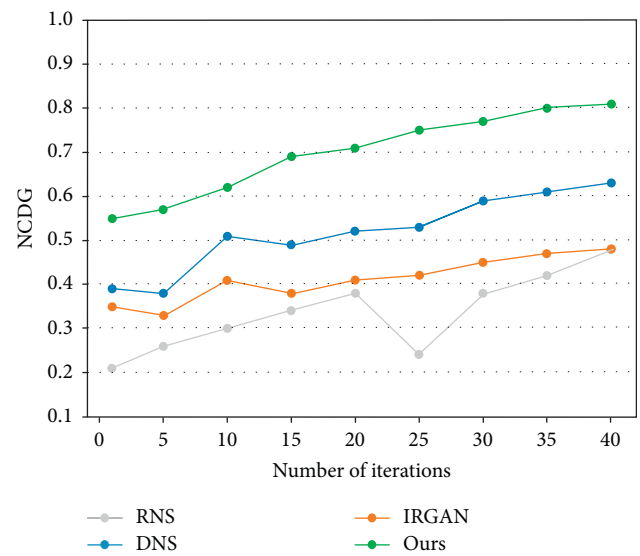


Figure 6: Results of comparative experiments using NCDG evaluation.

4.3. Evaluation Index. In the experiment of this article, the indicators used to evaluate the recommended performance are precision, normalized deviation cumulative gain, and average accuracy (MAP). The definitions of these three evaluation indicators are as follows:

$$\text{Precision} = \frac{|\text{true items}|}{N}, \tag{7}$$

where N is the first N number of recommended responses, and the numerator represents the number of true responses in the recommended response.

$$\begin{cases} \text{rel}(k) = \begin{cases} 1 & \text{if response } k \text{ is true with query} \\ 0 & \text{else} \end{cases} \\ \text{DCG} = \sum_{k=1}^{N} \dfrac{\text{rel}(k)}{\log_2(k+1)} \\ \text{NDCG} = \dfrac{\text{DCG}}{\text{DCG}_\pi} \end{cases}, \tag{8}$$

where $\text{rel}(k)$ represents whether the kth response in the recommended ranking response list is related to the query (the best response) and π represents the response list sorted according to actual relevance, that is, the best-case ranking response list.

$$\begin{cases} \text{AveP} = \dfrac{\sum_{k=1}^{C} \text{precision} \times \text{rel}(k)}{|\text{true items}|} \\ \text{MAP} = \dfrac{\sum_{q=1}^{O} \text{AveP}(q)}{|Q|} \end{cases}, \tag{9}$$

where c is the total number of recommended responses and Q is the recommended response list set corresponding to multiple queries.

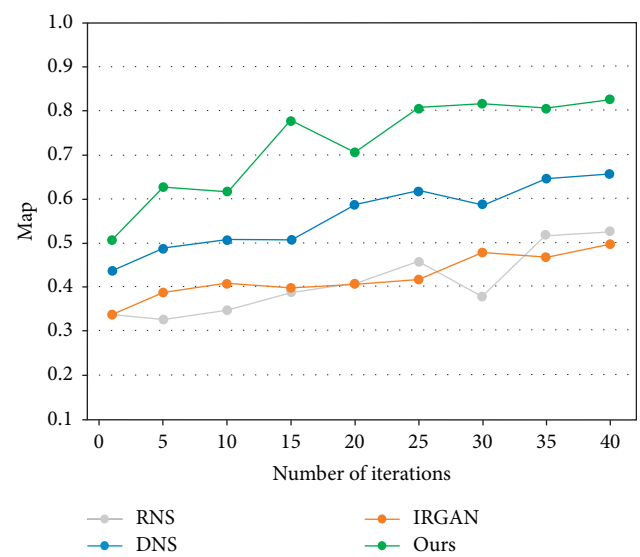


Figure 7: Results of comparative experiments using MAP evaluation.

4.4. Comparative Experiment. Figures 5–7 show the performance changes of all models on the precision indicator on the dataset. The performance of the proposed model is significantly better than benchmark comparison methods, including the DRGAN-nn model. As for the unstable performance curve, it is caused by more random sampling processes in the generation model during the training process of GAN. In fact, when training the generative model, it is necessary to perform the Softmax operation on the full set of responses, which is a very time-consuming operation. In the experiment, a compromise solution was adopted; that is, a smaller candidate response set was sampled from the complete set of responses for each query.

This article also summarizes the best performance results of the proposed model and benchmark method in the dataset

Table 2: Comparative experiment results.

Method	Precision	NCDG	MAP
RNS	0.43	0.48	0.53
DNS	0.56	0.63	0.66
IRGAN	0.31	0.48	0.50
DRGAN	0.51	0.59	0.62
Ours	**0.68**	**0.81**	**0.83**

1 of the performance indicators in the training process, as shown in Table 2. It can be seen that the method in this article is superior to the benchmark method in all evaluation indicators. It shows that the proposed scoring function (especially the scoring function (2)) is effective.

5. Conclusion

Health is considered to be an important foundation for students' success. With the fast-tracked pace of life, weak health awareness, increasing pressure from various parties, lack of exercise time, and other reasons, students' physical quality is usually declining, the incidence of health diseases is increasing, and the onset age tends to be younger. In this article, we build a student health information management model based on big data technology. The first step is to store and analyze the data on the students' health. It is necessary to integrate the data collection, supervision, data analysis, and data application of students' health to gradually improve the national student health monitoring and evaluation system. The student health check management platform is mainly used for school student information management and system, science, standardization, and automation of student health information relations. Its main task is to use computers in daily management to manage all previous medical information of students, such as query and modify, add, delete, and useful information for big data analysis to enhance students' health information management capabilities. In addition, we constructed a doctor recommendation model based on online question and answer and provided specific health advice for students of different physiques.

Data Availability

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

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

[1] C. Cadenas-Sanchez, T. Intemann, I. Labayen et al., "Physical fitness reference standards for preschool children: the PREFIT project," *Journal of Science and Medicine in Sport*, vol. 22, no. 4, pp. 430–437, 2019.

[2] K. Collins and K. Staples, "The role of physical activity in improving physical fitness in children with intellectual and developmental disabilities," *Research in Developmental Disabilities*, vol. 69, pp. 49–60, 2017.

[3] P. B. Júdice, A. M. Silva, J. Berria, E. L. Petroski, U. Ekelund, and L. B. Sardinha, "Sedentary patterns, physical activity and health-related physical fitness in youth: a cross-sectional study," *The International Journal of Behavioral Nutrition and Physical Activity*, vol. 14, no. 1, pp. 25–10, 2017.

[4] T. M. Barnett, A. McFarland, J. W. Miller, V. Lowe, and S. S. Hatcher, "Physical and mental health experiences among African American college students," *Social Work in Public Health*, vol. 34, no. 2, pp. 145–157, 2019.

[5] R. B. Lopez and B. T. Denny, "Negative affect mediates the relationship between use of emotion regulation strategies and general health in college-aged students," *Personality and Individual Differences*, vol. 151, Article ID 109529, 2019.

[6] X. Zhai, M. Ye, C. Wang et al., "Associations among physical activity and smartphone use with perceived stress and sleep quality of Chinese college students," *Mental Health and Physical Activity*, vol. 18, Article ID 100323, 2020.

[7] S. Lv, Y. Ding, H. Zhao, S. Liu, J. Zhang, and J. Wang, "Therapeutic potential and effective components of the Chinese herb Gardeniae fructus in the treatment of senile disease," *Aging and Disease*, vol. 9, no. 6, p. 1153, 2018.

[8] B. Behzadnia, P. J. C. Adachi, E. L. Deci, and H. Mohammadzadeh, "Associations between students' perceptions of physical education teachers' interpersonal styles and students' wellness, knowledge, performance, and intentions to persist at physical activity: a self-determination theory approach," *Psychology of Sport and Exercise*, vol. 39, pp. 10–19, 2018.

[9] C. Da-Wei, L. Chao, W. Shun, W. Xun-Ling, and W. Wen-Fang, "Research and application of multimedia digital platform in the teaching of college physical education course," *Journal of Intelligent & Fuzzy Systems*, vol. 34, no. 2, pp. 893–901, 2018.

[10] A. A. Ksenofontov, A. S. Ksenofontov, M. A. Kirpicheva, and P. V. Trifonov, "The use of modern management technology to improve business efficiency," in *IOP Conference Series: Materials Science and Engineering*vol. 483, no. 1, IOP Publishing, Article ID 012114, 2019.

[11] S. Khan and T. Yairi, "A review on the application of deep learning in system health management," *Mechanical Systems and Signal Processing*, vol. 107, pp. 241–265, 2018.

[12] D. Wang, K. L. Tsui, and Q. Miao, "Prognostics and health management: a review of vibration based bearing and gear health indicators," *IEEE Access*, vol. 6, pp. 665–676, 2017.

[13] S.-S. Lin and J.-J. Lin, "Development of a novel health promotion system based on wireless sensor network and cloud computing," *Sensors and Materials*, vol. 31, no. 3, pp. 939–952, 2019.

[14] M. del Rocío Moreno-Enguix, J. C. Gómez-Gallego, and M. Gómez Gallego, "Analysis and determination the efficiency of the European health systems," *The International Journal of Health Planning and Management*, vol. 33, no. 1, pp. 136–154, 2018.

[15] C. Thorley, *Not by Degrees: Not by Degrees: Improving Student Mental Health in the UK's Universities*, IPPR, London, UK, 2017.

[16] C. M. Sharkey, D. M. Bakula, K. L. Gamwell, A. J. Mullins, J. M. Chaney, and L. L. Mullins, "The role of grit in college student health care management skills and health-related quality of life," *Journal of Pediatric Psychology*, vol. 42, no. 9, pp. 952–961, 2017.

[17] W. Cai and Z. Wei, "PiiGAN: generative adversarial networks for pluralistic image inpainting," *IEEE Access*, vol. 8, pp. 48451–48463, 2020.

[18] C. Wang, X. Bai, X. Wang et al., "Self-supervised multiscale Adversarial regression network for stereo disparity estimation," *IEEE Transactions on Cybernetics*, 2020.

[19] X. Ning, K. Gong, W. Li, L. Zhang, X. Bai, and S. Tian, "Feature refinement and filter network for person Re-identification," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 99, 2020.

[20] W. Cai, Z. Wei, R. Liu, Y. Zhuang, Y. Wang, and X. Ning, "Remote sensing image recognition based on multi-attention residual fusion networks," *ASP Transactions on Pattern Recognition and Intelligent Systems*, vol. 1, no. 1, pp. 1–8, 2021.

[21] C. Yan, G. Pang, X. Bai et al., "Beyond triplet loss: person Re-identification with fine-grained difference-aware pairwise loss," *IEEE Transactions on Multimedia*, 2021.

[22] W. Cai, B. Liu, Z. Wei, M. Li, and J. Kan, "TARDB-Net: triple-attention guided residual dense and BiLSTM networks for hyperspectral image classification," *Multimedia Tools and Applications*, vol. 80, no. 7, pp. 11291–11312, 2021.

[23] R. Liu, X. Ning, W. Cai, and G. Li, "Multiscale dense cross-attention mechanism with covariance pooling for hyperspectral image scene classification," *Mobile Information Systems*, vol. 202115 pages, Article ID 9962057, 2021.

[24] X. Ning, Y. Wang, W. Tian, L. Liu, and W. Cai, "A biomimetic covering learning method based on principle of homology continuity," *ASP Transactions on Pattern Recognition and Intelligent Systems*, vol. 1, no. 1, pp. 9–16, 2021.

[25] X. Zhang, Y. Yang, Z. Li, X. Ning, Y. Qin, and W. Cai, "An improved encoder-decoder network based on strip pool method applied to segmentation of farmland vacancy field," *Entropy*, vol. 23, no. 4, p. 435, 2021.

[26] Y. Ding, X. Zhao, Z. Zhang, W. Cai, and N. Yang, "Multiscale graph sample and aggregate network with context-aware learning for hyperspectral image classification," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 14, pp. 4561–4572, 2021.

Hindawi
Scientific Programming
Volume 2021, Article ID 2934943, 8 pages
https://doi.org/10.1155/2021/2934943

Research Article

Group Behavior Pattern Recognition Algorithm Based on Spatio-Temporal Graph Convolutional Networks

Xinfang Chen [1] **and Venkata Dinavahi** [2]

[1] *College of Information Engineering, Institute of Disaster Prevention, Sanhe 065201, Hebei, China*
[2] *Department of Electrical & Computer Engineering, University of Alberta, Edmonton T6G 1H9, Alberta, Canada*

Correspondence should be addressed to Xinfang Chen; chenxinfang@cidp.edu.cn

Received 25 April 2021; Revised 29 May 2021; Accepted 5 June 2021; Published 8 July 2021

Academic Editor: Shah Nazir

With the rapid growth of population, more diverse crowd activities, and the rapid development of socialization process, group scenes are becoming more common, so the demand for modeling, analyzing, and understanding group behavior data in video is increasing. Compared with the previous work on video content analysis, factors such as the increasing number of people in the group video and the more complex scene make the analysis of group behavior in video face great challenges. Therefore, a group behavior pattern recognition algorithm based on spatio-temporal graph convolutional network is proposed in this paper, aiming at group density analysis and group behavior recognition in the video. A crowd detection and location method based on density map regression-guided classification was designed. Finally, a crowd behavior analysis method based on density grade division was designed to complete crowd density analysis and video group behavior detection. In addition, this paper also proposes to extract spatio-temporal features of crowd posture and density by using the double-flow spatio-temporal map network model, so as to effectively capture the differentiated movement information among different groups. Experimental results on public datasets show that the proposed method has high accuracy and can effectively predict group behavior.

1. Introduction

The growth of population and the diversity of crowd activities make group scenes become common. Group behavior [1–3] contains many important clues in interdisciplinary fields. Understanding the formation mechanism of group behavior has long been one of the important research topics in sociology and natural science. When the number of people in the video increases and the crowd scene becomes more complex [4], how to automatically and effectively model, analyze, and understand the group behavior data so as to better serve human beings becomes an important challenge. Research on group behavior analysis can provide support and corresponding solutions for many key engineering applications, such as intelligent video surveillance, crowd anomaly monitoring, and public facility planning. From the perspective of the cognitive mechanism of group behavior, this paper studies the effective computational framework and algorithm model of group behavior, trying to mine the dynamic group pattern and behavior in the real scene video data [5, 6], so as to solve the practical problems in the field of computer vision.

At present, the problem of group behavior analysis [7, 8] in video is based on ordinary surveillance video, and it uses computer vision technology to understand and analyze group behavior and events in the monitored scene. This changes the problem that traditional video surveillance relies too much on manpower. It can automatically realize the analysis and description of group behavior and realize the intelligent monitoring of large-scale crowd scenes. Group behavior analysis and understanding has become an important research branch of video surveillance, which has been widely applied in many fields such as public security, transportation, and facility planning. At the same time, the vigorous development of artificial intelligence [9–12], machine vision [13–15], cognitive science, and other cutting-

edge technologies also provides a guarantee for intelligent understanding of video content. Previous behavior understanding work [16] in video content analysis mainly focused on understanding individual behaviors, such as motion detection, target tracking, and object recognition, while ignoring the understanding of large-scale group behaviors. Compared with the analysis and recognition of individual behavior, group behavior is more real and complicated.

In the detection of crowd density [17, 18] and crowd behavior [19] by a computer, the movement of the crowd is complex and the scene is changeable. Due to the change of illumination, the blocking of the crowd, the perspective effect, the different shooting angles, and other factors, it will bring difficulties to the detection by the computer. Crowd behaviors have different semantics in different scenes. It is of great significance to quickly and effectively understand and distinguish the semantics of normal and abnormal behaviors of crowds and realize effective judgment, which is an urgent problem to be solved in the field of computer vision [20].

The processing of video images through computer vision can further replace manual monitoring to perform real-time and efficient monitoring of crowd density and crowd behavior. Recently, many scholars have used deep learning-based methods to conduct research on multiple tasks such as pedestrian detection, face recognition, and group behavior recognition and have made major breakthroughs. At present, when computer vision performs crowd detection, there are problems such as large crowd, poor detection accuracy, variable scenes, and high complexity. The existing technology can effectively overcome the abovementioned difficulties on the basis of deep learning and affect the distribution of the population. As well as by real-time monitoring of behavior, it provides solutions for crowd supervision, which has great practical significance and application value.

The main innovations and contributing points of this paper is to propose a group behavior pattern recognition algorithm [21, 22] based on spatio-temporal graph convolutional network, which can effectively recognize group behavior. The paper also proposed to use the dual-stream spatio-temporal map network model to extract spatio-temporal features of the crowd posture and density to effectively capture the differentiated movement information between different crowds.

The paper is organized as follows. Section 2 represents briefly the related work to the proposed research. Section 3 elaborates the methodology of the paper with details in sections. Experiments and results of the paper are given in Section 4. The paper is concluded in Section 5.

2. Related Work

The initial population research is mainly based on the detection of the crowd. The image is segmented before the target detection of the crowd using a sliding window, and finally, the crowd is counted based on the classifier. Detection-based methods include detection based on the whole [23] and detection based on parts of the human body [24]. The typical traditional method uses random forest matrix, SVM detector, and other methods to train the classifier and extracts various features such as pedestrian direction gradient histogram, edge, texture, and whole body wavelet. In scenes with highly dense crowds, crowds are severely occluded, and the method of detecting parts of the human body such as the head and shoulders is used instead of the method based on overall detection. The effect is improved, but the robustness of human detection is still not high.

Crowd density analysis and crowd counting based on regression are mainly used to learn the mapping relationship between image features and number of people [25]. Image segmentation is based on the regression method first, the image, texture, edge, and the prospect of gradient low-level features such as extraction and then the linear regression, Gaussian regression, ridge regression, and regression function are studied, such as learning exists in the mapping function of the number of low-level features and the image, generating a static background model, which is sensitive to illumination changes. The model needs to be retrained each time the scene is transformed, which is costly in terms of time and computation. Regression-based methods usually believe that the relationship between the number of people in the image and the foreground area can be approximately linear. However, such linear relationship is difficult to be established because of the problems of occlusion, overlap, and perspective of the crowd in the real scene.

In densely crowded images, deep learning usually uses convolutional neural networks to generate end-to-end models to extract features of different scales of pedestrians in the image, so as to generate crowd density maps through Gaussian kernel functions to achieve the effect of crowd counting [26]. The crowd density map can not only realize crowd counting but also provide rich spatial information, detect crowd density distribution, and further analyze crowd behavior through crowd density detection. Zhang et al. [27] proposed a multicolumn deep convolutional neural network MCNN model, which used different subnetworks with different convolutional kernel sizes to realize crowd count in the scene of serious crowd occlusion and height transformation. In the latest research, Sam et al. [28] proposed a switching network based on MCNN, which has multiple CNN subnetworks with different depths and different convolution kernel sizes of each subnetwork, thus improving the accuracy and robustness of crowd density analysis and crowd count results of high occlusion and multiscale scene transformation. Sindagi and Patel [29] proposed a context pyramid model CP-CNN. In order to extract global and local context information, the network learns the MCNN network of multicolumn architecture, designs two subnetworks to map the input image or video frame data to a high-dimensional feature map, and uses the CNN network to estimate the context at all levels. To reduce technical errors and generate higher quality density maps, Li et al. [30] proposed the deep neural network [31–35] model CSRNET, which abandoned the multicolumn framework, and believed that the multicolumn framework had no obvious advantages compared with the single-column framework. The front end of the model was the VGG-16 model, which abandoned the full connection layer and only retained the convolutional

layer and pooling layer, followed by the void convolution to expand the receptor field and obtain the features of different levels of images. Generate the population density distribution map, and obtain better detection results.

3. Methodology

3.1. Overall Framework. Crowd flow in video has the characteristics of time dynamic, space correlation, and uncertainty. Aiming at these characteristics, this paper proposes a kind of spatio-temporal dynamic graph convolutional network [36, 37] to study and predict crowd flow. Figure 1 shows the framework of spatiotemporal dynamic graph convolutional network (STDGCN) proposed in this paper. The STDGCN model consists of an input transformation layer, an STDGCN layer, and an output layer composed of a full connection layer.

The model uses the spatio-temporal data collected by the crowd flow sensor in the video and external factors to predict the crowd flow and other parameters in the future and comprehensively obtain the spatio-temporal network [38] prediction output. The input conversion layer embeds and converts crowd flow attribute data and exogenous factor data, among which three types of data are used for exogenous factors. The STDGCN layer contains a graph convolution module and a time-dimensional encoder-decoder structure. The output layer generates the prediction result of each node through a fully connected layer.

The core ideas of the STDGCN model can be summarized in the following two points. First, regard the sensor data at the same time as a graph data, connect the nodes and neighbor nodes to represent the spatial correlation of the crowd flow, and use the graph convolutional network to capture traffic. Second, treat the data at different moments of the same node as a time series and use the gated recurrent unit and attention mechanism to deal with the time dynamics of the traffic flow. The STDGCN layer structure is shown in Figure 2.

The spatio-temporal dynamic graph convolution module consists of two parts: graph convolutional network (GCN) and attention encoder network (AEN). Graph convolutional network is used to deal with the spatial dependence of crowd data, and attention encoder network is used for capture time dimension dynamics.

3.2. Spatial Feature Extraction of Crowd. Compared with the use of two-dimensional image convolution to obtain the patterns and characteristics of the crowd, the pedestrian sensor data with the characteristics of map data can obtain more primitive and real spatial attributes. In the proposed model STDGCN, graph convolution is directly applied to graph structure, and highly meaningful patterns and features are extracted in the spatial domain. Traditional convolutional neural networks can effectively extract local features of data, but they are not suitable for general graph structures. There are two types of methods to generalize convolutional neural networks to graph structures. One method is to expand the spatial definition of convolution, and the other is to use the Fourier transform of the graph to operate in the spectral domain.

The spectrogram method of graph convolution is to use the diagonalized linear operator defined in the Fourier domain to convolve the graph signal and use the convolution kernel g_θ. The convolution operation on the graph signal G_v can be expressed as

$$g_\theta \times G_v = g_\theta\left(U \wedge U^T\right)v = U g_\theta(\wedge) U^T v, \tag{1}$$

where U is the Fourier basis composed of eigenvectors and $g_\theta(\wedge)$ is the diagonal matrix composed of eigenvalues of L. Because the scale of the graph becomes larger, that is, when the crowd is large, the computational complexity of eigendecomposition of the Laplace matrix in equation (1) is very high, which can be approximated by Chebychev polynomial:

$$g_\theta \times G_v = g_\theta(L)v \approx \sum_{k=0}^{K-1} \theta_k T_k(\widetilde{L})v. \tag{2}$$

3.3. Time Feature Extraction of Crowd Flow. As shown in Figure 2, after the crowd flow data is extracted through the graph convolutional network for spatial feature extraction, the spatial feature sequence and the embedding representation of exogenous factors are used as the input of time dimension modeling. The AEN module is composed of two GRU networks with independent parameters. The GRU network on the left is the encoder module, and the GRU network on the right is the decoder module. The encoder encodes the input time sequence and initializes the decoder through the last moment of the encoder. Module, the decoder generates prediction output from the context vector in time steps:

$$h_{i,t} = \mathrm{GRU}\left(\left[y_{i,t}; e_{i,t}\right], h_{i,t-1}\right), \tag{3}$$

where $h_{i,t}$ is the output representation of sensor node i at time t, $y_{i,t}$ is the feature sequence obtained by the graph convolution operation, and $e_{i,t}$ is the exogenous factor.

A potential problem of the encoder-decoder model is that the model needs to be able to compress the context information of the source sentence into a fixed-length vector. This makes it difficult for the model to handle long sequences, especially those longer than the feature sequences in the training data. In other words, the encoder-decoder model may be difficult to grasp the longer periodic features in the crowd flow, such as weekly regularity.

3.4. Semantic Relevance of Group Behavior. In this paper, a spatio-temporal correlation model of video is designed to infer the behavioral semantics of group figures in video sequences. The model is composed of two layers of GRU. The first layer of GRU predicts feature mask sequence $E = \{E_1, E_2, \ldots, E_N\}$ to encode. The second layer GRU decodes the hidden codes of the first layer output one by one and outputs the action semantics of the characters and the

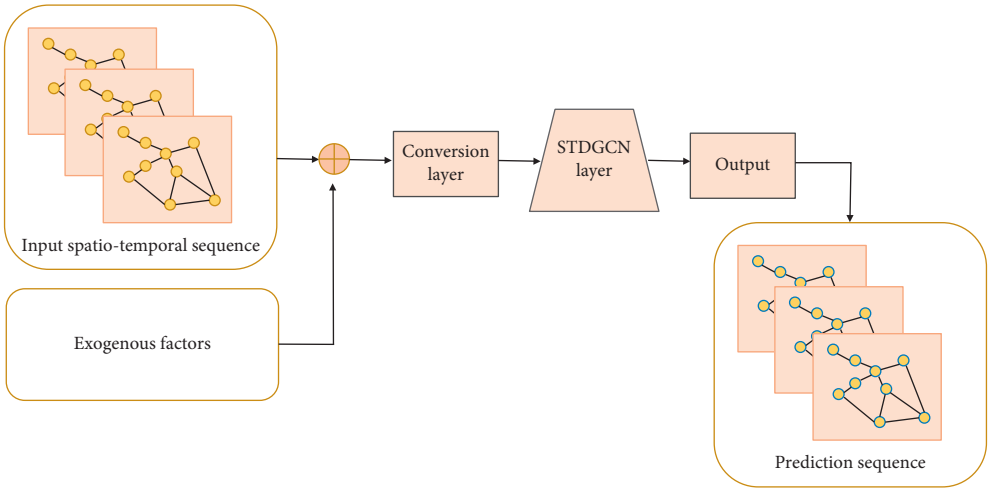


Figure 1: Schematic diagram of spatio-temporal dynamic graph convolutional networks.

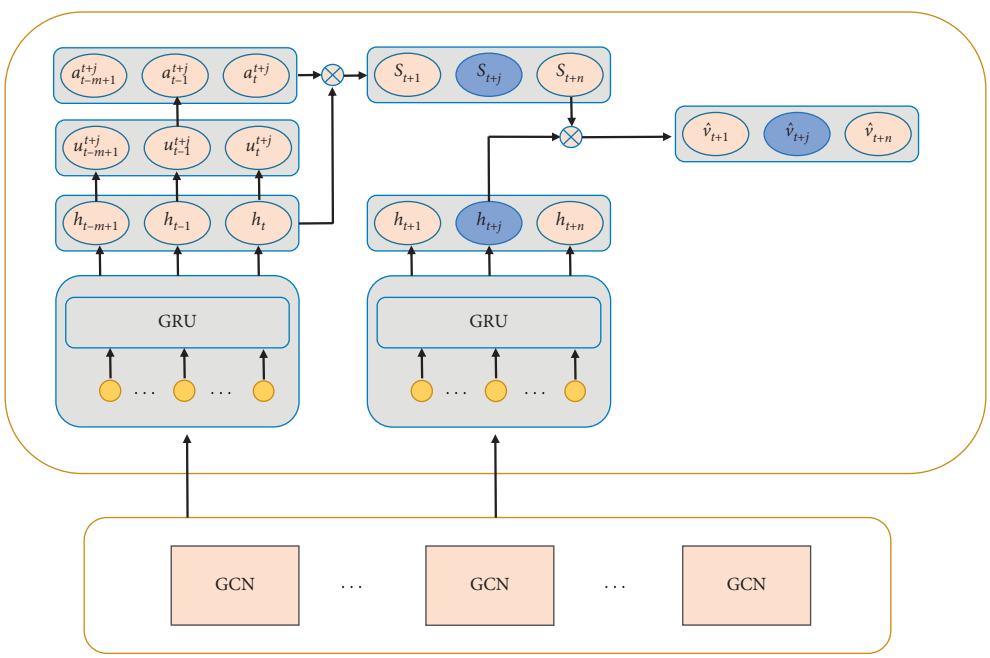


Figure 2: STDGCN layer structure.

behavior semantics of the group characters in the corresponding time sequence after spatio-temporal correlation. The model can be divided into two stages: encoding stage and decoding stage in the process of spatio-temporal correlation of group character behavior.

In the coding stage, the first part of the GRU structure of the first layer has a value of code_i^t, and the hidden layer information H_t is calculated. H_t includes the hidden layer information $h_t = \{h_1^t, h_2^t, \ldots, h_n^t\}$ of a single person and the hidden layer output g_t of the group of people. The equation for calculating the output of each character in the tth video frame is as follows:

$$p_i^t = x_i^t \oplus h_i^t, \tag{4}$$

where x_i^t is the predicted feature mask input of the person i in the tth frame of the first-layer encoding stage, h_i^t is the output result of the hidden layer of the person i in the tth frame of the video frame in the first-layer encoding stage, and $\oplus$ is the fusion function. The calculation equation of the hidden layer output structure of group character behavior is as follows:

$$g_t = p_1^t \vee p_2^t \vee \cdots p_N^t, \tag{5}$$

where p_i^t represents the character feature after fusion and $\vee$ is the maximum pooling operation.

In the decoding stage, H_t is output according to the semantic description of the action of the previous character and the hidden layer of the previous moment. Analyze the behavioral semantics of the group characters so that the group characters' behaviors after the GRU structure have temporal sequence information.

The obtained group character behavior semantic prediction probability set is calculated by a maximization equation, and the group character behavior semantics with the largest prediction probability is taken as the video group character behavior semantics:

$$\text{Group} = \arg\max\left(pt_{\text{group}}\right), \tag{6}$$

where pt_{group} is the set of semantic prediction probabilities of group character behavior we obtained.

4. Experiments and Results

4.1. Experimental Setup. The Volleyball data set is selected to verify the semantic extraction method of sparse group behavior based on the spatio-temporal trajectory of video. The Volleyball data set contains 55 real volleyball match videos and 4380 frame labels. The image size of each frame is 720×108, and each frame label contains the number of the current video frame. So, the position information of the player is composed of the coordinates of the upper left corner of the character's bounding box and the height and width of the bounding box. 3493 frame labels of the first 39 videos were used as the IJ} L training set, and 1137 video frames of the last 16 frames were used as the test set.

In the experiment, the length of the input video sequence is T, the individual action and group behavior semantics of N players are extracted, and $T = 10$ is defined. The first 4 frames and the next 5 frames, including the labeled video frames, are, respectively, taken as a video sequence fragment, and $N = 12$ is defined according to the characteristics of the volleyball match in the data set. All experiments in this section were developed using TensorFlow and run on Linux platform.

4.2. Evaluation Standard. In order to extract the semantics of sparse group behavior based on the spatial-temporal trajectory of video, the test is conducted with Volleyball data set and the mask position matching feature F_ code-B is used to complete the matching of people. The experimental results are compared with Inception and HDTM. It includes the comparison of the semantics of human action and the semantics of group action and takes the accuracy of the extracted semantic of group action and the semantic of individual action as the evaluation standard.

4.3. Experimental Results. Table 1 shows the accuracy comparison results of Inception, HDTM, and our algorithm, including two parts: group behavior semantics and individual action semantics. It can be seen from Table 1 that the algorithm in this paper is superior to the above two algorithms in terms of semantics of people's actions and group behaviors. Compared with the above two algorithms, the semantic accuracy of individual actions increases by 4.5% and 2.1%, and that of group actions increases by 8.3% and 1.8%. After integrating the related movement track of the group figures, the complete movement clues of the figures in the video sequence can be grasped by the accurate tracking of the group figures. Figure 3 is an example of successful semantic extraction of character actions and group behaviors in some videos of the data set. In Figure 3, the bounding box information and individual action semantics of each player in this video frame are specifically drawn, and the current group behavior semantics are marked.

4.4. Group Anomaly Recognition Experiments. Aiming at the evaluation of the detection effect of abnormal motion behavior in dense groups, this section uses the PETS 2009 data set containing sequence activities of different groups of people. The data set is divided into five parts: calibration, training, counting, density estimation, and crowd tracking. The video frame image has a resolution of 576×768, contains 9 videos, and has 152 abnormal data. The first 1134 frames of the experiment in this section are used as the training set, and the last 378 frames are used as the test set.

For the detection of abnormal gathering behaviors of dense groups, the experiments in this section are divided into image-level detection and pixel-level detection. Image-level detection gives abnormal aggregation detection results, and pixel-level detection can locate the abnormal gathering place and calculate the number of gatherings. Adopt the same data input and processing methods as the abnormal dispersion behavior.

The detection results of the abnormal dispersion behavior of dense groups on the PETS 2009 data set were quantitatively detected with the DBM algorithm based on optical flow and the D-IncSFA method based on deep learning. The results are shown in Table 2.

Experiments show that this paper has a good detection effect based on the abnormal dispersion behavior of crowd density distribution images and can detect abnormal video frames more accurately. Only when the movement speed exceeds a certain threshold can it be judged as abnormal. The specified movement speed is not less than 0.5 meters per second and more than 1.2 meters is abnormal dispersion behavior. The qualitative evaluation of abnormal dispersion behavior detection in dense groups is shown in Figures 4 and 5.

Table 1: Comparison of the accuracy of semantic extraction.

Methods	Personal action (%)	Group behavior (%)
Inception	78.1	75.5
HDTM	80.2	81.9
Ours	**85.6**	**86.2**

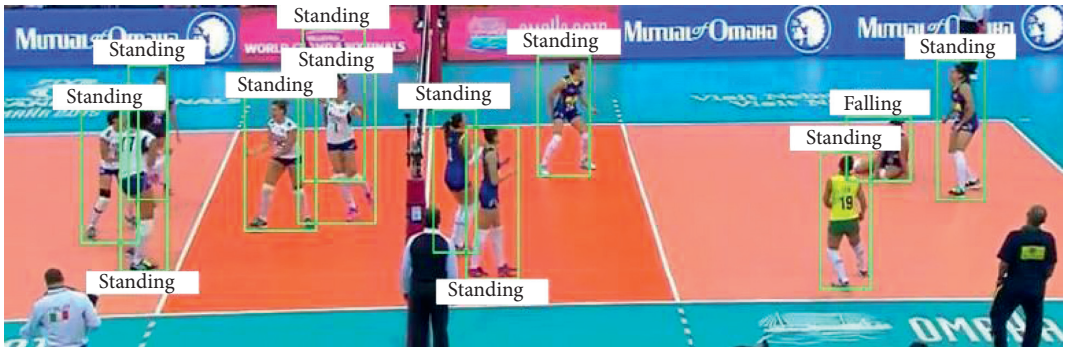
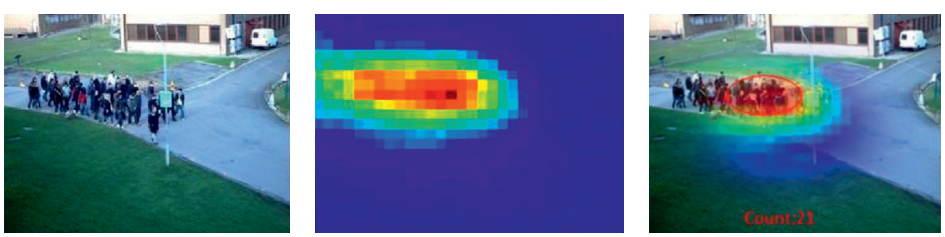


Figure 3: Successful example of group behavior semantic extraction.

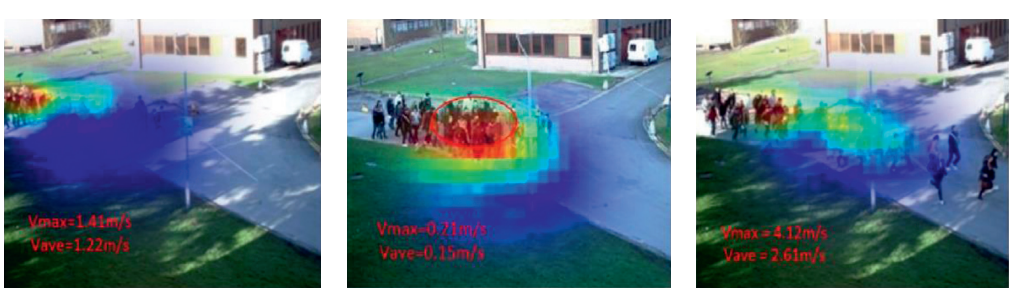
Table 2: Comparison of detection results of abnormal scattered behaviors of dense groups.

Methods	AUC
DBM	0.8770
D-IncSFA	0.9797
Ours	**0.9899**



(a) (b) (c)

Figure 4: Detection results of abnormal crowd behavior. (a) Test image. (b) Population density distribution map. (c) Density analysis and positioning.



(a) (b) (c)

Figure 5: Results of detection of abnormal scattered behaviors of dense groups. (a) Normal behavior, (b) abnormal aggregation behavior, and (c) abnormal dispersion behavior.

5. Conclusion

With the speedy growing population, various crowd activities, and the rapid development of the socialization process, group scenes are becoming more common. Due to this, the demand for analyzing, modeling, and understanding group behavior data in video is increasing. In this paper, we take group density analysis and group behavior recognition in video as the goal and propose a group behavior pattern recognition algorithm based on spatio-temporal graph convolutional network. We designed a crowd detection and positioning method based on density map regression guided classification and, finally, a crowd behavior analysis method based on density level division to complete crowd density analysis and video group behavior detection. In addition, this paper also proposes to use the dual-stream spatio-temporal map network model to extract spatio-temporal features of the crowd posture and density to effectively capture the differentiated movement information between different crowds. We have conducted experiments on public data sets, and the experimental results show that the method has high recognition accuracy and can effectively predict group behavior. The experimental results of the study have shown the effectiveness of the proposed research.

Data Availability

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

Conflicts of Interest

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

Acknowledgments

This work was supported by Special Funds for Basic Scientific Research in Central Universities (ZY20215126) and China Scholarship Fund.

References

[1] Y. Yuan, Y. Lu, and Q. Wang, "Tracking as a whole: multi-target tracking by modeling group behavior with sequential detection," *IEEE Transactions on Intelligent Transportation Systems*, vol. 18, no. 12, pp. 3339–3349, 2017.

[2] P. Ramdya, J. Schneider, and J. D. Levine, "The neurogenetics of group behavior in *Drosophila melanogaster*," *Journal of Experimental Biology*, vol. 220, no. 1, pp. 35–41, 2017.

[3] W. Yustisia, I. E. Putra, C. Kavanagh, H. Whitehouse, and A. Rufaedah, "The role of religious fundamentalism and tightness-looseness in promoting collective narcissism and extreme group behavior," *Psychology of Religion and Spirituality*, vol. 12, no. 2, pp. 231–240, 2020.

[4] Z. Pei, X. Qi, Y. Zhang, M. Ma, and Y.-H. Yang, "Human trajectory prediction in crowded scene using social-affinity long short-term memory," *Pattern Recognition*, vol. 93, pp. 273–282, 2019.

[5] U. Singh and M. K. Choubey, "Motion pattern recognition from crowded video," in *Proceedings of the 2020 8th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO)*, pp. 431–435, IEEE, Noida, India, June 2020.

[6] W. Wang, J. Shen, X. Lu, S. C. Hoi, and H. Ling, "Paying attention to video object pattern understanding," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2020, In press.

[7] H. Yao, A. Cavallaro, T. Bouwmans, and Z. Zhang, "Guest editorial introduction to the special issue on group and crowd behavior analysis for intelligent multicamera video surveillance," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 27, no. 3, pp. 405–408, 2017.

[8] H. Y. Swathi, G. Shivakumar, and H. S. Mohana, "Crowd behavior analysis: a survey," in *Proceedings of the 2017 International Conference on Recent Advances in Electronics and Communication Technology (ICRAECT)*, pp. 169–178, IEEE, Bangalore, India, March 2017.

[9] R. Liu, X. Ning, W. Cai, and G. Li, "Multiscale dense cross-attention mechanism with covariance pooling for hyperspectral image scene classification," *Mobile Information Systems*, vol. 2021, Article ID 9962057, 15 pages, 2021.

[10] Q. Liu, L. Cheng, A. L. Jia, and C. Liu, "Deep reinforcement learning for communication flow control in wireless mesh networks," *IEEE Network*, vol. 35, no. 2, pp. 112–119, 2021.

[11] X. Ning, X. Wang, S. Xu et al., "A review of research on co-training," in *Concurrency and Computation: Practice and Experience*John Wiley & Sons, Hoboken, NJ, USA, 2021.

[12] W. Cai, Z. Wei, R. Liu, Y. Zhuang, Y. Wang, and X. Ning, "Remote sensing image recognition based on multi-attention residual fusion networks," *ASP Transactions on Pattern Recognition and Intelligent Systems*, vol. 1, no. 1, pp. 1–8, 2021.

[13] Q. Luo, "Research on the teaching mode of physical education in colleges and universities," in *Proceedings of the 2017 2nd International Conference on Education, Sports, Arts and Management Engineering (ICESAME 2017)*, pp. 716–719, Atlantis Press, Beijing, China, June 2017.

[14] F. Muñoz-Bullón, M. J. Sanchez-Bueno, and A. Vos-Saz, "The influence of sports participation on academic performance among students in higher education," *Sport Management Review*, vol. 20, no. 4, pp. 365–378, 2017.

[15] Y. Cheng, G. Pang, B. Xiao et al., "Beyond triplet loss: person Re-identification with fine-grained difference-aware pairwise loss," *IEEE Transactions on Multimedia*, 2021, In press.

[16] M. G. Li, B. Jiang, Z. Che et al., "DBUS: human driving behavior understanding system," in *Proceedings of the 2019 IEEE/CVF International Conference on Computer Vision Workshop (ICCVW)*, pp. 2436–2444, IEEE, Seoul, Republic of Korea, October 2019.

[17] M. Marsden, K. McGuinness, S. Little, and N. E. O'Connor, "ResnetCrowd: a residual deep learning architecture for crowd counting, violent behaviour detection and crowd density level classification," in *Proceedings of the 2017 14th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS)*, pp. 1–7, IEEE, Lecce, Italy, August 2017.

[18] A. Almagbile, "Estimation of crowd density from UAVs images based on corner detection procedures and clustering analysis," *Geo-Spatial Information Science*, vol. 22, no. 1, pp. 23–34, 2019.

[19] W. G. Aguilar, M. A. Luna, J. F. Moya et al., "Real-time detection and simulation of abnormal crowd behavior," in *Proceedings of the International Conference on Augmented Reality, Virtual Reality and Computer Graphics*, pp. 420–428, Lecce, Italy, 2017.

[20] X. Ning, K. Gong, W. Li, L. Zhang, X. Bai, and S. Tian, "Feature refinement and filter network for person Re-identification," in *IEEE Transactions on Circuits and Systems for Video Technology*IEEE, Piscataway, NJ, USA, 2020.

[21] L. Zhou, X. Bai, X. Liu, J. Zhou, and E. R. Hancock, "Learning binary code for fast nearest subspace search," *Pattern Recognition*, vol. 98, Article ID 107040, 2020.

[22] C. Wang, X. Wang, X. Bai, Y. Liu, and J. Zhou, "Self-supervised deep homography estimation with invertibility constraints," *Pattern Recognition Letters*, vol. 128, pp. 355–360, 2019.

[23] H. Idrees, I. Saleemi, C. Seibert, and M. Shah, "Multi-source multi-scale counting in extremely dense crowd images," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 2547–2554, Portland, OR, USA, June 2013.

[24] V. Lempitsky and A. Zisserman, "Learning to count objects in images," *Advances in Neural Information Processing Systems*, vol. 23, pp. 1324–1332, 2010.

[25] L. Gao, Y. Wang, X. Ye, and J. Wang, "Crowd counting considering network flow constraints in videos," *IET Image Processing*, vol. 12, no. 1, pp. 11–19, 2017.

[26] P. F. Felzenszwalb, R. B. Girshick, D. McAllester, and D. Ramanan, "Object detection with discriminatively trained part-based models," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 32, no. 9, pp. 1627–1645, 2009.

[27] Y. Zhang, D. Zhou, S. Chen, S. Gao, and Y. Ma, "Single-image crowd counting via multi-column convolutional neural network," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 589–597, Las Vegas, NV, USA, June 2016.

[28] D. B. Sam, S. Surya, and R. V. Babu, "Switching convolutional neural network for crowd counting," in *Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 4031–4039, IEEE, Honolulu, HI, USA, July 2017.

[29] V. A. Sindagi and V. M. Patel, "Generating high-quality crowd density maps using contextual pyramid CNNS," in *Proceedings of the IEEE international conference on computer vision*, pp. 1861–1870, Venice, Italy, October 2017.

[30] Y. Li, X. Zhang, and D. Chen, "CSRNet: dilated convolutional neural networks for understanding the highly congested scenes," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1091–1100, Salt Lake City, UT, USA, June 2018.

[31] Y. Tong, L. Yu, S. Li, J. Liu, H. Qin, and W. Li, "Polynomial fitting algorithm based on neural network," *ASP Transactions on Pattern Recognition and Intelligent Systems*, vol. 1, no. 1, pp. 32–39, 2021.

[32] Z. Chu, M. Hu, and X. Chen, "Robotic grasp detection using a novel two-stage approach," *ASP Transactions on Internet of Things*, vol. 1, no. 1, pp. 19–29, 2021.

[33] X. Ning, Y. Wang, W. Tian, L. Liu, and W. Cai, "A biomimetic covering learning method based on principle of homology continuity," *ASP Transactions on Pattern Recognition and Intelligent Systems*, vol. 1, no. 1, pp. 9–16, 2021.

[34] Z. Huang, P. Zhang, R. Liu, and D. Li, "Immature apple detection method based on improved Yolov3," *ASP Transactions on Internet of Things*, vol. 1, no. 1, pp. 9–13, 2021.

[35] Y. Ding, X. Zhao, Z. Zhang, W. Cai, and N. Yang, "Multiscale graph sample and aggregate network with context-aware learning for hyperspectral image classification," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 14, pp. 4561–4572, 2021.

[36] W. Cai and Z. Wei, "Remote sensing image classification based on a cross-attention mechanism and graph convolution," *IEEE Geoscience and Remote Sensing Letters*, 2020, In press.

[37] A. Feng, Z. Gao, X. Song, K. Ke, T. Xu, and X. Zhang, "Modeling multi-targets sentiment classification via graph convolutional networks and auxiliary relation," *Computers, Materials & Continua*, vol. 64, no. 2, pp. 909–923, 2020.

[38] S. Wang, X. Yu, L. Liu, J. Huang, and T. Jiang, "An approach for radar quantitative precipitation estimation based on spatiotemporal network," *Computers, Materials & Continua*, vol. 65, no. 1, pp. 459–479, 2020.

Hindawi
Scientific Programming
Volume 2021, Article ID 9957323, 9 pages
https://doi.org/10.1155/2021/9957323

Research Article

Cluster-Based Antiphishing (CAP) Model for Smart Phones

Mohammad Faisal [1] **and Sa'ed Abed** [2]

[1]*Department of Computer Science and IT, University of Malakand, KPK, Pakistan*
[2]*Computer Engineering Department, College of Engineering and Petroleum, Kuwait University, Kuwait City, Kuwait*

Correspondence should be addressed to Sa'ed Abed; s.abed@ku.edu.kw

Received 30 March 2021; Accepted 24 June 2021; Published 8 July 2021

Academic Editor: Shah Nazir

Different types of connectivity are available on smartphones such as WiFi, infrared, Bluetooth, GPRS, GPS, and GSM. The ubiquitous computing features of smartphones make them a vital part of our lives. The boom in smartphone technology has unfortunately attracted hackers and crackers as well. Smartphones have become the ideal hub for malware, gray ware, and spyware writers to exploit smartphone vulnerabilities and insecure communication channels. For every security service introduced, there is simultaneously a counterattack to breach the security and vice versa. Until a new mechanism is discovered, the diverse classifications of technology mean that one security contrivance cannot be a remedy for phishing attacks in all circumstances. Therefore, a novel architecture for antiphishing is mandatory that can compensate web page protection and authentication from falsified web pages on smartphones. In this paper, we developed a cluster-based antiphishing (CAP) model, which is a lightweight scheme specifically for smartphones to save energy in portable devices. The model is significant in identifying, clustering, and preventing phishing attacks on smartphone platforms. Our CAP model detects and prevents illegal access to smartphones based on clustering data to legitimate/normal and illegitimate/abnormal. First, we evaluated our scheme with mathematical and algorithmic methods. Next, we conducted a real test bed to identify and counter phishing attacks on smartphones which provided 90% accuracy in the detection system as true positives and less than 9% of the results as true negative.

1. Introduction

A phishing attack is used to obstruct and limit legitimate user access to resources of service providers on global networks. A phish exploits the victim's systems resources to acquire confidential and nonconfidential data. The phishing attack can be single and standalone on the system resources or can be distributed, known as a phishing Distributed Denial of Services (DDOS) attack. Phishing attack focuses on a single system by using different launching pads [1]. Although it is not mandatory for both standalone and distributed phishing attacks to harm the data permanently and directly, it is certain that they deliberately compromise all the resource availability for cornerstone security services. Phishing attacks craft congestion in networks by spawning tremendous data traffic in the vicinity of the victim's system, which is adequate to thwart a legitimate packet from reaching its destination. In phishing attacks, the attacking traffic not only contaminates legitimate users but also attacks the target system, either to downgrade the system performance or, in some cases, to stop the service availability. Compared to other cyberattacks, phishing attacks are harder and more complex to circumvent [2]. Most often, phishing attacks exploit network bandwidth and connectivity, downgrading performance in systems they are compromising, whether during network-based attacks or host-based attacks, as shown in Figure 1. Consequently, phishing attacks are successful at halting, interrupting, and demoting the real-time performance of the system by draining all its resources [3].

Figure 1 shows multiple-incident level attacks that can occur on smartphones. Among these attacks, a phishing attack is the worst as its damages smartphones compared to other attacks such as ransomware, backdoors, Denial of Service (DOS)/(DDOS), bot activity, and worm propagation. A phishing attack harms smartphones in both active and

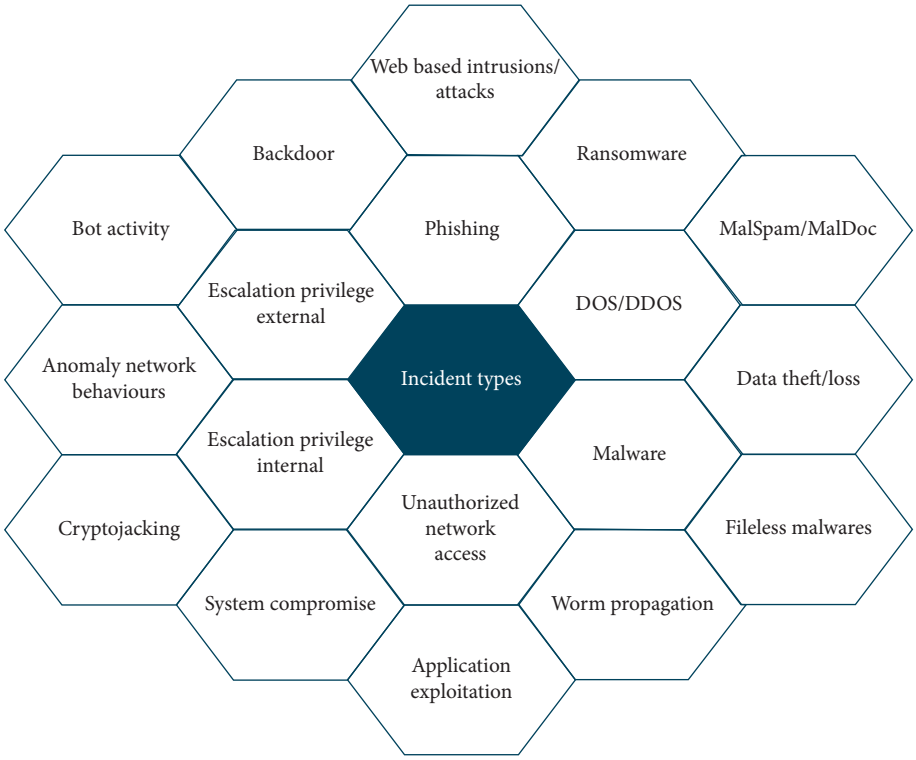


Figure 1: Smartphones incident types.

passive attack scenarios. In active attack scenarios, a phishing attack damages all the data content of the smartphones, while in passive attack scenarios, a phishing attack uses the smartphone as a launching pad against other systems after compromising the system. The additional attacks mentioned in Figure 1 target the system in either active or passive mode. For example, other attacks in Layer 1 are DOS/DDOS, escalation privileges (internal/external), unauthorized access, and malware, which are active attacks only. Similarly, if we examine Layer 2 of Figure 1, different attacks are seen, which can be easily classified as active or passive. In this work, the challenge in all the various incident types is the phishing attack, where an intruder can harm the victim in both an active and passive manner.

Confidentiality, integrity, and availability (CIA) are the main trivet underlying security services. Avoiding any one of these tenets deliberately or inadvertently might lead to open security breaches and unresolved vulnerabilities, consequently leading to a loss of credentials, reputations, and financial gains. Hence, the focus on web security against phishing attacks is necessary and must counter the latest exploitation techniques [3]. Among all cyberattacks, phishing attacks appear friendly but target financial transactions and highly confidential data. Their exploitation might result in the disappearance of financial gains and critical losses. Phishing attacks employ spam emails that attack online banking and money transfer accounts, all of which contain secret pins and passwords for user authentication, which is why hackers focus on banking and money transfer websites [4].

Cluster-based searching is vital, particularly in cybersecurity, because the algorithms use Markov chains to rank the data into clusters. Also, cluster-based searches work on probability, in which clustering is more appropriate for the query in process. Cluster-based searches also save battery life and execution time for smartphones.

The following are the objectives of this work:

(1) To counter the phishing attacks on smartphones.

(2) To falsify the fake websites containing phishing scams into a cluster form.

(3) To report the true positives and true negatives.

(4) To gray list and then blacklist the phishing attack links

The paper is organized as follows: Section 2 elaborates the literature review with critical analysis along with the security trial parameters. Section 3 explains the proposed scheme. Experimental results are evaluated with the help of detailed algorithms and mathematical and statistical methods. With assistance from the Weka tool and JavaScript language, a real test bed experiment was conducted using datasets from UCI and Mendeley, which are discussed in Section 4. Finally, Section 5 summarizes the key findings and presents our future trends.

2. Literature Review

Tools such as the Global System for Mobile (GSM), General Packet Radio Services (GPRS), enhanced data rates for GSM evolution (EDGE), Universal Mobile Telecommunication

System (UMTS), Bluetooth, and infrared make smartphones as a device of connectivity. However, this connectivity also serves as a gateway for malware, gray ware, and spyware. The GSM global communications expertise of another generation-2G enables the messaging among smartphones and sorted locations by exchanging the subsystem's replaced first-generation (1G) analog-centered facilities for a numerical, complete duplex and circuit substituted network for voice telephony [4].

GPRS 2.5-generation technology developed to improve the data rates and decrease the connection access time for 2G. Implementing the packet switching mechanism and introducing the Wireless Access Protocol (WAP) and Multimedia Messaging Services (MMS). EDGE improved GPRS features with an enhancement of its data rates and service reliability [5].

The UTMS, developed in 2002 attained a data rate perimeter of 2 Mbps together for packet and circuit-switching networks sustained concurrently. Several facilities can be entered instantaneously by the consumer such as streaming, discussions, and collaborations with colleagues. Bluetooth was industrialized in 1999, grounded on radio-broadcast small wavelengths customary for data communication as well as private area networks. Bluetooth provided an optimal level of security and a small array of communications up to 100 meters with negligible charge and ingestion [6, 7].

Based on the legality, delivery methods, and user authentication, mobile threats are classified into three main categories: malware, gray ware, and spyware, with regard to assorted attack vectors, motivations, and defense mechanisms [8].

For launching cyberattacks, all illegitimate activities such as spam emails and messages and twitter messages are exploited for smartphones. These cyberattacks either damage all the data contacted in the smartphones or compromise the smartphone to use it as a launching pad against other platforms. In [9–11], all schemes used data-mining techniques in different ways to counter cyberattacks on smartphones.

In [12, 13], the authors highlighted phishing attacks targeting smart grids to launch a phishing attack and compromise the system data to roll back the system. The authors proposed a data-mining technique to identify and counter phishing attacks and falsify fake pages and messages.

Table 1 shows the critical evaluation of the literature review with respect to the classifications. The table consists of six columns, in which the first column states the approach of the scheme with its reference, while the second column explains the classification of the category where the scheme is occurring.

In the third column, the mechanism of the antiphishing scheme and how the scheme is working are briefly stated. In the fourth column, the contribution or the strength of the scheme is highlighted with its achievements. In the fifth column, the weaknesses or the limitations of the scheme are mentioned with its possible vulnerabilities. In the sixth column, we mentioned the implementation scenario that in which scenario the scheme is implemented. In the last column, there is the tool/technology used by the scheme mentioned in the literature.

The main emphasis of malware is to annoy the genuine consumers, damaging the platform, cutting the reserved data, or misusing the scheme or policy susceptibilities irrespective of any notice to the victim users. Malware includes viruses, worms, Trojans, rootkits, and botnets. A computer virus is defined as a self-replicating piece of code, and a worm is a self-copying program [19], Trojans impersonate software that appears to provide services but in reality is a malicious program. A rootkit installs Trojans after which it then disables firewalls and antiviruses. Finally, botnets are a complete set of device viruses that infect victims for organized crime, consisting of a group of "zombies"; each is an infected computer or device [20].

Malware is prohibited in numerous countries, such as the United States, and in some cases of malware sharing, jail sentences have been administered [21].

Determining the position of a node and retrieving its history for a specific span of time is the main objective of the spyware. Depending on the practical circumstances, spyware may be genuine or illicit. For example, if a person is going to install personal spyware on his children or spouse's smartphone, the spyware is not going to cheat the victim. However, if the spyware is installed without the user's consent and successfully gains access to the device, sending confidential information to the intruder rather than the real author, then it is illegitimate [22].

Accumulating consumer evidence for the sole tenacity of summarizing and then advertising are the main intentions of gray ware, as indicated in Table 2. The gray ware distributor's corporate objectives are not to harm the user but rather to provide some sort of functionality and importance to the host user. If a user finds that the data collection process of gray ware is questionable, the user can complain and block the services of the gray ware. In contrast to malware and spyware, the illegal use of gray ware is punished by fines rather than any personal statements in countries where there is a rule of law. Therefore, gray ware is sometimes identified as laying at the boundary of legitimacy and illegitimacy. Based on the dogma of confidentiality and the consumer's rights of grievances, gray ware companies must disclose their compilation practices [23–26].

Another novelty of this research is that we classify schemes in a unique way that can be easily detected. We organized phishing attacks into multiple classes so that each category was tested against our proposed scheme of the CAP model. Based on the literature review, we classified the phishing attacks as shown in Figure 2. The main categories include Internet Protocol (IP), Uniform Resource Locator (URL), Domain Name System (DNS), certificate-based, social engineering, and technical maneuver [27–29].

In the first category, IP-based phishing attacks are classified. In the second category, URL-based phishing techniques are classified, which are subclassified into abnormal URL, URL of anchor, URL of long address, and repeating the same characters of the URLs. In the third category, we classified all phishing attacks based on social engineering. In the fourth category, we classified all phishing attacks that can be caused by technical maneuvers. In the fifth category, we classified phishing attacks that can occur from DNS poisoning. In the last category, we classified

Table 1: Critical review of the literature schemes.

Approach	Classification diagram	Mechanism	Contribution	Limitations	Implementation scenario	Technology/ algorithms used/tools
Network-level protection [14]	IP	Internet service providers database is used	Attack detection good, offender address list is up to date	Rule tuning, message content not verified	DNS	Snort
Authentication [15]	Certificates	User, domain, e-mail, and transaction based authentication based on digital signatures and hashing	Less complexity, no need of inter e-mail domains cooperation, enhance security	Vulnerable to man-in-the-middle attacks, technology constraints	Hotmail, Yahoo, Gmail	PGP, S/MIME
Client-side tools [16]	URL	Whitelisting and blacklisting	Legitimate e-mail will be acceptable only, best for already known phishing websites	High false positive and false negative rate for white- and blacklisting respectively	Mozilla, Firefox, and Internet Explorer browsers	Net craft, eBay toolbar, IE phishing filter
User education [17]	Social engineering	Online material, online test, and contextual training	Authority, attractive and impressive	False negative	All scenarios	Smart OS
Server-side filters and classifiers [18]	Technical maneuvers	Compare multiple classifiers and clustering techniques	Discover phishing attacks with narrow earlier knowledge	Time and space tradeoff	Internet browsers	Support vector machines

Table 2: Security trials of smartphones platforms (OS).

Security trials	Android	Nokia-Ovi	Apple iOS (iPhone OS)
Software installation markets	Allowed from both official Android market and unofficial sites	Allowed from both official Symbian market and unofficial sites	Allowed only from official sites Apple App Store
Grant permissions	Installation time-extensive fully	Only for unsigned Symbian	Runtime user permission only
Spyware and gray ware permissions	Allow and available at Android market	Allow	Disallow both
Premium SMS	Did not require user confirmation	Did not require user confirmation	Require user confirmation
SMS spam	Possible	Possible	Not easily allowed

attacks that can occur via digital certificates. This category is subclassified into Secure Socket Layer (SSL) and centralized authority certificates.

3. Proposed Schemes

A novel cluster-based antiphishing (CAP) mechanism, which addresses the following concerns, is necessary to secure users from cybersecurity attacks. For example, to thwart admittance of phishing websites/phishing attacks, to shield vital e-mail communication from phishes, to perceive deceptive website via (a) Appropriate Domain Name System and IP/MAC addresses toning and (b) authentication followed by authorization of website, data trickle consequential from device damage or robbery, inadvertent confession of data from smartphones, assault on decommissioned smartphones, mitigate spyware attacks, supervise network spoofing attacks, and financial malware attacks.

Our scheme operates in three main phases. In the first phase, the classification of incoming data is observed on the basis of a map provided in the second phase. In the second phase, packets are being clustered into their classified groups. In the third phase, digital forensics of the malicious packets are investigated, tracking back the culprits for future blacklists or recovery that can be used as a honey pot. As our scheme focuses on smartphones, which are capable of using only lightweight software, we designed our novel solution for implementations on base stations (centralized) rather than smartphones (distributed). This focus allows for placing all of the mechanisms in one package because the smartphone market is full of variants with different architectures of software and hardware. For smartphones, not only the attack (phishing) is distributed but also the tools and techniques are multiple and distributed in nature, such as social engineering and website spoofing techniques.

4. Results and Discussion

The CAP mechanism is elaborated in the following three evaluation methods: (1) algorithms, (2) mathematical and statistical formulae and tools (e.g., SPSS), and (3) test bed

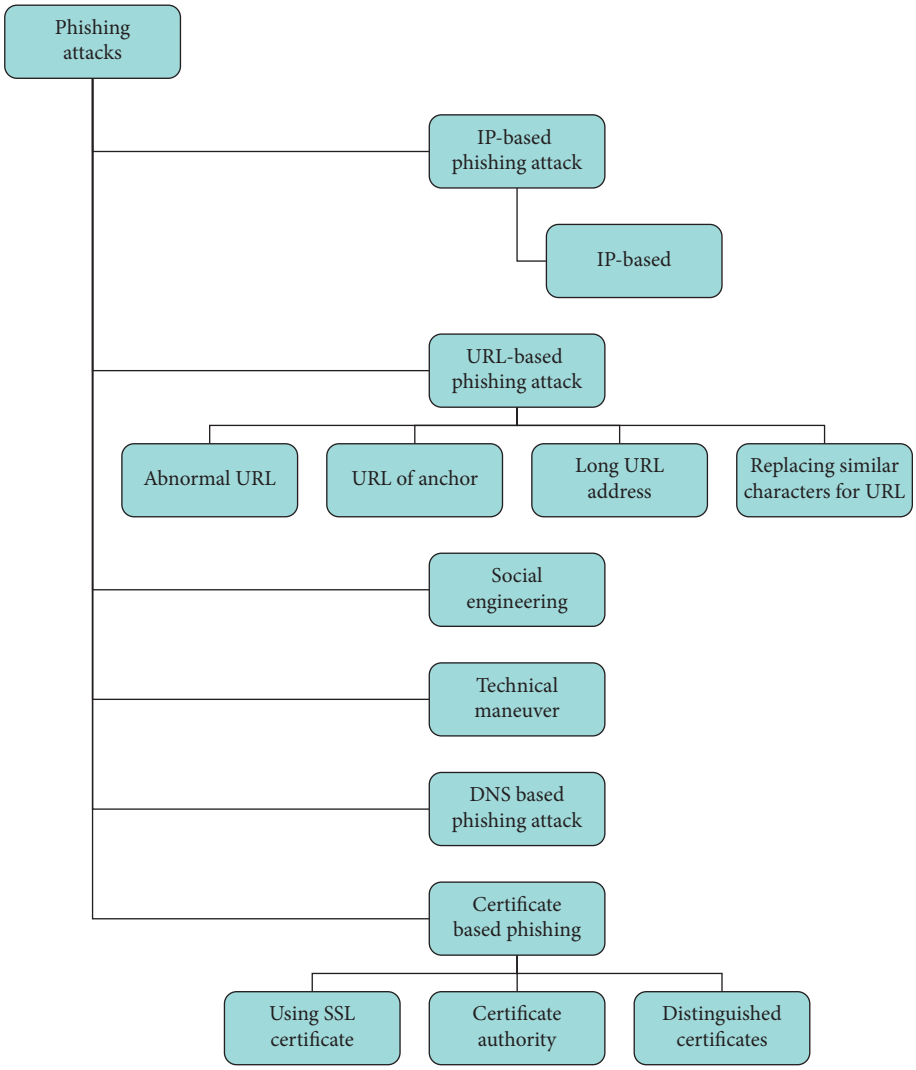


Figure 2: Classification of phishing attacks.

implementation via the Wireshark tool. The evaluation matrix consists of the following four main components: "true positive" measures the rate of correctly detected phishing attacks relative to overall prevailing phishing attacks; "true negative" measures the proportion of appropriately noticed genuine occurrences in relation to completely prevailing genuine occurrences; "false positive" measures the proportion of genuine occurrences that are inaccurately identified as phishing attacks relative to completely prevailing genuine occurrences; and finally, "false negative" measures the rate of phishing attacks that are inaccurately noticed as genuine relative to entirely prevailing phishing attacks.

4.1. Algorithm. In this section, the pseudocode of the CAP mechanism algorithm exploiting the IRC messenger of a smartphone is shown in Algorithm 1. In Step 1, we defined all the parameters involved in the execution of our scheme: PS, N, S, E, C, M, and P, all of which are labelled as described. In Step 2, the values received by the scheme as inputs will be validated, and each entry is executed as with FOUR subsequent IF and ELSE conditions. For example, if the argument E is received, then it means end times of communication or NULL value; if the argument C is received, then the channel name is identified through which the communication is required; if argument M is received, then the Internet Relay Chat (IRC) messenger will be communicated; and if argument P is received, then the port number is received through which the communication will be considered. In Step 3, once the communication is initiated with the IRC messenger, then a connection object is created in Step 4. Subsequently, a channel is created for communication between the nodes. In Step 5, an event handler is activated that is already defined in JavaScript to monitor malicious activities. The communication is countered if it is malicious; otherwise, it will proceed as normal.

4.2. Mathematical and Statistical Model. In our mathematical and statistical model, we attempted to best generalize the CAP model to all possible incoming phishing attacks.

//PS: packet size defined in RFC, N: number of packets (date rate) defined in RFC, S: start time of the data communication, E: end time of data communication, C for channel name, M for IRC messenger/IRC name, P for port number.

(1) Assignment and validating entries/values of the parameters defined
$\sum(S, E, C, M, P)$ IF $S \geq 0$ THEN $\int T = 0$ IF $E = 1$ THEN ends time of communication ELSE 0, IF $C = 1$ THEN channel name ELSE 0, IF $M = 1$ THEN IRC messenger name ELSE0, IF $P = 1$ THEN port number ELSE 0.

(2) $\int \text{IRC} \sum \text{IRC} \geq 0$

(3) $\int C = 1\, E \leq D$

(4) $€E \int DP$

(5) $\int E = M$ then $\sum M = 0$

Algorithm 1: Phishing attack detection using the CAP algorithm.

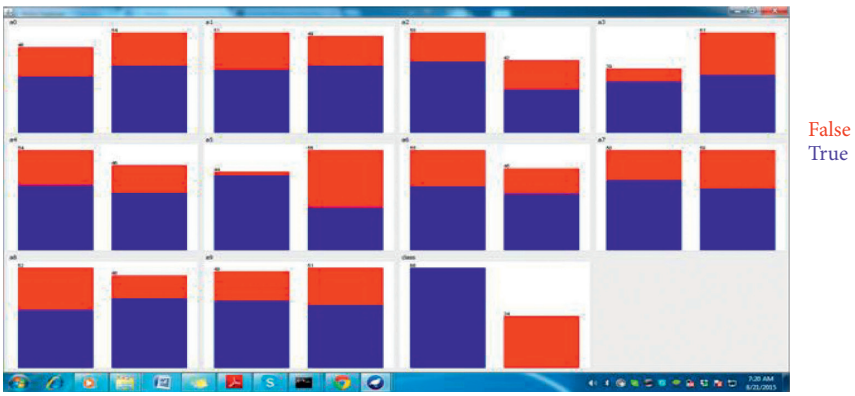


Figure 3: Weka tool results after a phishing attack on a smartphone.

We checked the maximum size and threshold value of the packets as well as their starting and ending limits. Furthermore, we tested combinations of various types of attacks from multiple resources and finally, the probability of the attack to occur:

For all $\sum$ of all bits = packet maximum size (defined in RFC).

$\int$ starting and ending of packets ratio = threshold (defined n RFC).

"Starting" indicates the lower limit and "ending" means the upper limit.

nCr combination of different packets from various resources.

$n\text{Cr} = n!/r!\ (n - \text{r})!$

Where n = for all incoming packets and r = malicious packets.

nPr permutation of different malign and nonmalign packets.

$n\text{Pk} = n!/(n - k)!$

Hence, the probability of launching a successful attack.

4.3. Implementation via Weka Tool with Results. Figure 3 shows the results taken by the Weka tool after the successful launching of a phishing attack on a smartphone. To validate our scheme, we considered the following system setup. The principal server where the DNS is laid on is considered as the main target of a phishing attack. The clients are exhausting a JavaScript code to implement bots on the server, while keeping the JavaScript code in an indefinite loop within the malicious code. A Wireshark instrument mounted on the server notices the scheme's position and deployment before and after the attack. Results are displayed in both scenarios in the real test bed implementation phase.

Table 3: Techniques' descriptions.

A1DE	Average one dependency estimator
NB	Naïve Bayes
MLP	Multilayer perceptron
QDA	Qualitative Descriptive Analysis
SVM	Support Vector Machine
WiSARD	WiSARD
CHIRP	COMPOSITE HYPERCUBE ON ITERATED RANDOM PROJECTION
DecisionTable	DecisionTable
DTNB	DecisionTable\Naïve Bayes
Ridor	Ridor
J48	Decision Tree

JavaScript code used for phishing attack: the malicious code below was used against our CAP model to test its capability for countering a phishing attack.

```
<html> <head>
hello me PHISHING attack.
</head> <SCRIPT language = JavaScript>
var name = prompt ("R u ready", "Name");
</SCRIPT> <BODY> </BODY> </html>
```

The results are shown in Figure 3, in which the blue shows the true positive data while the red shows the false positive data successfully encountered by our CAP model.

Table 4: Statistical data from UCI datasets.

Serial number	Technique	Instances			Precision	Accuracy (%)
		Correctly classified	Incorrectly classified	Total		
1	A1DE	1140	213	1353	0.81	84.2572
2	NB	1107	246	1353	0.792	81.8182
3	MLP	1153	200	1353	0.847	85.218
4	QDA	1130	223	1353	0.828	83.5181
5	SVM	1131	222	1353	0.83	83.592
6	WiSARD	1197	156	1353	0.886	88.4701
7	CHIRP	1171	182	1353	0.862	86.5484
8	DecisionTable	1133	220	1353	0.816	83.7398
9	DTNB	1172	181	1353	0.867	86.6223
10	Ridor	1204	149	1353	0.893	88.9874
11	J48	1215	138	1353	0.899	89.8004

Table 5: Statistical data from Mendeley datasets.

Serial number	Technique	Instances			Precision	Accuracy (%)
		Correctly classified	Incorrectly classified	Total		
1	A1DE	8724	276	10000	0.872	87.24
2	NB	8515	1485	10000	0.864	85.15
3	MLP	8694	306	10000	0.869	86.94
4	QDA	8657	1343	10000	0.881	86.57
5	SVM	8388	612	10000	0.839	83.88
6	WiSARD	8733	1267	10000	0.893	87.03
7	CHIRP	8633	367	10000	0.863	86.33
8	DecisionTable	8526	474	10000	0.853	85.26
9	DTNB	8641	359	10000	0.864	86.41
10	Ridor	8644	356	10000	0.864	86.44
11	J48	8731	269	10000	0.873	87.31

The CAP model produces more than 90% accuracy in its detection system, which has been classified as true positive (blue packets). Similarly, the CAP model also reduces the percentage of false negatives (red packets) to a single digit, that is, less than 9%.

4.4. Real Test Bed Experiment Conducted Using Datasets from UCI and Mendeley. As per the standard for assessment when evaluating the performance, a 10-time validation is completed for each classifier. A standard methodology examines a dataset by dividing data into 10 equal sizes, in which one set is used for testing and the other is used to train the data until each subset has been used for testing [30]. Investigating antiphishing revealed the following techniques for analysis, evaluation, and experimentations, as provided in Table 3. The 10 different techniques were then used for testing the UCI and Mendeley datasets.

4.4.1. Dataset Taken from UCI. Source Neda Abdelhamid, Auckland Institute of Studies, nedah '@' ais.ac.nz.

(1) Dataset Information. In online communications such as e-banking and e-commerce, phishing attacks are considered a threat. From 1353 websites, we have collected different issues related to legal and phishing websites (http://www.phishtank.com), where anyone can collect information about phishing attacks. The website used is Yahoo, in which 548 websites out of 1353 found were legitimate using a web script developed in PHP, 702 phishing URLs and 103 suspicious URLs. The results are shown in Table 4.

In the second column of Table 4, different techniques are considered for testing on the UCI datasets. As shown in the third column, the highest correctly classified technique used is J48, with the lowest incorrectly classification specified. However, regarding percentages, J48 is still less effective than our CAP model, which produces 90% true positive and 9% false negative results. As shown in the sixth column, the precision value of the J48 is 0.899, while the accuracy level in the last column is 89.8%. Similarly, if we consider any other technique applied to the UCI datasets shown in Table 4, none can reach the value of 90% accuracy. Therefore, we can deduce that the accuracy of our CAP model is much superior to any other of the latest techniques tested on UCI datasets.

4.4.2. Dataset Taken from Mendeley. Source Phishing web page: Phish Tank, Legitimate web page source: Alexa, Common Crawl.

(1) Dataset Information. In this scenario, the dataset under consideration is extracted from 10,000 websites; 48 features were extracted from 5000 phishing and legitimate websites. The results are shown in Table 5.

As a second test case, different techniques were considered for Mendeley datasets. As seen in the third column of

Table 5, J48 shows the highest precision (0.873) and accuracy at 87.31%, yet it is still less than our CAP model. We can conclude that by either changing the datasets or the technique, no method can achieve better than our proposed CAP model.

5. Conclusion

Fake commercial advertisements can play the role of honey pots for phishing attacks, as they behave like original finance and business sector advertisements. As the users follow, the fake websites and log on once, it is enough for the phishing hackers to steal the passwords and transact according to their own wishes and possibly changing login details. Currently, all phishing hackers must pass through some sort of Internet service providers (ISP), for which the administrator is responsible for countermeasures. Techniques such as content filtering, heuristics engines, IP blacklisting, and fingerprinting are currently employed; however, the problem of spamming followed by phishing is not yet contained. Neither of the schemes in the broad classification above focus on smartphone platforms. All the schemes in this work are silent about the digital forensics of the phishing attack. Our scheme successfully identified and maligned all the phishing attack packets in one single solution. The test bed results showed that the CAP model successfully identified and countered phishing attacks on smartphones. The CAP model produces more than 90% accuracy in its detection system, which has been classified as true positive. Similarly, the CAP model also reduces the percentage of the true negative to a single digit, namely, less than 9%. As a future trend, our model can be extended for Edge, FOG, and cloud computing environments as the CAP model is a lightweight scheme that can easily be integrated into such energy deficient computing zones.

Data Availability

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

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

[1] S. Afzal, M. Asim, A. R. Javed, M. O. Beg, and T. Baker, "URLdeepDetect: a deep learning approach for detecting malicious URLs using semantic vector models," *Journal of Network and Systems Management*, vol. 29, no. 3, pp. 1–27, 2021.

[2] A. K. Jain and B. B. Gupta, "A survey of phishing attack techniques, defence mechanisms and open research challenges," *Enterprise Information Systems*, pp. 1–39, 2021.

[3] A. G. Martín, A. Fernández-Isabel, I. M. de Diego, and M. Beltrán, "A survey for user behavior analysis based on machine learning techniques: current models and applications," *Applied Intelligence*, 2021.

[4] M. H. Siddiqi Abdullah, Y. S. Alhwaiti, I. Alrashdi, A. Ali, and M. Faisal, "Segmentation and classification of heart angiographic images using machine learning techniques," *Journal of Healthcare Engineering*, vol. 2021, Article ID 6666458, 9 pages, 2021.

[5] X. Liao, M. Faisal, Q. Chang, and A. Ali, "Evaluating the role of big data in IIOT-industrial Internet of things for executing ranks using the analytic network process approach," *Scientific Programming*, vol. 2020, Article ID 8859454, 7 pages, 2020.

[6] A. Das, S. Baki, A. El Aassal, R. Verma, and A. Dunbar, "SOK: a comprehensive reexamination of phishing research from the security perspective," *IEEE Communications Surveys & Tutorials*, vol. 22, no. 1, pp. 671–708, 2019.

[7] A. Rana, "Phishing attacks survey: types, vectors, and technical approaches," *Future Internet*, vol. 12, no. 10, p. 168, 2020.

[8] S. Jindal and M. Misra, "Multi-factor authentication scheme using mobile app and camera," *Advances in Communication and Computational Technology*, Springer, New York, NY, USA, 2021.

[9] R. Zhang, X. Wang, X. Yang, and X. Jiang, "Billing attacks on SIP based VoIP systems," in *Proceedings of the 1st USENIX Workshop on Offensive Technologies (WOOT)*, pp. 1–8, Berkeley, CA, USA, August 2007.

[10] Z. Zhu, G. Cao, S. Zhu, S. Ranjan, and A. Nucci, "A social network based patching scheme for worm containment in cellular networks," in *Proceedings of the 28th IEEE International Conference on Computer Communications, Joint Conference of the IEEE Computer and Communications Societies (INFOCOM 2009)*, pp. 1476–1484, Rio de Janeiro, Brazil, April 2009.

[11] K. A. Otunaiya and G. Muhammad, "Performance of data-mining techniques in the prediction of chronic kidney disease," *Computer Science and Information Technology*, vol. 7, no. 2, pp. 48–53, 2019.

[12] E. U. Soykan, M. Bagriyanik, and G. Soykan, "Disrupting the power grid via EV charging: the impact of the SMS Phishing attacks," *Sustainable Energy, Grids and Networks*, vol. 26, Article ID 100477, 2021.

[13] A. K. Ghazi-Tehrani and H. N. Pontell, "Phishing evolves: analyzing the enduring cybercrime," *Victims & Offenders*, vol. 16, no. 3, pp. 316–342, 2021.

[14] H. Shahriar and L. Etienne, "Presentation attack detection framework," *Machine Intelligence and Big Data Analytics for Cybersecurity Applications*, Springer, New York, NY, USA, pp. 297–311, 2021.

[15] A. Saxena, N. Sharma, P. Agarwal, and R. Barotia, "Phishing website prediction by using cuckoo search as a feature selection and random forest and BF-tree classifier as a classification method," *Rising Threats in Expert Applications and Solutions*, Springer, New York, NY, USA, 2021.

[16] G. K. Bhageria, V. Ekambaram, and S. K. Rakshit, *Protecting against Notification Based Phishing Attacks*, 2021.

[17] G. Shrivastava, K. Sharma, and S. Rai, "The detection & defense of DoS & DDoS attack: a technical overview," in *Proceedings of the ICC*, pp. 274–282, Kampala, Uganda, 2010.

[18] T. Peng, C. Leckie, and K. Ramamohanarao, "Survey of network-based defense mechanisms countering the Phishing and Phishing problems," *ACM Computing Surveys*, vol. 39, p. 3, 2007.

[19] D. C. W. M. Wozniak, N. C. N. Meghanathan, and D. Nagamalai, *Advances in Network Security and Applications, Springer, Chennai, India*, 2011.

[20] A. Shabtai, U. Kanonov, Y. Elovici, C. Glezer, and Y. Weiss, ""Andromaly": a behavioral malware detection framework for android devices," *Journal of Intelligent Information Systems*, vol. 38, no. 1, pp. 161–190, 2012.

[21] A. P. Felt, M. Finifter, E. Chin, S. Hanna, and D. Wagner, "A survey of mobile malware in the wild," in *Proceedings of the 1st ACM Workshop on Security and Privacy in Smartphones and Mobile Devices*, pp. 3–14, Chicaga, IL, USA, October 2011.

[22] P. Porras, H. Saïdi, and V. Yegneswaran, "An Analysis of the iKee.B iPhone Botnet," in *Security and Privacy in Mobile Information and Communication Systems*, pp. 141–152, Springer, New York, NY, USA, 2010.

[23] A. Apvrille, "Symbian worm Yxes: towards mobile botnets?" *Journal in Computer Virology*, vol. 8, no. 4, pp. 117–131, 2012.

[24] M. La Polla, F. Martinelli, and D. Sgandurra, "A survey on security for mobile devices," *IEEE Communications Surveys & Tutorials*, vol. 15, no. 1, pp. 446–471, 2013.

[25] G. Kambourakis, C. Kolias, S. Gritzalis, and J. H. Park, "DoS attacks exploiting signaling in UMTS and IMS," *Computer Communications*, vol. 34, no. 3, pp. 226–235, 2011.

[26] A. Bose and K. G. Shin, "Proactive security for mobile messaging networks," in *Proceedings of the 5th ACM Workshop on Wireless Security (WiSe'06)*, pp. 95–104, ACM, New York, NY, USA, September 2006.

[27] A. Bose and K. G. Shin, "On mobile viruses exploiting messaging and Bluetooth services," in *Proceedings of the Securecomm and Workshops*, pp. 1–10, Baltimore, MD, USA, 2006.

[28] A. Bremler-Barr, R. Halachmi-Bekel, and K. Kangasharju, "Unregister attacks in SIP," in *Proceedings of the 2nd IEEE Workshop on Secure Network Protocols*, pp. 32–37, Santa Barbara, CA, USA, November 2006.

[29] G. Zhang, S. Ehlert, T. Magedanz, and D. Sisalem, "Denial of service attack and prevention on SIP VoIP infrastructures using DNS flooding," in *Proceedings of the 1st International Conference on Principles, Systems and Applications of IP Telecommunications (IPTCOMM)*, pp. 57–66, New York, NY, USA, July 2007.

[30] W. Conner and K. Nahrstedt, "Protecting SIP proxy servers from ringing-based denial-of-service attacks," in *Proceedings of the 10th IEEE International Symposium on Multimedia (ISM)*, pp. 340–347, Berkeley, CA, USA, December 2008.

Hindawi
Scientific Programming
Volume 2021, Article ID 7510641, 7 pages
https://doi.org/10.1155/2021/7510641

Research Article

A Stock Trend Forecast Algorithm Based on Deep Neural Networks

Yingying Yan and Daguang Yang

Business School of Northeast Normal University, Changchun, Jilin 130000, China

Correspondence should be addressed to Daguang Yang; yangdg797@nenu.edu.cn

Received 26 April 2021; Accepted 17 June 2021; Published 5 July 2021

Academic Editor: Shah Nazir

As a recognized complex dynamic system, the stock market has many influencing factors, such as nonstationarity, nonlinearity, high noise, and long memory. It is difficult to explain it simply through mathematical models. Therefore, the analysis and prediction of the stock market have been a very challenging job since long time. Therefore, this paper adopts an encoder-decoder model of attention mechanism, adding attention mechanism from two aspects of feature and time. Both encoder and decoder use LSTM neural network. This method solves two problems in time series prediction; the first problem is that multiple input features have different degrees of influence on the target sequence, the feature attention mechanism is used to deal with this problem, and the weights of different input features can be obtained. A more robust feature association relationship is obtained; the second problem is that the data before and after the sequence have a strong time correlation. The time attention mechanism is used to deal with this problem, and the weights at different time points can be obtained to obtain more robustness and good timing dependencies. The simulation and experimental results show that the introduction of the attention mechanism can obtain lower forecast errors, which proves the effectiveness of the model in dealing with stock forecasting problems.

1. Introduction

The high yield of the stock market has attracted the majority of investors [1], making stock investment [2, 3] become one of the most common ways of investment and financial management. In order to avoid the high risk that comes with high returns, investors are in tireless pursuit of accurate analysis and prediction of the stock market. The amount of trading data generated by daily stock trading is considered to reflect the actual situation of the market, which is often used by investors to analyze and forecast the market. However, because the stock market is affected by many factors such as market supply and demand, exchange rate, company operating conditions, policy changes, and market interest rates, stock prices show the characteristics of fluctuation, which makes the analysis and prediction of stock prices [4–6] face great difficulties.

As a recognized complex dynamic system, the stock market has many influencing factors, such as nonstationarity, nonlinearity, high noise, and long memory. It is difficult to explain it simply through mathematical models. Therefore, the analysis and prediction of the stock market have been a very challenging job since very long time. Random walk theory [7] believes that the fluctuation of stock prices is completely random, and there is no rule to follow. However, many researchers have found a certain rule in the fluctuation of stock prices, which shows that the stock market does have its own specific rules of operation, which lays the foundation for stock price prediction. Traditional stock market forecasting methods include fundamental analysis, technical analysis, multiple regression, and autoregressive moving translation method (ARIMA) [8]. The above methods are usually used for the simulation analysis of stationary or linear time series. The upper limit of the data introduced is small, and the data need preprocessing such as difference to smooth the nonstationary series. Therefore, traditional analysis and forecasting methods have certain discomforts in stock market forecasting [9–12].

The advent of the big data era and the continued advancement of artificial intelligence technology in recent

years [13–16], which has been gradually applied to various fields, brings profound changes to various industries. More and more experts and scholars begin to study the neural network [17–20] and apply it to the prediction of stock and other time series. Neural network has strong self-learning ability, can deal with massive data, and can make more accurate prediction for unstable and nonlinear time series. A long time ago, some researchers used BP neural network to study stock prediction related problems and achieved good prediction results. However, due to the disadvantage that BP neural network does not contain the concept of time sequence when dealing with problems, the prediction effect is still limited. However, the cyclic neural network solves this problem, it has the concept of time series, and it has better performance for stock prediction. However, the cyclic neural network cannot solve the problem of long-term dependence; so on the basis of its structural improvement, LSTM neural network is proposed. LSTM neural network has a great advantage in dealing with time series prediction and further improves the performance of stock prediction. In order to predict the stock price of the next day, the attention mechanism is added into the LSTM neural network and further optimized. It is hoped that we can further improve the handling ability of stock price prediction [21] and obtain better forecast results to meet the needs of current investors.

Following are the main innovations points of this paper:

(1) This paper proposes an encoder-decoder model of attention mechanism and adds attention mechanism from two aspects of feature and time.

(2) This paper uses the time attention mechanism to deal with this problem, which can get the weights of different time points and at the same time obtain the more robust timing dependence.

(3) The simulation experiment results show that the introduction of the attention mechanism can obtain lower prediction errors, which proves the effectiveness of the model in dealing with stock prediction problems.

The organization of the paper is as follows. Section 2 discusses the related work to the proposed research. Methodology of the paper is given in Section 3 with details of the work done in the proposed research. Experiments and results are given in Section 4. The paper is concluded in Section 5.

2. Related Work

Some scholars used neural network technology to study financial time series in the 1990s and predicted the daily rate of return of IBM stocks. However, due to the gradient explosion problem of the traditional BP neural network [22, 23], the result will converge to a local minimum. With the advent of the big data era and the widespread application of deep learning, many scholars have also tried to apply the newly proposed recurrent neural network (RNN) model and its improved model LSTM model in financial research. In particular, LSTM has become the first choice of many researchers due to its unique potential in time series modeling.

Siami and Namin [24] comparatively studied the two financial time series analysis methods, ARIMA and LSTM, and the results show that the LSTM model is 85% more accurate than the ARIMA model. Skehin et al. [25] used wavelet analysis to denoise the stock price data of the five groups listed on Nasdaq: Facebook, Apple, Netflix, Alphabet, and Amazon and established ARIMA models. Compared with the LSTM model, the prediction performance of these two models is compared. Xiong et al. [26] selected Google domestic trends, which can represent macroeconomic factors and public psychology, as input indexes, and used the neural network of long- and short-term memory to study the impact of the above indexes on the volatility of the S & P 500 index from 2004 to 2015. Fischer and Krauss [27] applied the LSTM network to predict the out-of-sample changes of the constituent stocks of the S & P500 index from 1992 to 2015. The prediction results show that the daily return rate of 0.46% can be obtained by using the model, and the annualized Sharpe rate is 5.8.

In addition, Baek and Kim [28] proposed the MOD AUG NET framework, which consists of overfitting prevention LSTM module and prediction LSTM module, to build the model. The validity of the model is evaluated by using two different representative stock market data (S & P500 and KOSPI200). The results show that the test error of MOD-AUGNET C is lower than that of LSTM prediction model alone. Yao et al. [29] used the LSTM network to build a short-term stock price change model and tested it by experimenting with some stocks randomly selected from the 300 Shanghai and Shenzhen stocks. Experiments show that the accuracy, recall rate, and critical error of LSTM are better than random prediction.

The application of the deep learning method to prediction [30] can improve the prediction accuracy of the model, but it does not mean that deep learning method has an inevitable relationship with high accuracy. Compared with traditional methods, this method has many unique advantages in the analysis and prediction of financial time series. However, when facing different practical financial problems, the specific characteristics of different financial problems should be considered to select the application.

3. Methodology

3.1. Attention Mechanism. Attention mechanism was originally used in machine translation and has now become an important method in the field of neural networks. In the field of artificial intelligence, the attention mechanism has become an important part of its structure and has a large number of applications in many fields such as natural language processing, time series prediction, speech, and computer. The attention mechanism is very similar to the human observation mechanism of external things. When we observe external things, we usually only pay attention to a certain part of them first, and everyone has different concerns. For example, when we observe a person, the first thing we notice may be his face, height, body shape, clothing, and

so on. After obtaining the information of each part, you can arrange and combine them and finally get the overall feeling of the person. The attention mechanism generally consists of two parts of tasks, determining which inputs require more attention and extracting features of key parts to obtain important information. Therefore, the attention mechanism can give different weights to each part of the input to filter the main features, which is why the attention mechanism is so widely used.

When processing the input, the calculation of the attention value is the most critical of all operations. Calculating the attention value usually requires two major operations. The first is to use all the inputs to calculate the attention weight and then calculate the weighted average of all inputs based on it.

First, record all the inputs as a whole as X and then select a query variable as q. The function of q is to find and select part of the information in the whole X. Here, a soft attention mechanism is used to select all inputs, give more attention to important inputs, and give less attention to less important inputs.

There are a total of N inputs, q is the query variable, and s is the attention calculation function, and then the attention weight of the i-th input can be expressed as follows:

$$a_i = \text{softmax}\left(s\left(x_i, q\right)\right) = \frac{\exp\left(s\left(x_i, q\right)\right)}{\sum_{j=1}^{N} s\left(x_i, q\right)}. \tag{1}$$

There are many kinds of attention calculation functions. The following four methods are introduced:

$$\begin{aligned} s\left(x_i, q\right) &= v^T \tanh\left(W x_i + U q\right), \\ s\left(x_i, q\right) &= x_i^T q, \\ s\left(x_i, q\right) &= \frac{x_i^T q}{\sqrt{d}}, \\ s\left(x_i, q\right) &= x_i^T W q. \end{aligned} \tag{2}$$

These four methods are called additive method, dot product method, scaled dot product method, and bilinear method where W, U, and v are the network parameters that need to be learned and d is the dimension of the input information.

After getting the attention weight, the attention value can be calculated by weighted average:

$$\text{att}\left(X \Delta q\right) = \sum_{i=1}^{N} a_i x_i. \tag{3}$$

3.2. Encoder-Decoder. The encoder-decoder model is a method to realize the problem of mapping the input sequence to the output sequence. It consists of two parts: an encoder and a decoder. Encoders and decoders can use neural networks such as CNN, RNN, LSTM, and GRU as shown in Figure 1.

The main purpose of the encoder is to extract features from the input time series data. The encoder extracts the information and encodes it into an intermediate vector as the input of the decoder. In the decoder, the intermediate vector is decoded and combined with the input data at the current moment to predict the sequence data at the next moment.

In the encoder-decoder model, the intermediate vector corresponding to each encoder is the same, which means that each feature vector in the input sequence has the same effect on each feature vector in the output sequence. This will cause two problems. For a sequence with a huge amount of information, the intermediate vector cannot fully display it, and the information also has a first-come-first-served relationship; that is, the newly added information will dilute the information before it. This phenomenon is particularly prominent in long time series. This makes the decoder unable to obtain sufficient information from the input sequence when decoding, and the accuracy of the prediction will also be affected.

To solve the above problems, some scholars have proposed to use the attention mechanism to optimize the encoder-decoder model. There is not only one intermediate vector in the model. The increase in intermediate vectors enables the model to optimize the attention mechanism. The selection of the input sequence is realized, and the final output is obtained by learning according to the selection information. The model structure is shown in Figure 2.

First, the input sequence X is passed to the recurrent neural network of the encoder to calculate the hidden state as follows:

$$s_t = f_{\text{RNN}}\left(s_{t-1}, x_t\right). \tag{4}$$

After calculating the hidden state at each moment, they need to be synthesized and stored in the form of an intermediate vector as follows:

$$c = q\left(s_1, s_2, \ldots, s_T\right). \tag{5}$$

The decoding process is actually the inverse operation of the encoding process, mainly using the intermediate vector and the output y_1 at the previous moment to predict y_2 to be output. The recurrent neural network of the encoder first uses the predicted output at the previous time, the hidden state at the previous time, and the intermediate vector to calculate the hidden state at the current time as follows:

$$h_t = f_{\text{RNN}}\left(y_{t-1}, h_{t-1}, c\right). \tag{6}$$

The predicted output at the previous moment, the hidden state at the current moment, and the intermediate vector are passed to a multilayer perceptron g, and the predicted output at the current moment is calculated as follows:

$$y_t = g\left(y_{t-1}, h_t, c\right). \tag{7}$$

3.3. LSTM Based on Attention Mechanism. Time series forecasting mainly faces two major problems. The first is that the target sequence in the time sequence is affected by multiple input feature sequences. Therefore, the influence of multiple input features on the target sequence changes over time, and the degree of influence of different input features is

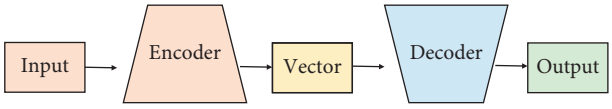


Figure 1: Encoder-decoder model.

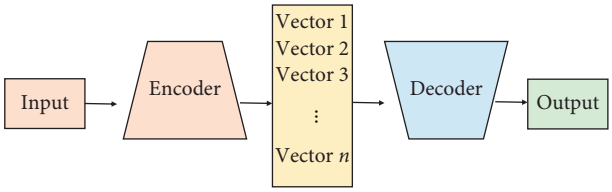


Figure 2: Encoder-decoder model of attention mechanism.

different. The second is the time correlation of the time series. The data before and after each series have a strong time correlation. Within a certain period of time, there is a strong interaction between the series. These time variables have different effects on the target sequence. When using traditional models for analysis, the impact of this interaction will be ignored, resulting in insufficient generalization capabilities of the model.

To solve the abovementioned problems in time series prediction, the attention mechanism can be added to the encoder and decoder parts of the encoder-decoder model. The feature attention mechanism is introduced into the encoder to calculate the attention weight of the input feature at the current moment. This weight indicates the importance of the input feature to the current target task. The sum of the attention weights of all input features is equal to 1, which completes the enhancement of key information and dilutes the information of general importance. The original input features are weighted according to the size of the attention weight, and the intermediate vector representing the updated information is obtained. A time attention mechanism is introduced into the decoder, and a neural network is used to comprehensively process the time sequence of the intermediate vector and input information at the current moment, thereby obtaining the predicted output at the next moment.

4. Experiments and Results

4.1. Hyperparameter Settings. The LSTM neural network model based on attention mechanism proposed in the previous section is referred to as ATT-LSTM model. In the prediction research of stock index in this section, the input characteristics of the LSTM model and ATT-LSTM model are the closing price, opening price, maximum price, minimum price, rising/falling price, and trading volume of the stock index, and the output is the predicted closing price of the next trading day. The number of neurons in the input and output layers is 6 and 1, respectively. In this section, the research on predicting the closing price of individual stocks is also added. In addition to the six features used in predicting the stock index above, the input features of the model are also added: the turnover ratio of individual stocks, volume ratio, price-earnings ratio, price-to-book ratio, price-to-sales ratio, and total market value. The number of neurons in the input layer and output layer is 12 and 1, respectively. In the study of stock index and individual stock, the number of neurons in the hidden layer of both models is 128. ReLU function is used for activation function, MSE is used for loss function, and the Adam algorithm is used for the model learning algorithm. In order to verify the influence of time step on the predicted results, the LSTM model and ATT-LSTM encoder and decoder used LSTM time step which are set to groups for comparative analysis.

4.2. Experimental Results. The forecast results of the closing price of the Shanghai and Shenzhen 300 Index in this paper are fitted into a graph as shown in Figure 3 (the abscissa is time, the ordinate is stock index prices, the blue curve is the predicted value, and the yellow curve is the true value). Observing the curve of the fitted graph and comparing the predicted result with the actual result, we can find the following.

(1) The predicted value of the forecast model for the closing price of the Shanghai and Shenzhen 300 Index is basically consistent with the true value in the overall trend. The predicted value is very close to the actual value, and the degree of deviation is low. However, it can be clearly seen from Figure 3 that during the period of the 2015 stock market crash, although the predicted value of the model is basically the same as the real value in trend, there is a large gap between the predicted value and the actual value. Considering that China's stock market is a developing policy market, the trading mechanism is immature and easily affected by policies, so the model has a certain degree of failure.

(2) There is a certain lag between the predicted value of the model and the actual value, which is manifested as the right deviation of the overall true value curve of the predicted value curve. The predicted value in the rising interval is mostly lower than the true value, and the predicted value in the falling interval is mostly higher than the true value. Since the prediction of the model is usually based on past historical data, there will inevitably be a lag in the prediction results whether the lagging line affects the forecast results and the elimination of lagging still needs to be studied in depth.

In addition, the model trained with 160 stocks of the CSI 300 is recorded as M1, the model trained with 16 stocks of the banking category is recorded as M2, and the model trained with 14 stocks of the securities category is recorded as M3. It can be seen from Figures 4 and 5 that the M1 and M3 models have converged after about 100 iterations. From Figure 6, it can be seen that the M2 model converges after about 80 iterations. In the subsequent iterative training process, the loss function value of the verification set begins to oscillate, indicating that the model has overfitting. Therefore, when testing the model with the test set, the checkpoint model with the minimum loss function is selected.

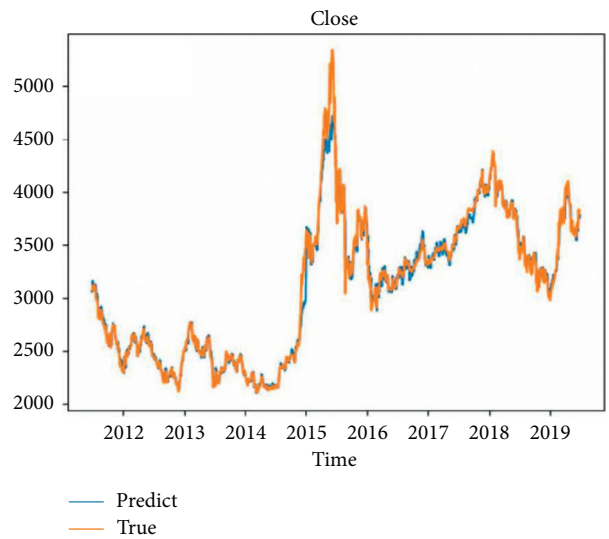


Figure 3: 2011–2019 Shanghai and Shenzhen 300 index forecast result and actual value fitting curve.

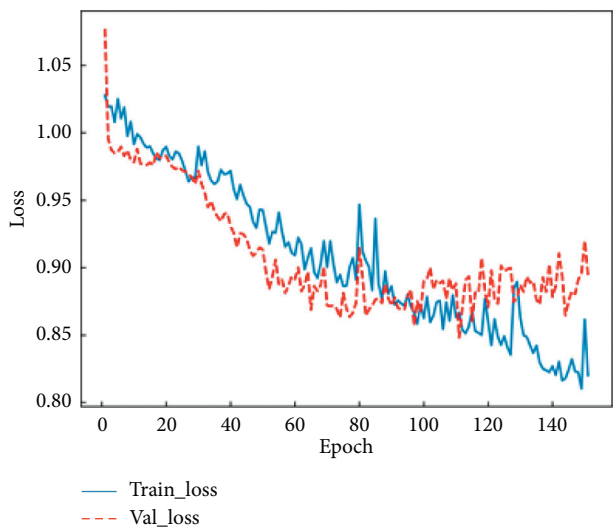


Figure 4: M1 model loss function.

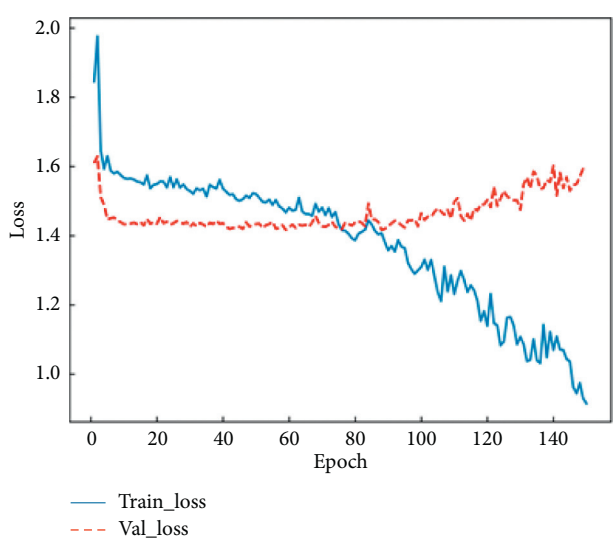


Figure 5: M2 model loss function.

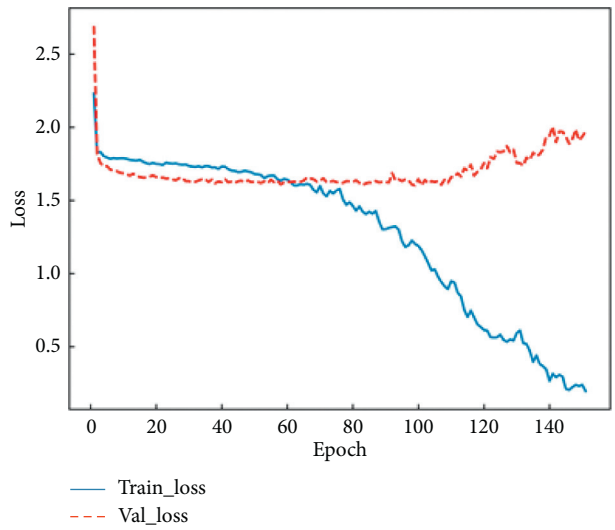


Figure 6: M3 model loss function.

5. Conclusion

In this paper, we adopt an encoder-decoder model of attention mechanism, adding attention mechanism from both features and time. Both encoder and decoder use LSTM neural network. This method solves two problems in time series prediction. The first problem is that multiple input features have different degrees of influence on the target sequence. The feature attention mechanism is used to deal with this problem, and the weights of different input features can be obtained. A more robust feature association relationship is obtained; the second problem is that the data before and after the sequence have a strong time correlation. The time attention mechanism is used to deal with this problem, and the weights at different time points can be obtained to obtain more robustness and good timing dependencies. The simulation experiment results show that the introduction of the attention mechanism can obtain lower forecast errors, which proves the effectiveness of the model in dealing with stock forecasting problems.

Data Availability

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

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

[1] S. Tuysuz, "Conditional correlations between stock index, investment grade yield, high yield and commodities (gold and oil) during stable and crisis periods," *International Journal of Economics and Finance*, vol. 5, no. 9, pp. 28–44, 2013.

[2] Y. Yuniningsih, S. Widodo, and M. B. N. Wajdi, "An analysis of decision making in the stock investment," *Economic Times: Journal of Economic and Islamic Law*, vol. 8, no. 2, pp. 122–128, 2017.

[3] L. Kengatharan and N. Kengatharan, "The influence of behavioral factors in making investment decisions and performance: study on investors of colombo stock exchange, Sri Lanka," *Asian Journal of Finance & Accounting*, vol. 6, no. 1, p. 1, 2014.

[4] L. Zhang, F. Wang, B. Xu, W. Chi, Q. Wang, and T. Sun, "Prediction of stock prices based on LM-BP neural network and the estimation of overfitting point by RDCI," *Neural Computing & Applications*, vol. 30, no. 5, pp. 1425–1444, 2018.

[5] R. Batra and S. M. Daudpota, "Integrating stocktwits with sentiment analysis for better prediction of stock price movement," in *Proceedings of the 2018 International Conference on Computing, Mathematics and Engineering Technologies (iCoMET)*, pp. 1–5, IEEE, Sukkur, Pakistan, March 2018.

[6] S. C. Nayak, B. B. Misra, and H. S. Behera, "ACFLN: artificial chemical functional link network for prediction of stock market index," *Evolving Systems*, vol. 10, no. 4, pp. 567–592, 2019.

[7] S. Rehman, I. U. Chhapra, M. Kashif, and R. Rehan, "Are stock prices a random walk? an empirical evidence of asian stock markets," *ETIKONOMI*, vol. 17, no. 2, pp. 237–252, 2018.

[8] M. Afeef, A. Ihsan, and H. Zada, "Forecasting stock prices through univariate ARIMA modeling," *NUML International Journal of Business & Management*, vol. 13, no. 2, pp. 130–143, 2018.

[9] S. Carta, A. Ferreira, A. S. Podda, R. D. Reforgiato, and A. Sanna, "Multi-DQN: an ensemble of deep Q-learning agents for stock market forecasting," *Expert Systems with Applications*, vol. 164, Article ID 113820, 2021.

[10] R. Efendi, N. Arbaiy, and M. M. Deris, "A new procedure in stock market forecasting based on fuzzy random auto-regression time series model," *Information Sciences*, vol. 441, pp. 113–132, 2018.

[11] A. H. Bukhari, M. A. Z. Raja, M. Sulaiman, S. Islam, M. Shoaib, and P. Kumam, "Fractional neuro-sequential ARFIMA-LSTM for financial market forecasting," *IEEE Access*, vol. 8, Article ID 71326, 2020.

[12] S. P. Chatzis, V. Siakoulis, A. Petropoulos, E. Stavroulakis, and N. Vlachogiannakis, "Forecasting stock market crisis events using deep and statistical machine learning techniques," *Expert Systems with Applications*, vol. 112, pp. 353–371, 2018.

[13] C. Yan, G. Pang, X. Bai, J. Zhou, and L. Gu, "Beyond triplet loss: person re-identification with fine-grained difference-aware pairwise loss," *IEEE Transactions on Multimedia*, 2021, In press.

[14] W. Cai and Z. Wei, "Remote sensing image classification based on a cross-attention mechanism and graph convolution," *IEEE Geoscience and Remote Sensing Letters*, pp. 1–5, 2020, In press.

[15] X. Ning, Y. Wang, W. Tian, L. Liu, and W. Cai, "A biomimetic covering learning method based on principle of homology continuity," *ASP Transactions on Pattern Recognition and Intelligent Systems*, vol. 1, no. 1, pp. 9–16, 2021.

[16] Z. Huang, P. Zhang, R. Liu, and D. Li, "Immature apple detection method based on improved Yolov3," *ASP Transactions on Internet of Things*, vol. 1, no. 1, pp. 9–13, 2021.

[17] X. Ning, K. Gong, W. Li, L. Zhang, X. Bai, and S. Tian, "Feature refinement and filter network for person re-identification," *IEEE Transactions on Circuits and Systems for Video Technology*, 2020, In press.

[18] J. Zhang, Y. Liu, H. Liu, and J. Wang, "Learning local-global multiple correlation filters for robust visual tracking with kalman filter redetection," *Sensors*, vol. 21, no. 4, p. 1129, 2021.

[19] Z. Chu, M. Hu, and X. Chen, "Robotic grasp detection using a novel two-stage approach," *ASP Transactions on Internet of Things*, vol. 1, no. 1, pp. 19–29, 2021.

[20] X. Ning, X. Wang, S Xu et al., "A review of research on co-training," *Concurrency and Computation: Practice and Experience*, 2021, In press.

[21] J. M. T. Wu, Z. Li, N. Herencsar, B. Vo, and J. C. W. Lin, "A graph-based CNN-LSTM stock price prediction algorithm with leading indicators," *Multimedia Systems*, pp. 1–20, 2021, In press.

[22] L. Huang, G. Xie, W. Zhao, Y. Gu, and Y. Huang, "Regional logistics demand forecasting: a BP neural network approach," *Complex & Intelligent Systems*, pp. 1–16, 2021, In press.

[23] G. Zhao, Y. Zhang, Y. Shi, H. Lan, and Q. Yang, "The application of bp neural networks to analysis the national vulnerability," *Computers, Materials & Continua*, vol. 58, no. 2, pp. 421–436, 2019.

[24] N. S. Siami and A. S. Namin, "Forecasting economics and financial time series: ARIMA vs. LSTM," 2018, http://arxiv.org/abs/1803.06386.

[25] T. Skehin, M. Crane, and M. Bezbradica, "Day ahead forecasting of FAANG stocks using ARIMA, LSTM networks and wavelets," in *Proceedings of the 26th AIAI Irish Conference on Artificial Intelligence and Cognitive Science*, Dublin, Ireland, December 2018.

[26] R. Xiong, E. P. Nichols, and Y. Shen, "Deep learning stock volatility with google domestic trends," 2015, http://arxiv.org/abs/1512.04916.

[27] T. Fischer and C. Krauss, "Deep learning with long short-term memory networks for financial market predictions," *European Journal of Operational Research*, vol. 270, no. 2, pp. 654–669, 2018.

[28] Y. Baek and H. Y. Kim, "ModAugNet: a new forecasting framework for stock market index value with an overfitting prevention LSTM module and a prediction LSTM module," *Expert Systems with Applications*, vol. 113, pp. 457–480, 2018.

[29] S. Yao, L. Luo, and H. Peng, "High-frequency stock trend forecast using LSTM model," in *Proceedings of the 2018 13th International Conference on Computer Science & Education (ICCSE)*, pp. 1–4, IEEE, Colombo, Sri Lanka, 2018, August.

[30] W. Sun, P. Zhang, Z. Wang, and D. Li, "Prediction of cardiovascular diseases based on machine learning," *ASP Transactions on Internet of Things*, vol. 1, no. 1, pp. 30–35, 2021.

Hindawi
Scientific Programming
Volume 2021, Article ID 3058141, 8 pages
https://doi.org/10.1155/2021/3058141

Research Article

Aerobics Action Recognition Algorithm Based on Three-Dimensional Convolutional Neural Network and Multilabel Classification

Qian Wang[1] and Mingzhe Wang[2]

[1]*College of Art, Xi'an Physical Education University, Xi'an 710068, China*
[2]*Mechanical and Electrical Department, Hebei Vocational College of Rail Transportation, Shijiazhuang 050000, Hebei, China*

Correspondence should be addressed to Mingzhe Wang; 107057@tea.xaipe.edu.cn

Received 20 April 2021; Accepted 16 June 2021; Published 5 July 2021

Academic Editor: Shah Nazir

In the context of modern people increasingly paying attention to health and promoting aerobics, the amount of data and audiences of aerobics videos has grown rapidly, and its potential application value has attracted widespread attention from scientific research and industry perspectives. This article has integrated computer vision and deep learning related knowledge to realize the intelligent recognition and representation of specific human movements in aerobics video sequences. The study proposes an automatic recognition method for floor exercise videos based on three-dimensional convolutional networks and multilabel classification. Since two-dimensional convolutional neural networks (CNNs) lose time information when extracting features, so to overcome this, the proposed research uses three-dimensional convolutional networks to perform video recognition. The feature is taken in time and space, and the extracted features are subjected to multiple binary classifications to achieve the goal of multilabel classification. Various comparison and simulation experiments are conducted for the proposed research, and the experimental results prove the effectiveness and superiority of the approach.

1. Introduction

With the rapid development of related technologies such as computers [1–3], networks [4–6], and multimedia, multimedia data have shown an exponential growth trend. Video [7] is a common form of multimedia data, and it is also an important part of multimedia data, which is closely related to our daily lives. The video contains the most abundant data information, with a complex structure and a large amount of data. Faced with such a huge amount of video data, automatic video description can better manage and utilize these rich video resources and can help users improve the indexing speed and search quality of online videos so that they can play a greater role. For visually impaired people, through the automatic description of the video combined with text-to-speech technology, the text in the computer is converted into continuous natural language for communication [8]. It can help them better understand the content in the video, thus making the life of the visually impaired more convenient. In the field of video automatic description research, video automatic analysis and understanding based on human actions has gradually become a hot research problem in the field of computer vision and pattern recognition in recent years. It has a wide range of application prospects in the fields of intelligent life assistance, advanced human-computer interaction, and content-based video retrieval [9] and is closely watched by researchers at home and abroad.

Faced with the low-level video features in the current aerobics video analysis research that cannot accurately reflect human high-level semantic concepts, the action recognition algorithm in traditional RGB video has high time complexity and low recognition accuracy, and the use of a single feature cannot meet the massive amount of existing

video data. Growth of complex and other issues, its automatic description research has important theoretical research significance and extensive practical application value. In terms of theoretical research, the research on automatic video description of floor exercise is a cross-cutting subject that integrates machine learning, pattern recognition [10–12], video analysis, computer vision, and cognitive science and provides a good basis for research in these fields. In-depth research can promote the development of related disciplines.

Regarding the problem of automatic identification of aerobics videos [13], which is a difficult point in visual research, in practical applications, the research of automatic identification based on aerobics videos has a wide range of application prospects and potential economic value. In addition to the abovementioned video retrieval and convenience to the visually impaired, potential application areas include sports assisted training, human-computer interaction, and project promotion. First of all, it can satisfy human-computer interaction [14]. In the complex floor exercise set of movements, it is particularly important to identify various human movements quickly. When watching aerobics competitions, commentators often have problems such as delay and error in interpretation of decomposed movements. In this paper, we strive to achieve a higher recognition accuracy in the automatic understanding of human movements based on video and even realize real-time movement recognition and interpretation. For non-professionals, if automatic recognition can be realized, it can not only improve the feeling of watching the game but also make it more convenient for them to understand and learn aerobics.

Secondly, it can assist sports training. In the aerobics video, the movement of the human body is very complex with strong skills. Compared with the daily exercise, the analysis of aerobics video is more difficult and challenging. The analysis of aerobics videos can not only bring more watching effects to sports games but also help coaches to analyze the games and assist athletes in training. Through the research on the automatic understanding of aerobics, while improving the accuracy of aerobics movement recognition, this paper analyzes the movement data so as to excavate the regularity characteristics of gymnastics technology innovation and development and realize the function of auxiliary training. For example, taking related athletes as the main research object, the paper analyzes the differences between the difficulty, arrangement, and quality of the complete sets of movements between the winners and ordinary athletes, studies the development and innovation trend of aerobics, and adjusts the training countermeasures so as to improve the skill level of athletes [15].

Finally, project promotion can be carried out. Taking aerobics as a typical research object, knowledge transfer can be used to effectively identify and locate human movements in aerobics videos. By referring to the method of aerobics movement recognition, it can be applied to other sports so as to expand the research results. Following are the main innovations points of this paper:

(i) To improve the existing algorithm model, the accuracy of automatic aerobics recognition is improved.

(ii) Automatic aerobics video recognition is transformed into a multilabel classification problem. In order to extract the temporal and spatial feature representation in the video, a three-dimensional CNN [16–18] is used as a feature extractor. Then, a two-class classifier for a single decomposition action is constructed, and each video will perform two-class calculations for all categories to complete the multilabel classification process.

(iii) To conduct comparison and ablation experiments, the experimental results prove the effectiveness and superiority of our algorithm.

2. Background

2.1. Convolutional Neural Network. A typical CNN schematic diagram is shown in Figure 1. It consists of three parts: the first part is the input layer, the second part is the several hidden layers, and the third part is the one output layer. Each layer is composed of multiple neural units. CNN [19, 20] has two key ideas, which determine its performance in solving problems related to computer vision field which is particularly outstanding. The first point is that CNN makes use of the two-dimensional structure of images. Since pixels in adjacent areas are usually highly correlated, CNN does not need to establish one-to-one connection between pixel units like traditional neural networks but can directly use grouped local connections. The second point is that the CNN architecture relies on feature sharing, where each channel is generated by convolution using the same filter at all locations.

In the specific CNN network structure, the hidden layer usually includes a convolutional layer, an activation function, a pooling layer, and a fully connected layer. The function of the convolutional layer is to extract the features of the input layer. It is composed of many convolutional units, and the parameters of the convolutional unit are optimized through the backpropagation of the convolutional network. In the process of recognition, the human brain first perceives each feature locally and then comprehensively sorts the local features to obtain global information. Therefore, the feature extraction of the convolutional layer plays a central role in the CNN. A CNN usually contains multiple convolutional layers. The shallow convolutional layer usually can only extract lower-level features. Commonly used CNN usually uses multiple layers in order to obtain deeper feature maps. Convolutional layer is used to iterate. The function of the activation function is to increase the nonlinear segmentation ability of the network. As an activation function, it generally satisfies the properties of nonlinearity, continuous differentiability, monotonicity, best unsaturated range, and approximate linearity at the origin. Commonly used activation functions include ReLU and Maxout. The pooling layer is also called the down-sampling layer, usually after the convolutional layer. The

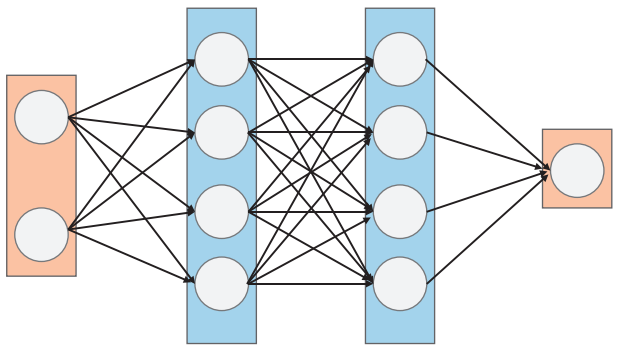


Figure 1: Schematic diagram of a typical neural network.

feature dimensions obtained by the convolutional layer are relatively large, and compressing and sampling the feature maps obtained by the convolutional layer can not only reduce the computational complexity of the network and improve the recognition of features but also avoid the overfitting problem to a certain extent. Common pooling methods include average pooling and maximum pooling. The fully connected layer will connect all the features by weighting, and the output value obtained is used in the calculation of the classifier.

2.2. Recurrent Neural Network. Recurrent neural network (RNN) [21–24] is a special neural network structure inspired by human beings' reliance on past experience and memory in the cognitive process. RNN is called a recurrent neural network. RNN not only gives the input of the previous moment the memory function but also gives the input of the next moment referring to the memory of the previous moment; that is, the current output of a sequence is composed of its input and the previous sequence. The output is jointly determined. The specific process performance will be applied to the previously memorized output when calculating the current output. RNN is different from CNN. In RNN, the input data have a time sequence, thus forming a sequence. This is the most critical point that distinguishes RNN from other neural networks [25–27], and it is also the fundamental reason why the "loop" can be established. The nodes between the hidden layers are unconnected in CNN and become connected in RNN, and the input of the hidden layer includes the output of the input layer and the output of the hidden layer at the previous moment. The hidden layer of the simplest structure of RNN is expanded in time, and its structure is shown in Figure 2. X represents the input sample, 0 represents the output, U and U, respectively, represent the weight of the sample input and output at the moment, t represents the time series, and the memory of the input sample at time t is expressed as follows:

$$S_t = f\left(W * S_{t-1} + U * X_t\right), \tag{1}$$

where W represents the weight entered at the previous moment.

In the actual application process, with the deepening of the network model, the problems of gradient explosion and gradient disappearance appear when the RNN model is

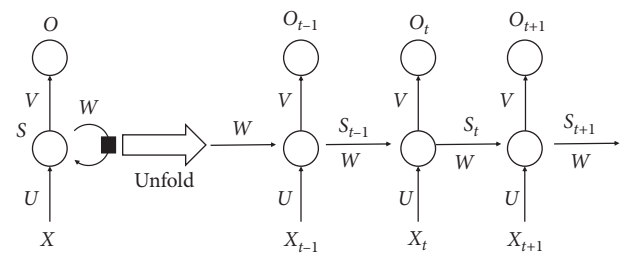


Figure 2: Hierarchical expansion diagram of RNN hidden layer.

trained often trouble researchers. Once the gradient disappears or the gradient explodes, the transfer performance of the training gradient will be greatly reduced, and the original purpose of the RNN model design cannot be achieved. That is, the training gradient cannot be transmitted in a long sequence, which eventually leads to a large deviation in the detection accuracy of the long sequence by the RNN. In order to achieve the long-term dependence problem that needs to be achieved during the training of the RNN model, a long- and short-term memory network is proposed. This network model improves the traditional RNN model by introducing a memory unit and a gate control memory unit. The memory unit can store historical information and the network. In the long-term state, the gate control determines the flow of information through linear intervention, which can selectively increase or decrease the transmission of information. The hidden layer of LSTM [28–31] is different from RNN, and its internal structure is more complicated.

3. Methodology

3.1. Attention Mechanism. The human brain pays attention to different parts of the brain differently when processing signals, known as the visual attention mechanism [32, 33]. Human vision can quickly scan the global image to obtain the target area that needs to be focused on, which is generally known as the focus of attention and then invest more attention resources in this area to obtain more detailed information of the target that needs to be focused on and suppress other useless information. The reason why this paper needs to use the attention mechanism is very intuitive. The decisive video frame for automatic description of the decomposing movements of floor exercises should be the method, direction, and angle of the athlete's body turning, and the weight of these video frames should be greater. This paper uses an attention mechanism which allows the decoder to weight each time feature vector of floor exercise video. Figure 3 shows the network structure after the attention mechanism is introduced.

This paper adopts the dynamic weighted sum of time feature vectors, and the formula is as follows:

$$\phi_t(X) = \sum_{i=1}^{n} \alpha_i^{(t)} x_i, \tag{2}$$

where $\sum_{i=1}^{n} \alpha_i^{(t)} = 1$, $\alpha_i^{(t)}$ is the proportion of the matching score between the hidden layer output and the entire video

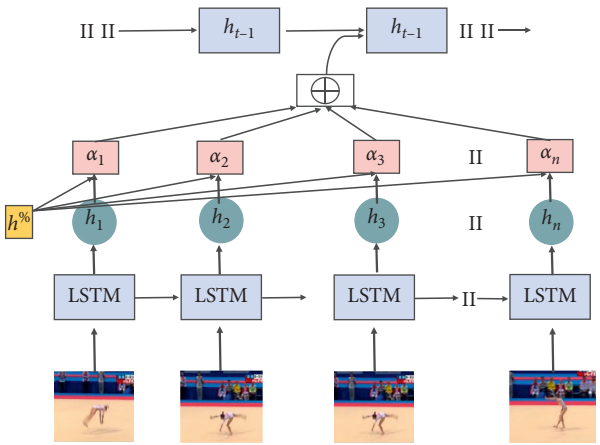


Figure 3: Diagram of attention mechanism.

representation vector at the moment in the overall score, and the calculation formula is as follows:

$$\alpha_i^{(t)} = \frac{\exp\left(\text{score}\left(x_i, h_i\right)\right)}{\sum_{j=1}^{n} \exp\left(\text{score}\left(x_i, h_j\right)\right)}, \tag{3}$$

where $\text{score}(x_i, h_i)$ represents the score value of the output h_i of the i-th hidden layer in the video feature vector x_i. The larger the score, the greater the attention of the input at this moment in the video, and its calculated as follows:

$$\text{score}\left(x_i, h_i\right) = \omega^T \tanh\left(W x_i + U h_i + b\right), \tag{4}$$

where ω, W, and U are the weight vectors and b is the bias.

3.2. Network Framework. In this chapter, the automatic recognition problem of aerobics videos is transformed into a video multilabel classification problem, and finally the classification results are further transformed into real floor exercise recognition. In order to achieve this process, the basic framework used in this section is shown in Figure 4. The framework in the figure can be divided into two parts based on the three-dimensional convolutional network to extract the multilabel video features of aerobics, and then SVM is used to extract multilabel classification that is performed on the pictures, and finally the mapping of the results of multilabel classification to natural language is completed, and the automatic description of the aerobics video is finally completed.

3.2.1. Feature Extraction. Compared with 2D convolutional networks, 3D convolutional networks can better model time information through 3D convolution and 3D pooling operations. In a two-dimensional convolutional network, the process of convolution and pooling is completed in space. In a three-dimensional convolutional network, they perform in time and space. In the introduction of 3D convolutional network above, it was proposed that images should be output when 2D convolutional network is processing images, and images should also be output when multiple images (which are regarded as different channels) are operated. Therefore, the time information of input data will be lost after each convolution operation in the two-dimensional convolutional network. Only three-dimensional convolution can preserve the time information of the input signal and produce the output quantity. The same principle can be applied to 2D pooling and 3D pooling.

3.2.2. Multiclassification of Video Based on SVM. After the video features are obtained, this article will establish a two-class classifier for each decomposition action to determine whether the video contains this type of action. For the establishment of the second-class classifier, the SVM classifier is used in the work of this article. In order to obtain the optimal linear interface of the SVM classifier, the basic idea of solving the problem is to transform the input space into a high-dimensional feature space through nonlinear transformation, which can be regarded as a linear classifier in a broad sense. As shown in Figure 5, the two types of training samples in the figure are represented by "$*$" and "$\diamond$" respectively, x_1 and x_2 represent the two feature items of the sample, and H is the interface of H', H_1, and H_2, respectively, which represent the closest to the interface of the two types of samples. The point is parallel to the plane of the interface. In order to ensure that the empirical risk is minimized in the support vector classification model, not only the optimal dividing line is required to correctly separate the two types of data but also the two types of classification interval (M in the figure) must be maximized. Therefore, although H' is also a boundary that can be classified correctly, it is not suitable as a boundary. The principle of interface selection is to make the support vector machine show better generalization ability.

Use $\{(a_1, c_1), \ldots, (a_N, c_N)\}$ to represent the linearly separable sample set of the two types of problems, where $a_i \in R^d$ and d represents the dimension. The category label $c_i \in \{-1, 1\}, i \in [1, N]$, w is a d-dimensional vector, and b is a constant. From this, the linear discriminant function can be obtained as follows:

$$c(a) = w^T a + b. \tag{5}$$

In order to obtain the maximum classification interval M, the interface needs to meet the following requirements:

$$w^T x + b \begin{cases} > \dfrac{M}{2}, & \text{for } y_i = 1, \\ < \dfrac{M}{2}, & \text{for } y_i = -1. \end{cases} \tag{6}$$

Normalize formula (6) so that all samples can satisfy $|c(a)| \geq 1$, and the sample with the smallest distance from the interface satisfies $|c(a)| = 1$; thus,

$$c_i\left(w^T a_i + b\right) \geq 1. \tag{7}$$

It can be deduced that $M = 2/\|W\|$. When $\|W\|$ is the smallest, the classification interval is the largest. And in order to satisfy that the objective function becomes a

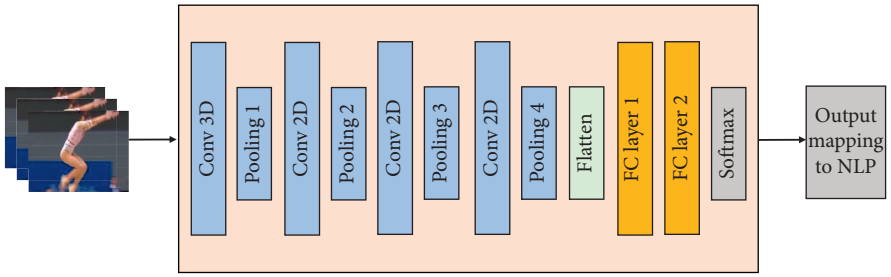


Figure 4: Network framework.

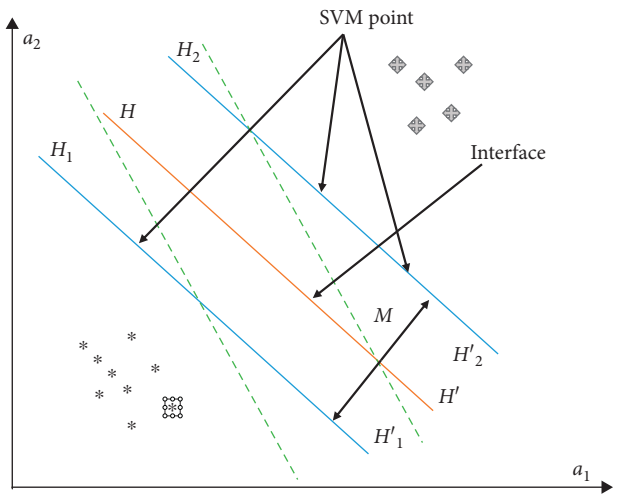


Figure 5: Schematic diagram of SVM interface.

quadratic programming problem, the minimum value of $\|W\|^2$ is taken. Thus, we can get

$$\min_{w,b} \frac{1}{2} w^T w$$
$$\text{s.t.} c_i\left(w^T a_i + b\right) \geq 1. \tag{8}$$

According to the Lagrange method, the corresponding Lagrange function is obtained:

$$L(w, b, a) = \frac{1}{2} w^T w - \sum_{i=1}^{N} a_i\left(c_i\left(w^T a_i + b\right) - 1\right). \tag{9}$$

The previous description is the case of linear separability, and the actual problems that need to be dealt with are often linear inseparable data sets, that is, nonlinear separable problems. Therefore, some misclassifications will inevitably occur in the classification process. One of the solutions is to transform the nonlinearity into a linearly separable problem. Specifically, the kernel function is introduced to map the nonlinearly separable problem in the input space to a higher-dimensional feature space. The Gauss radial basis kernel function is used in this chapter:

$$K(a_1 a_2) = \exp\left(-q\|a_1 - a_2\|^2\right). \tag{10}$$

For N two-classification problems of N decomposition actions, N two-class SVMs are constructed, and each SVM is trained in a one-to-many manner; that is, the video containing the current decomposition action is taken as a positive sample, and other data are taken as a negative sample.

3.2.3. Automatic Recognition Based on Multiple Classifications. In order to form a contrast experiment, this chapter maps the multilabel classification results into automatic recognition statements of aerobics. Different from the previous operation, this step does not involve the processing of video data but only compares the classification results with the test descriptions marked in Chapter 3. Through multiple binary classification SVM classifier, each video will get multiple 0-1 classification results, to identify the category of the assembled into a sentence because most of the video data are only less categories, and the logical relationship between each category is not obvious, so here temporarily do not consider the semantic information, The classification results are directly connected to form the description statement of aerobics movements, and the process of automatic recognition is completed.

4. Experiments and Results

4.1. Experimental Setup. The experiment in this section is carried out on the operating system application Ubuntu 16.04 version. The code to realize the experiment is based on TensorflowL.6.0 framework, and the language used is Python 2.7. The network model was trained on two Nvidia Titan 1080 graphics cards with 11 GB of memory. The input video data are sampled every 5 frames. The input of the C3D feature extraction model is a 16-frame long segment with 8 frames overlapping between two continuous segments. The FC1 activation of these segments is averaged to obtain a 4096-dimensional video descriptor.

4.2. Aerobics Multicategory Data Set. This paper collects a large number of high-standard events of professional athletes, including the Olympic Games, the World Championships, the National Games, and other male and female heavyweight events. First of all, these games are preprocessed. A complete aerobics video is completed by the participation of many athletes. During the video, there will be playback of wonderful moments, slow motion

commentary, judges' ranking, and other links. In the massive videos, athletes are taken as the unit to cut, and only the complete sets of aerobics movements are reserved. 298 videos were included in the training data, and 45 videos were included in the test data.

4.3. Evaluation Index

(1) Accuracy is the most common performance metric in classification models. It is suitable for two-class models and can also be used for multiclass models. The calculation of accuracy is also relatively simple. Assuming that the classification model is g and the test set contains N data in D, the accuracy rate calculation formula is as follows:

$$A = \frac{1}{N} \sum_{i=1}^{N} \left(f\left(a_i\right) = \text{label}_i \right). \tag{11}$$

(2) Bleu is currently the closest indicator to a human score. Bleu adopts the matching principle of N-gram. N-gram represents a sentence as a sequence of n consecutive words. When $N = 1$, the result is Bleu 1.

4.4. Experimental Results. The ultimate goal of this article's multiclassification is to realize the automatic recognition of videos. This article takes the average value of Blue 1 to Blue 4, Blue, as the evaluation index and compares the recognized description of aerobics with the correct description. The experimental results are shown in Table 1. It can be clearly seen that the method used in this article to automatically recognize aerobics videos using the conversion of video multilabel classification has outstanding performance. In addition, Figures 6–8 also show the visual results of the experiment.

4.5. Ablation Study. The experimental results as shown in Table 2 compared from the table are the average values of Bleu 1 to Bleu 4, Blue. You can see three data sets in the table. Two of them are self-built. The two data sets are different when labeling the video description. Ours A is the most direct natural language due to the decomposition of aerobics. The professional requirements of description are very high. Ours B adjusts the description sentence according to professional terminology when describing the mark. From the experimental results of these three data sets, it can be seen that the model with the attention mechanism introduced in this article has better performance regardless of the experimental results on the MSVD data set or the experimental results on the self-built data set. The use of planned sampling in the experiment can also improve the experimental results to a certain extent.

Table 1: Classification accuracy of each category.

Category	Acc
Result 1	0.75
Result 2	0.75
Result 3	0.68
Result 4	0.72
Result 5	0.69
Result 6	0.68
Result 7	0.73
Result 8	0.73
Result 9	0.78
Result 10	0.80

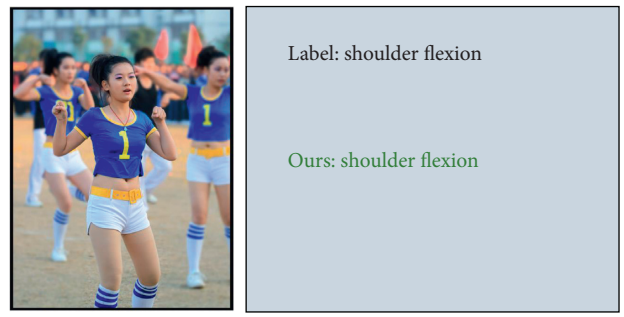


Figure 6: Visualization results of test example 1.

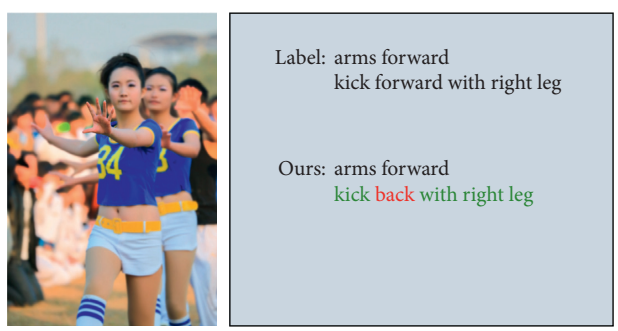


Figure 7: Visualization results of test example 2.

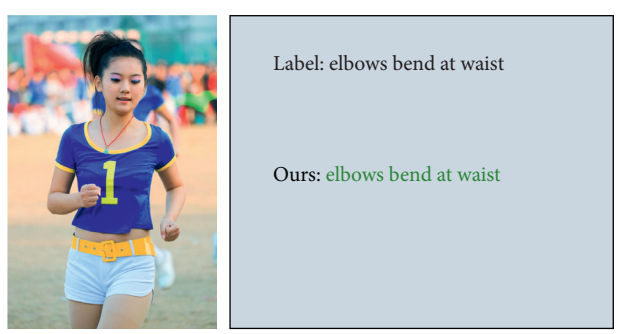


Figure 8: Visualization results of test example 3.

Table 2: Results of ablation research experiments.

Method	S2VT	Attention	Plan to sample
MSVD	17.2	17.9	18.8
Ours A	8.7	10.2	11.6
Ours B	10.9	11.3	13.2

5. Conclusion

In this paper, we will combine computer vision and deep learning related knowledge to realize the intelligent recognition and representation of specific human movements in aerobics video sequences. Therefore, this article proposes an automatic recognition method for floor exercise videos based on three-dimensional convolutional networks and multilabel classification. Since two-dimensional CNN loses time information when extracting features, this paper uses three-dimensional convolutional networks to perform video recognition. The feature is taken in time and space, and the extracted features are subjected to multiple binary classifications to achieve the goal of multilabel classification. We will conduct comparison and simulation experiments, and the experimental results prove the effectiveness and superiority of our algorithm.

Data Availability

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

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

[1] M. A. R. Ahad, *Computer Vision and Action Recognition: A Guide for Image Processing and Computer Vision Community for Action Understanding*, Springer Science & Business Media, Berlin, Germany, 2011.

[2] S. Arseneau and J. R. Cooperstock, "Real-time image segmentation for action recognition," in *Proceedings of IEEE Pacific Rim Conference on Communications, Computers and Signal Processing*, pp. 86–89, Victoria, Canada, Augest 1999.

[3] I. T. Toudjeu and J. R. Tapamo, "Slope pattern spectra for human action recognition," in *Proceeding of the International Conference Image Analysis and Recognition*, pp. 381–389, Springer, Póvoa de Varzim, Portugal, June 2018.

[4] X. Ning, K. Gong, W. Li, L. Zhang, X. Bai, and S. Tian, "Feature refinement and filter network for person re-identification," *IEEE Transactions on Circuits and Systems for Video Technology*, 2020, In press.

[5] W. Cai and Z. Wei, "Remote sensing image classification based on a cross-attention mechanism and graph convolution," *IEEE Geoscience and Remote Sensing Letters*, 2020, In press.

[6] 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.

[7] C. Feichtenhofer, A. Pinz, and R. P. Wildes, "Spatiotemporal multiplier networks for video action recognition," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 4768–4777, Honolulu, HI, USA, July 2017.

[8] S. Pal, P. K. D. Pramanik, T. Majumdar, and P. Choudhury, "A semi-automatic metadata extraction model and method for video-based e-learning contents," *Education and Information Technologies*, vol. 24, no. 6, pp. 3243–3268, 2019.

[9] J. Dong, X. Li, C. Xu et al., "Dual encoding for zero-example video retrieval," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 9346–9355, Long Beach, CA, USA, June 2019.

[10] Y. Zhang, Y. Chen, X. Bai et al., "Adaptive unimodal cost volume filtering for deep stereo matching," in *Proceedings of the AAAI Conference on Artificial Intelligence*, pp. 12926–12934, New York, NY, USA, February 2020.

[11] L. Zhou, X. Bai, X. Liu, J. Zhou, and E. R. Hancock, "Learning binary code for fast nearest subspace search," *Pattern Recognition*, vol. 98, Article ID 107040, 2020.

[12] C. Wang, X. Wang, X. Bai, Y. Liu, and J. Zhou, "Self-supervised deep homography estimation with invertibility constraints," *Pattern Recognition Letters*, vol. 128, pp. 355–360, 2019.

[13] M. Woitas, ""Exercise teaches you the pleasure of discipline"–the female body in jane fonda's aerobics videos," *Historical Social Research/Historische Sozialforschung*, vol. 43, no. 2, pp. 148–164, 2018.

[14] B. K. Chakraborty, D. Sarma, M. K. Bhuyan, and K. F. MacDorman, "Review of constraints on vision-based gesture recognition for human–computer interaction," *IET Computer Vision*, vol. 12, no. 1, pp. 3–15, 2017.

[15] D. Alfermann, M. J. Lee, and S. Würth, "Perceived leadership behavior and motivational climate as antecedents of adolescent athletes' skill development," *Athletic Insight: Online Journal of Sport Psychology*, vol. 7, no. 2, pp. 14–36, 2005.

[16] Z. Chu, M. Hu, and X. Chen, "Robotic grasp detection using a novel two-stage approach," *ASP Transactions on Internet of Things*, vol. 1, no. 1, pp. 19–29, 2021, In press.

[17] C. Yan, G. Pang, X. Bai et al., "Beyond triplet loss: person re-identification with fine-grained difference-aware pairwise loss," *IEEE Transactions on Multimedia*, 2020.

[18] Y. Ding, X. Zhao, Z. Zhang, W. Cai, and N. Yang, "Multiscale graph sample and aggregate network with context-aware learning for hyperspectral image classification," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 14, pp. 4561–4572, 2021.

[19] J. Zhang, Y. Liu, H. Liu, and J. Wang, "Learning local–global multiple correlation filters for robust visual tracking with Kalman filter redetection," *Sensors*, vol. 21, no. 4, p. 1129, 2021.

[20] X. Ning, X. Wang, and S. Xu, "A review of research on co-training. concurrency and computation: practice and experience," *Journal of Healthcare Engineering*, 2021, In press.

[21] L. R. Medsker and L. C. Jain, "Recurrent neural networks," *Design and Applications*, vol. 5, 2001.

[22] X. Guo, H. Zhang, L. Ye, and S. Li, "An approach to learning users' intention to legal consultation with normalized tensor decomposition and BI-LSTM," *Computers, Materials & Continua*, vol. 63, no. 1, pp. 315–336, 2020.

[23] M. Hasnain, S. R. Jeong, M. F. Pasha, and I. Ghani, "Performance anomaly detection in web services: an rnn-based approach using dynamic quality of service features," *Computers, Materials & Continua*, vol. 64, no. 2, pp. 729–752, 2020.

[24] G. Yang, J. Zeng, M. Yang, Y. Wei, and X. Wang, "Ott messages modeling and classification based on recurrent neural networks," *Computers, Materials & Continua*, vol. 63, no. 2, pp. 769–785, 2020.

[25] Y. Tong, L. Yu, S. Li, J. Liu, H. Qin, and W. Li, "Polynomial fitting algorithm based on neural network," *ASP Transactions on Pattern Recognition and Intelligent Systems*, vol. 1, no. 1, pp. 32–39, 2021.

[26] X. Ning, Y. Wang, W. Tian, L. Liu, and W. Cai, "A biomimetic covering learning method based on principle of homology continuity," *ASP Transactions on Pattern Recognition and Intelligent Systems*, vol. 1, no. 1, pp. 9–16, 2021.

[27] Z. Huang, P. Zhang, R. Liu, and D. Li, "Immature apple detection method based on improved Yolov3," *ASP Transactions on Internet of Things*, vol. 1, no. 1, pp. 9–13, 2021.

[28] D. Zhu, Y. Sun, X. Li, and R. Qu, "Massive files prefetching model based on LSTM neural network with cache transaction strategy," *Computers, Materials & Continua*, vol. 63, no. 2, pp. 979–993, 2020.

[29] W. Fang, F. Zhang, Y. Ding, and J. Sheng, "A new sequential image prediction method based on lstm and dcgan," *Computers, Materials & Continua*, vol. 64, no. 1, pp. 217–231, 2020.

[30] B. Yan, X. Tang, J. Wang, Y. Zhou, and G. Zheng, "An improved method for the fitting and prediction of the number of Covid-19 confirmed cases based on LSTM," *Computers, Materials & Continua*, vol. 64, no. 3, pp. 1473–1490, 2020.

[31] Z. Qu, B. Cao, X. Wang et al., "Feedback lstm network based on attention for image description generator," *Computers, Materials & Continua*, vol. 59, no. 2, pp. 575–589, 2019.

[32] W. Cai, B. Liu, Z. Wei, M. Li, and J. Kan, "TARDB-net: triple-attention guided residual dense and BiLSTM networks for hyperspectral image classification," *Multimedia Tools and Applications*, vol. 80, no. 7, pp. 11291–11312, 2021.

[33] R. Liu, X. Ning, W. Cai, and G. Li, "Multiscale dense cross-attention mechanism with covariance pooling for hyperspectral image scene classification," *Mobile Information Systems*, vol. 2021, Article ID 9962057, 15 pages, 2021.

Hindawi
Scientific Programming
Volume 2021, Article ID 3396878, 8 pages
https://doi.org/10.1155/2021/3396878

Research Article

Research on Sports Training Action Recognition Based on Deep Learning

Peng Wang

College of Physical Education, Zhengzhou University, Zhengzhou 450000, China

Correspondence should be addressed to Peng Wang; lyt@zznu.edu.cn

Received 28 April 2021; Revised 31 May 2021; Accepted 9 June 2021; Published 29 June 2021

Academic Editor: Shah Nazir

With the rapid development of science and technology in today's society, various industries are pursuing information digitization and intelligence, and pattern recognition and computer vision are also constantly carrying out technological innovation. Computer vision is to let computers, cameras, and other machines receive information like human beings, analyze and process their semantic information, and make coping strategies. As an important research direction in the field of computer vision, human motion recognition has new solutions with the gradual rise of deep learning. Human motion recognition technology has a high market value, and it has broad application prospects in the fields of intelligent monitoring, motion analysis, human-computer interaction, and medical monitoring. This paper mainly studies the recognition of sports training action based on deep learning algorithm. Experimental work has been carried out in order to show the validity of the proposed research.

1. Introduction

In recent years, human motion recognition has become a hot issue in the field of the application system and academic research. As early as 1973, a psychologist named Johansson carried out the motion perception experiment of moving light spot, which is the first modern research on human motion recognition. Since then, until the 1990s, people began to pay more attention to this field. So far, many researchers around the world have done a lot of research on human motion recognition technology. The traditional research on human motion recognition can be divided into the following parts: representation of motion information and recognition and classification of motion information.

Computer vision is the field of artificial intelligence which is mostly used for creating systems to prepare computer for understanding and eliminating issue with artificial images and sense [1]. Human action recognition is having subtask of collective activity recognition for which the available datasets are commonly inadequate. The study has been presented to look into the issues presenting the collective sports dataset containing multitask recognition for sports and collective activity categories.. A novel protocol of evaluation called unseen sports is presented in which the training and test are carried out on disjoint sets of sports categories [2]. The study proposed human action recognition through deep multimodal feature fusion algorithm [3]. The research fuses visual feature, probability maps, skeleton, and audio signal into hybrid feature utilized for representing human action. Categories of human and nonhuman are classified through the use of convolutional neural network [4]. Research has been done for action recognition of swimming sports based on wireless sensor and field programmable gate array [5].

In recent ten years, deep learning has become a research hotspot in the field of artificial intelligence, and various research results based on deep learning methods have been applied to practice. The sudden boom of deep learning is not accidental, but the reward of decades of intensive work of researchers in this field. From the 1940s to 1960s [6–9], the rudiment of deep learning began to appear in cybernetics. In 1958, Rosenblatt designed neuron perceptron and relized the training of a single neuron. In the 1990s, the emergence of the backpropagation algorithm made it possible to train neural networks with one or two hidden neurons. It was not until 2006 that the concept of deep learning was formally

established by Hinton et al., which set off the third wave of deep learning.

Following are the main contributions of the study:

(i) To study the existing approaches in the context of sports training recognition

(ii) To study the recognition of sports training action based on deep learning algorithm

(iii) Experimental work has been carried out in order to show the validity of the proposed research

2. Human Motion Video Image and Motion Information Representation

2.1. Motion History Image. The motion history image was first proposed by Davis and Bobrick. Before that, they first proposed a binary motion energy image, which is the predecessor of the motion history image. So, let us take a look at the motion energy image. The motion energy image mainly describes how the object moves and space changes, to recognize the moving object. It can describe the outline of the object movement and the spatial distribution of the energy [10–12].

As shown in Figure 1, we take the action of sitting down as an example. The upper line is the keyframe of the action, and the next line shows the binary motion image accumulated from the start frame to the corresponding frame. We can observe that the blank area in the image is the target motion area. By observing the shape of the moving area of the target, the occurrence of the movement and the observation angle is judged [13–15].

We call the accumulated binary motion image the motion energy graph, as shown in the following equation:

$$E_\tau = (x, y, t) = \sum_{i=0}^{\tau-1} D(x, y, t-i), \tag{1}$$

where $E_\tau = (x, y, t)$is the binary motion energy image, $D = (x, y, t)$is the frame difference between the t frame and the $t-1$ frame, and the motion energy image $E = (x, y, t)$is the cumulative sum of the frame differences.

Although a motion energy map can reflect the spatial information of motion, it cannot reflect its temporal information. Therefore, the motion image emerges as the times requirement based on the motion energy image. By calculating the pixel changes in the same position at a certain time, it presents the target motion in the form of image brightness. This method belongs to the template method based on vision. The gray value of each pixel in the motion history image shows the motion of position pixels in the video sequence. If the last moving time of the pixel is closer to the current frame, the higher the gray value. Compared with the motion energy image, it can not only show the sequence of action but also contain more details. Therefore, the motion history image can represent the movement of the human body in a movement process, which makes it widely used in the field of motion recognition. Let s be the intensity value of pixels in the motion history map, and $S_\tau = (x, y, t)$ is the update function.

$$S_\tau(x, y, t) = \begin{cases} \tau, & \text{if } \Psi(x, y, t) = 1, \\ \max\left(0, S_\tau(x, y, t-1) - \delta\right), & \text{otherwise}, \end{cases} \tag{2}$$

where (x, y) represents the position of the pixel and t is the time; t is the duration, which determines the time range of motion from the angle of frame number; 8 is the attenuation parameter. The update function $S(x, y, t)$ can be defined by optical flow, interframe difference, or image difference, and the interframe difference method is the most commonly used. Its application is shown in formulas (3) and (4):

$$\Psi(x, y, t) = \begin{cases} 1, & \text{if } D(x, y, t) \geq \xi, \\ 0, & \text{otherwise}, \end{cases} \tag{3}$$

where

$$D = (x, y, t) = |I(x, y, t) - I(x, y, t \pm \Delta)|, \tag{4}$$

where $I(x, y, t)$ is the intensity value of the coordinate (x, y) pixel in the T frame of the video image sequence, Δ is the interframe distance, and $\xi$is a difference threshold given by a human, which can be adjusted with the change of video scene.

Figure 2 shows the effect pictures of motion history images corresponding to different T values. It can be seen from Figures 2(a) and 2(b) that when the value of T is too small, the whole motion trajectory of the action cannot be obtained completely. As shown in Figure 2(d), when the value of R is too large, the change of the intensity value of the motion track in the captured motion history map is not obvious, which leads to the loss of the information of the action time dimension. We cannot distinguish that the value of t must be considered in the motion history map obtained because the value is too small. As for the difference threshold, if the value is too small, the acquired motion history map will exhibit a lot of messy noise. As shown in Figure 2(e), the obtained image cannot distinguish the foreground from the background well; if the value is too large, the area with a smaller pixel intensity value will disappear, and empty holes will appear, resulting in loss of action information. With the increase of the value, the void area will be larger and larger, until the final motion history image only contains the contour edge. Through the experiment, the optimal value is $t = 50$, $\xi = 40$, which can obtain the most sufficient and effective motion trajectory information.

2.2. Rainbow Coding. The pseudocolor processing of the image can transform the image information into a form that is easier to recognize by humans or machines and enhance the useful information in the image. Pseudocolor processing refers to the technical process of converting a black and white gray image or multiband image into a color tone image. The commonly used pseudocolor coding methods are density segmentation, filtering, and gray level color transformation.

The density segmentation method is mainly used to deal with the image with discontinuous hue, which is the simplest

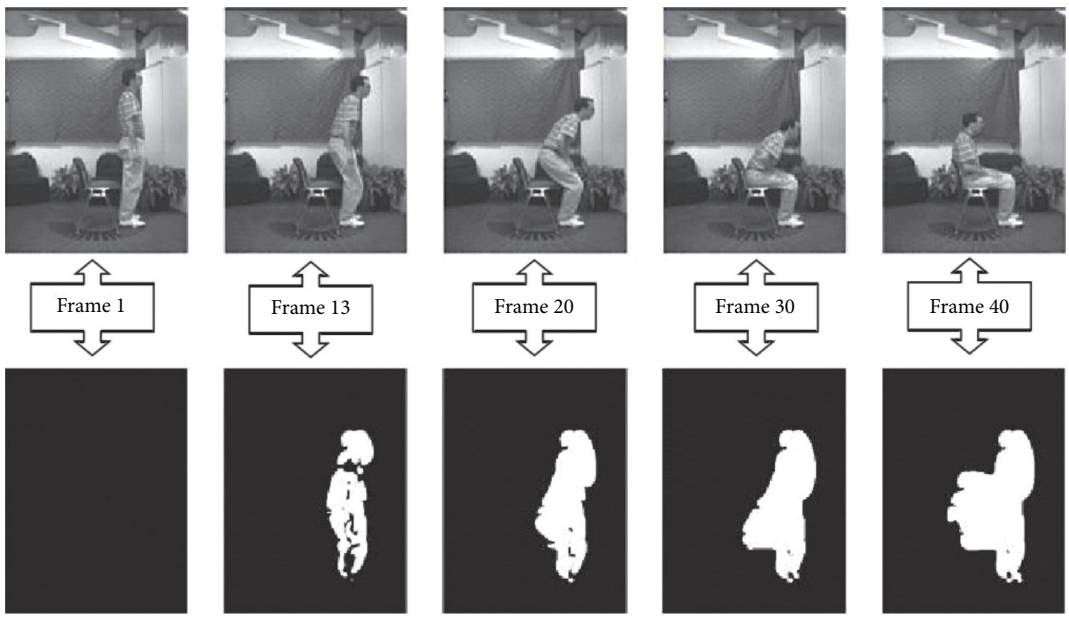


Figure 1: Keyframe and motion energy diagram of action "sit down."

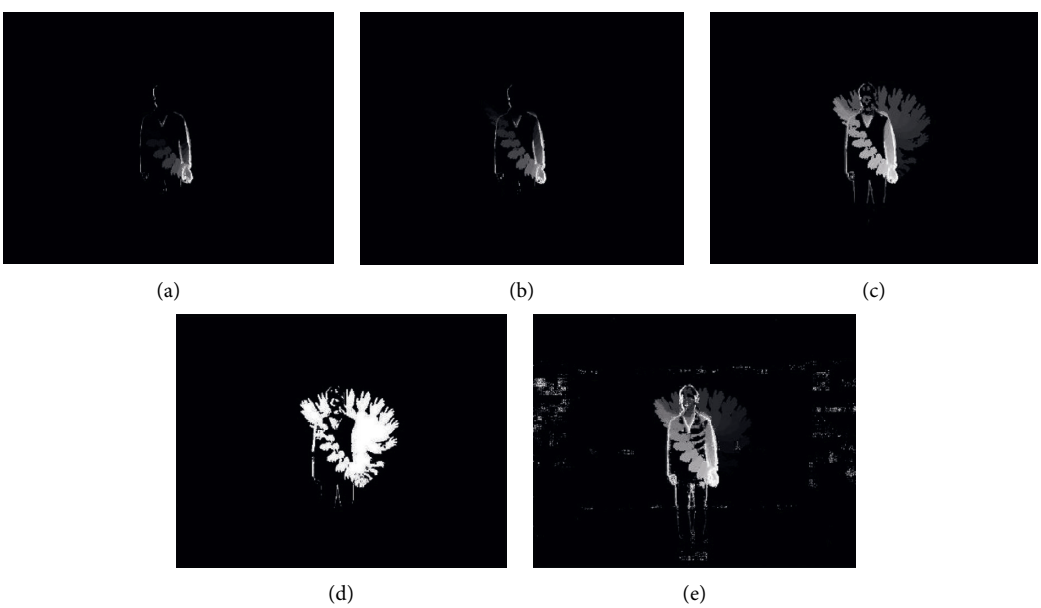


Figure 2: Motion history of different parameters. (a) $\tau = 10$, $\xi = 40$. (b) $\tau = 20$, $\xi = 40$. (c) $\tau = 50$, $\xi = 40$. (d) $\tau = 100$, $\xi = 40$. (e) $\tau = 50$, $\xi = 10$.

method of pseudocolor enhancement. It divides the gray level of a gray image from 0 to 255 into m intervals g, $I = 1, 2, \ldots, M$, and then assigns a specific color C; to each interval, a color image is obtained from a gray image. However, the disadvantage of this method is that the change of hue is not continuous, and the image has obvious blocks, and the number of colors is not rich.

The filtering method is a method based on the frequency domain. It does not rely on the gray level of the image to generate pseudocolor but is determined by the different spatial frequencies of the gray image. As shown in Figure 3, the gray image is first transformed into the frequency domain by Fourier transform, and then it is separated into three independent variables by using three filters with different characteristics in the frequency domain. Then, three single-channel images with different frequency components are obtained by inverse Fourier transform of these three variables. Then, they are processed, such as histogram equalization. Finally, we synthesize our pseudocolor images as RGB tricolor components.

There are many color transformation methods based on gray levels, such as gray mapping, rainbow coding, and so

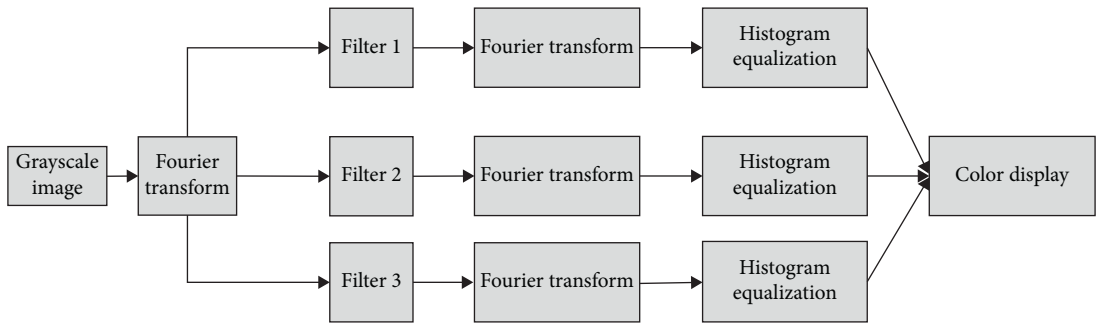


Figure 3: Flowchart of the filtering method.

on. But the central idea is based on the principle of color; according to different coding formulas, the gray value of the image is generated into three-channel values of red, green, and blue, and then the color is synthesized. In RGB color space, any color can be composed of red, green, and blue in different proportions. Therefore, what we need to set is the transformation function of the three color channels. The color matching equation is shown in formulas (5)–(7):

$$R(x, y) = T_R\{f(x, y)\}, \tag{5}$$

$$G(x, y) = T_G\{f(x, y)\}, \tag{6}$$

$$B(x, y) = T_B\{f(x, y)\}, \tag{7}$$

where $R(x, y)$, $G(x, y)$, and $B(x, y)$ are the values of red, green, and blue, respectively, $f(x, y)$ is the gray value of (x, y) points on the gray image, and T_R, T_G, and T_B are the corresponding mapping functions. The pseudocolor image we need can be obtained by driving the color display with the three-channel values. It can be seen that the red, green, and blue mapping functions are very important, which determine the quality of pseudocolor after transformation. Different mapping functions will result in different pseudocolor images.

2.3. Improved Motion History Image. It is not effective to extract motion history images from RGB video and send it to the network for training. In this paper, we propose a human motion recognition method based on the improved motion history image, mainly from the following aspects.

2.3.1. Removing Redundant Motion Sequences. In the experiment, it is found that the performer usually has a reaction time of about 1 second when the action execution command is issued. Similarly, after the execution of the action, there is a period of static time, which means that there are useless still frames at the beginning and end of the dataset video. These frames contain useless redundant information and even cover the important information of keyframes, which directly affects the quality of the extracted motion history and then the image. Therefore, before extracting the motion history image from the video, the first step is to remove 10 frames of each video, and then the motion history image is obtained.

2.3.2. Applying Rainbow Coding. According to Abidi et al. in the report, better perceptual quality and more information can be obtained by encoding gray texture with human perceptible color. Inspired by this, in this paper, we use the rainbow coding to enhance the motion model of the motion image. The larger the gray value, the closer it is to red; otherwise, the smaller the gray value, the closer it is to blue. The motion history image encoded by the rainbow has a rich color. The distribution of color reflects the level of motion and the information of the time dimension, which can more effectively represent the motion information of action.

3. Overview of Deep Learning

3.1. Deep Learning Method. Deep learning is a method of learning data representation in machine learning. Through learning multilevel combination, we can get the recognizable feature representation and finally map the feature representation to the task target. As a kind of machine learning, deep learning is superior to machine learning in that it can automatically learn the feature representation of data. As shown in Figure 4, deep learning avoids the trouble of manual feature design in machine learning; in the traditional machine learning process, data are extracted from input to manual feature extraction, and then the extracted features are mapped to learning objectives. Deep learning simplifies this process. Using the end-to-end idea, the deep learning model can directly convert the input to the output, and the process of feature extraction and feature mapping to the target output is automatically completed by the model, which eliminates many complicated intermediate processes in traditional machine learning.

Like other machine learning methods, the essence of deep learning is to use algorithms to learn knowledge from a large number of data, but it is called "depth." On the one hand, the depth of the deep learning model is the stack of multiple layers of modules, and the number of layers is large. The data from input to target output need multilayer transformation, and the model is deeper; on the other hand, the feature extraction of deep learning is a process of abstraction and fusion from generalization features to semantic features. The shallow features are some basic patterns, the

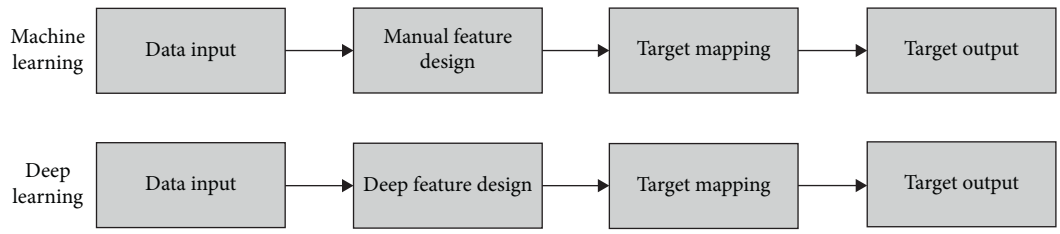


Figure 4: Machine learning method and deep learning method flow.

middle-level features begin to have some fuzzy semantics, and the deep-seated features are recognizable semantic features, which can be mapped to the target output; the deep-seated features have the following features: the obvious progressive process from shallow to deep, and what the model finally learns is this deep feature representation method.

Since the development of deep learning, it includes mathematical analysis, linear algebra, probability theory, mathematical statistics, optimization theory, and numerical calculation. It also includes regularization methods such as random deactivation and batch standardization to ensure the generalization performance of the model, model learning method combining backpropagation and random gradient descent, and distributed representation strategy in the absorption representation learning field. These have injected strong vitality into the deep learning method.

3.1.1. Forward Propagation and Backward Propagation. In deep learning, feedforward neural networks have a forward propagation process. When the input information passes through a layer of hidden units, after layer by layer conversion, the final output is generated. Such a flow of information is called forward propagation; forward propagation can be regarded as a process of network input processing. When the network is training, the input data flow through the network and produce output and calculate the loss function with the target output. Combined with the backpropagation algorithm, the network is updated; when the network weight parameters are fixed, the network training has been completed, and the forward propagation is a prediction process, and the final prediction results are obtained by the forward propagation of the input.

In contrast to forwarding propagation, backpropagation is a process in which the information of cost function flows forward through the network and calculates the gradient. The nonlinearity of the deep learning model makes the learning of the model nonconvex optimization. Generally, the gradient-based method is used to iterate the training model, so that the cost function of the model converges to the minimum value. The backpropagation algorithm points out a way to calculate the gradient. The gradient points out the optimization direction and then combines the stochastic gradient descent algorithm to update the weight parameters of the model. The core idea of backpropagation is to recursively calculate the gradient of the cost function concerning hidden layer output and weight using chain rule. Firstly, the gradient of the cost function about the output of the last hidden layer and the gradient of the output of the last layer concerning the weight parameters are calculated, and then the gradient of the cost function concerning the weight parameters of this layer can be obtained by using the chain rule, and then the gradient of the output of this layer concerning the input is calculated and multiplied.

Based on the gradient of the previous cost function on the output of this layer, the derivative of the cost function concerning the output of the penultimate hidden layer is obtained, and so on, until the hidden layer of the lowest layer, to obtain the gradient of the cost function concerning the weight parameters of all layers.

For the deep learning model with L-layer hidden layer, the weight parameter of each layer is w, and the input of each layer is X. To calculate simply, it is assumed that the output of the former layer is the input of the latter layer, X is the output of the last hidden layer, and the cost function is J.

Calculate $(\partial J/\partial X_{L+1})$ and keep it as the next operation. For each layer $l, l = \{L, L-1, \ldots, 1\}$, the calculation process is as follows:

$$\begin{aligned}
&\text{calculate } \frac{\partial X_{i+1}}{\partial W_i} \text{ and } \frac{\partial J}{\partial X_{i+1}}, \quad \text{then } \frac{\partial J}{\partial W_i} = \frac{\partial J}{\partial X_{i+1}} \frac{\partial X_{i+1}}{\partial W_i}, \\
&\text{calculate } \frac{\partial X_{i+1}}{\partial X_i}, \quad \text{then } \frac{\partial J}{\partial X_i} = \frac{\partial J}{\partial X_{i+1}} \frac{\partial X_{i+1}}{\partial X_i}, \\
&\text{reserve } \frac{\partial J}{\partial X_i} \text{ for the next operation.}
\end{aligned} \tag{8}$$

In the above operation process, the propagation process is simplified. In practice, in addition to the gradient of the weight parameters, if there are bias terms and regular terms, the gradient of the cost function concerning bias and regularity also needs to be calculated. Moreover, before the hidden layer output, there is generally an activation process, and the gradient of the cost function on the layer output needs to be converted to the gradient before the activation.

The backpropagation algorithm is not only used to calculate the gradient of cost function about parameters but also used to calculate the gradient of other outputs on parameters to analyze the model. The backpropagation algorithm can be used to calculate the gradient of any function, which is a very practical method to calculate the gradient.

The backpropagation algorithm combined with random gradient descent has always been the most commonly used learning method for deep learning model law.

3.1.2. Distributed Representation. When it comes to deep learning, we have to mention distributed representation. As a kind of representation learning, deep learning is unique in that it can automatically learn the distributed feature representation of data according to different learning tasks; as an important tool of representation learning, distributed representation is a representation of concepts expressed by a combination of multiple separated features.

In the deep learning model neural network, neuron to the semantic concept is many-to-many mapping. The semantic concept may be represented by activation patterns distributed in different neurons, and a neuron can participate in the representation of different semantic concepts. For example, the semantic concept "cat" can be represented by the combination of features such as "ear," "four legs," and "fur." In convolutional neural networks, these features are the activation modes of several convolutional neurons, and the feature of "fur" can also be a local feature of the semantic concept "dog," "leopard," or "tiger." The neurons that generate this activation mode can also participate in the representation of these semantic concepts; the advantage of this feature is that fewer learning samples can be used to achieve the same as nondistributed tables.

It shows the same learning effect. For example, for input samples such as "white cat," "black cat," "white dog," and "black dog," when distributed representation is not applied, four separate neurons are needed to learn the concepts of color and category at the same time. After using distributed representation, only two kinds of neurons are needed, one is used to describe categories, one is used to describe colors, and the other is to describe colors. Color neurons can learn color concepts from input samples of "cat" and "dog," instead of using specific neurons to learn from specified samples.

4. Analysis and Recognition of Sports Video in the Process of Sports Training

4.1. Adaptive Threshold Moving Object Separation Based on Particle Filter Prediction. The separation of moving objects in sports training videos can collect and process the moving objects from the dynamic background, which is the basis of sports video analysis. Sports training video sequence separation in the sports target is the athlete in the video; the adaptive threshold moving object separation algorithm based on particle filter prediction is used to enhance the accuracy of moving object acquisition. The specific process is as follows: firstly, the foreground image in the video is separated by the three-frame difference method, and the background is projected into adjacent video frames according to the camera steady motion model, and the background separation map of each frame is obtained. The method of background subtraction is used to further separate moving objects. Because of the similarity between the foreground image and the background image, to avoid the separation of the moving objects in the video foreground image mistakenly fused into the background image, it is necessary to separate the coordinate range of the foreground image and obtain the frame threshold of the background image which is not in the coordinate range of the foreground image according to the particle filter method, to complete the adaptive threshold separation of the moving object.

The foreground target obtained by the three-frame difference method is interfered with by noise, which will lead to false separation. After filtering, it can be used as the operation standard of adaptive threshold separation. Then, through the prediction scheme of the particle filter, the foreground coordinate interval of other frames is predicted according to the result of frame separation at this time. The offset state between the image pixel and the foreground coordinate interval is added, which is set as the foreground separation probability of the pixel in other frames. Based on this probability, the adaptive separation threshold can be calculated.

(1) The probability of separating background points by the three-frame difference method is as follows:

$$P_{L,j} = \begin{cases} 0.1, & \text{the pixel belongs to the foreground,} \\ 0.5, & \text{the background of the pixel is not clear,} \\ 0.9, & \text{the pixel belongs to the background.} \end{cases} \tag{9}$$

The probability of pixel separation into the background can be obtained by formula (9). For pixels adjacent to the image boundary, it is not clear that they are the background, so a median value of 0.5 needs to be set. The mean filtering method can restrain the bad invasion of noise. After filtering the collected image pixels with a 3×3 filter, the new background separation probability is obtained as follows:

$$P_f(i, j) = \frac{1}{9} \sum_{m=-1}^{1} \sum_{n=-1}^{1} P(i+m, j+m). \tag{10}$$

(2) The probability of background points is obtained by particle filter. From the schematic diagram of particle filter prediction described in Figure 5, it can be seen that the particle set composed of weighted particles can be regarded as the foreground range of the moving target, and the prediction of foreground range is completed. The particle types in the particle set are vectors (including the x-coordinate and y-coordinate of the upper left corner of the moving object, the x-coordinate and y-coordinate of the lower right-hand corner of the moving object, and the horizontal and vertical movement speed of the upper left-hand corner and the lower right-hand corner of the moving target). After separating the foreground coordinate interval, the actual weights of various types of particles are calculated according to the foreground coordinate interval, and the particle

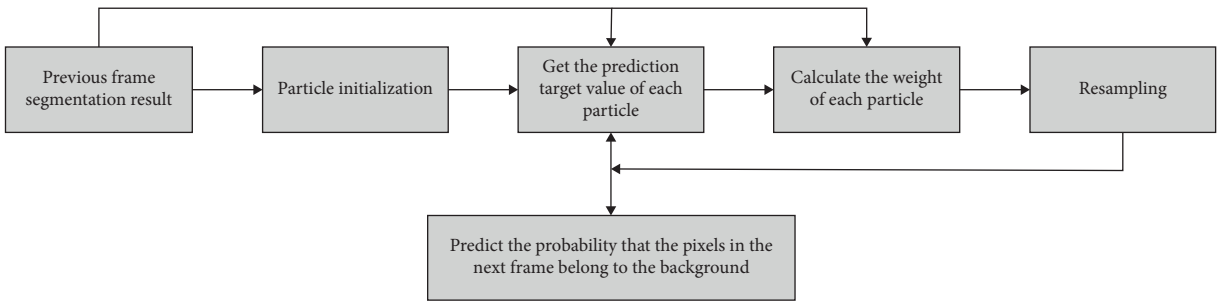


Figure 5: Schematic diagram of particle filter prediction.

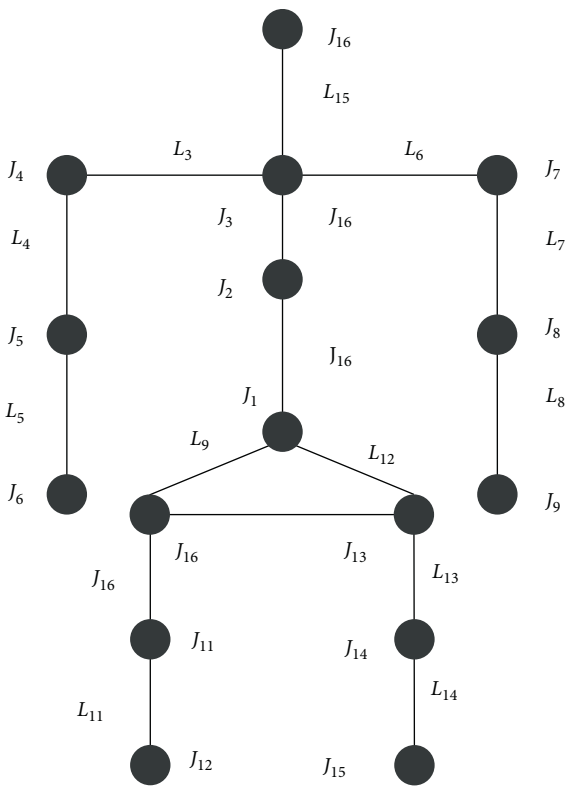


Figure 6: Two-dimensional nodes of human skeleton model.

samples are extracted; the particles with higher weight are more likely to be output as samples. The new particle set is formed by sampling the particles, and then the foreground range of subsequent frames is predicted to obtain the probability of different pixels as background in the subsequent frames. If the moving object pixel in the moving video falls in the background probability area, the point belongs to the background; otherwise, it belongs to the moving foreground, and finally, the moving object separation is completed.

4.2. Moving Object Tracking in Moving Video. The purpose of sports target separation in sports training video is to accurately track the sports target, and the purpose of tracking is to collect the motion parameters of athletes' body joints from the sports training video. The adaptive particle filter algorithm is used to track the sports video in the process of sports training, and the human skeleton model is created, as shown in Figure 6.

The human motion model can predict the immediate motion state according to the motion state of the previous frames. In the process of sports training, the sports trend and the number of sports have regularity.

5. Conclusion

The key to the success of human motion recognition is to capture the spatiotemporal motion patterns of all parts of the human body at the same time. In this paper, a deep learning method based on the motion characteristics of local joints is proposed to recognize the motion samples. Then, the effectiveness of the method is evaluated on two datasets. Aiming at the shortcomings of the model, the average error of the model is reduced and the spatial configuration information is introduced to improve the recognition accuracy of the model. The work of this paper is summarized as follows: after the in-depth investigation on the related tasks of motion recognition, the research prospect and application significance of the deep learning method are pointed out, and the recent research status based on the deep learning method is introduced according to the data mode used, which provides the basis for the later work of action recognition based on local motion features.

Data Availability

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

Conflicts of Interest

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

Acknowledgments

This study was supported by the Outstanding Young Teachers Project in Colleges and Universities of Henan Province (no. 2017GGJS164).

References

[1] X. Xie, "Image recognition of sports training based on open IoT and embedded wearable devices," *Microprocessors and Microsystems*, vol. 82, Article ID 103914, 2021.

[2] C. Zalluhoglu and N. Ikizler-Cinbis, "Collective sports: A multi-task dataset for collective activity recognition," *Image and Vision Computing*, vol. 94, Article ID 103870, 2020.

[3] E. Zhou and H. Zhang, "Human action recognition toward massive-scale sport sceneries based on deep multi-model feature fusion," *Signal Processing: Image Communication*, vol. 84, Article ID 115802, 2020.

[4] W. Yang, Y. Gao, and F. Zhai, "Simulation of sports action picture recognition based on FPGA and convolutional neural network," *Microprocessors and Microsystems*, vol. 80, Article ID 103593, 2021.

[5] T. Liang, "Swimming sports action recognition based on wireless sensor and FPGA," *Microprocessors and Microsystems*, Article ID 103433, 2020, In press.

[6] C. Gong, H. Wang, R. Li et al., "Video moving object trajectory tracking algorithm based on state-dependent detection," *Modern Electronic Technology*, vol. 39, no. 7, pp. 51–56, 2016.

[7] H. Song, F. Wang, X. Liu et al., "Group motion analysis and abnormal behavior detection in a geographical environment," *Geography and Geographic Information Science*, vol. 31, no. 4, pp. 1–5, 2015.

[8] M. Zhu, "Research and application of a three-axis acceleration sensor system in the field of sports technology analysis," *Ice and Snow Sports*, vol. 37, no. 2, pp. 89–96, 2015.

[9] F. Wei, "Loose design of key parts of children's trousers under sports condition," *Xi'an University of Technology Acta Sinica*, vol. 29, no. 5, pp. 550–554, 2015.

[10] H. Guo, "Table tennis decision-making system based on dual-channel target motion detection," *Journal of Guangzhou Institute of Physical Education*, vol. 34, no. 6, pp. 67–70, 2014.

[11] D. Sui and D. Hou, "Research and simulation of dynamic image tracking optimization algorithm," *Modern Electronic Technology*, vol. 39, no. 6, pp. 98–100, 2016.

[12] B. Li, F. Liu, D. Li et al., "Abnormal behavior detection and recognition of moving objects in campus intelligent video surveillance," *Software Guide*, vol. 15, no. 2, pp. 134–136, 2016.

[13] M. Wu and L. Lin, "Research on video recognition and comparison system of competitive sports," *Journal of Guangzhou Institute of Physical Education*, vol. 34, no. 4, pp. 59–61, 2014.

[14] C. Li and X. Lian, "Research on video analysis and intelligent diagnosis system of national traditional sports antagonism," *Journal of Guangzhou Institute of Physical Education*, vol. 36, no. 1, pp. 52–56, 2016.

[15] W. Wei, K. Wu, L. Guo et al., "Video tracking of gradient expansion template motion of memory watershed disc small," *Journal of System Simulation*, vol. 28, no. 2, pp. 462–466, 2016.

Hindawi
Scientific Programming
Volume 2021, Article ID 9292496, 11 pages
https://doi.org/10.1155/2021/9292496

Research Article

Determining the Degree of Characteristics for Internet of Healthcare Devices Using Fuzzy ANP

Lin Wen

Department of Medical Equipment, Sichuan Provincial People's Hospital, University of Electronic Science and Technology of China, Chengdu, China

Correspondence should be addressed to Lin Wen; linwen044@gmail.com

Received 30 April 2021; Accepted 9 June 2021; Published 29 June 2021

Academic Editor: Shah Nazir

With the revolution in Internet and digital technology, every organization is adopting digital things to carry out their day-to-day activities. Internet of Things (IoT) is a concept that is used to connect various devices over the Internet to increase the production and quality of service, deliver a huge amount of data in seconds, and automate the processes. IoT implementation in the health sector has changed the typical setup into smart and intelligent setup. With the "smart" and "Intelligent" abilities of IoT devices such as sensors and with the collaboration of humans and computers, physical processes can be monitored and, based on the received data, optimal decisions can ultimately be taken. IoT applications in healthcare will increase flexibility, patient's care, quality of health, and control of diseases. As the IoT combines various heterogeneous devices and the inappropriate determination of the degree of characteristics for Internet of healthcare (IoH) devices may affect the efficiency of services, this research is carried out with the use of decision support system and application of fuzzy analytic network process (fuzzy ANP) was used for optimal determination of the degree of characteristics for IoH devices based on unique properties which would greatly increase the efficiency of industry. The experimental results are efficient and show the usefulness of the approach.

1. Introduction

The evolution of information technology to the IoT has demonstrated that a highly intelligent ecosystem can be built by connecting an information network autonomously via routers from OS built gadgets and computers [1]. However, there are some security issues with IoT, such as understanding the transition from offline to online infrastructure, handling temporal aspects of security, resolving the implementation gap for best practice, and dealing with development complexity [2]. Industrial processes have been turned into completely digitized smart processes as a result of Industry 4.0, which involves IoT, Industrial Internet, smart manufacturing, and cloud-based manufacturing and is still a futuristic yet practical idea. It has placed strict human constraints in the manufacturing process in order to achieve quality improvement and an emphasis on value-added activities while eliminating waste [3]. Due to heterogeneous devices, multiple channel impairments, attenuation, spontaneous variations, and shadowing impacts occur in industrial scenarios where devices interact dynamically and in a stable and efficient manner with one another using wireless technology [4]. Flexibility, mass customization, service efficiency, and productivity are all growing as a result of Industry 4.0 [5]. To control physical processes and make intelligent choices, modern industrial production systems depend on real-time cooperation and consultation with humans, machines, sensors, and other systems. Resources are transformed into intelligent objects so that they can feel, function, and behave in a smart world, which is an important part of Industry 4.0 [6]. Industrial Internet of Things (IIoT) has revolutionized the digital industry by allowing for product accountability, integrity control, and delivery cost [7].

IoT solutions lead to the formation of innovative and efficient innovations in today's manufacturing environments, with the aim of increasing operational efficiency in growing fast responding factories [8]. Flexible

manufacturing systems allow all devices to communicate via the Internet of Things (IIoT). Automotive automakers' willingness to adapt processes to achieve ultimate versatility would help them maintain efficiency in an increasingly global marketplace [9]. During the production process, numerous devices capture signal and upload it to a cloud server via the Internet. Since a cloud server cannot be fully trusted, data authenticity should be maintained before data is outsourced to the cloud server, ensuring that only approved users or devices have access to the authentic data from various topography areas [10]. The IoT combines various heterogeneous devices that may produce certain risks to data such as data loss, decreased throughput, time delays, latency problem, and security and privacy issues.

Various research studies have been carried out on IoT and IIoT which addressed different areas of research including smart communication, intelligent factories, and Industry 4.0. Research works have been done in the area and some researchers suggested approaches for enhancing the efficiency and effectiveness of industry with IoT applications. However, no study revealed the problem of determining the degree of characteristics of IoH devices in the IoT setup. To overcome this issue, the proposed research is carried out as a decision support system and the applications of fuzzy ANP were used for optimum determination of the degree of characteristics for IoH devices based on exclusive properties which would significantly increase the efficiency of industry.

The remainder of this paper is organized as follows: work related to the area of the research is given in Section 2. The methodology of the paper with subheadings is briefly described in Section 3. Section 4 discusses the results and discussion of the paper. The conclusion of the paper is presented in Section 5.

2. Related Work

Researchers are trying to come across approaches, tools, and solutions for effective IoT in industry. Various approaches have been devised. The asynchronous altering direction method of multipliers algorithm was used to solve the global consensus cross-layer optimization problem that is associated with the control of industrial health status [11]. For observing the issue of a single architecture that complies with Industry 4.0 standards as they apply to businesses, the research provided a fundamental idea for free, safe, stable, near real-time, standardised communication interfaces [12]. DDoS mitigation employs the fog computing philosophy. Real-time traffic filtering through field firewall tools is used in the mitigation process, and data from decentralized dedicated servers is correlated for accurate decision-making [13]. In order to break the deadlock in Industry 4.0, a job shop scheduling method is adopted, and a new assessment feature is created, using the Asset Administrative Shell (AAS) [14]. Logistics 4.0 is the result of the IIoT revolutionizing the logistic domains by offering accountability (supply chain visibility), integrity management, and commodity delivery cost. In the scope of Logistics 4.0, some insights on the necessary criteria and concerns for organizations to be effective and fully operational are highlighted [7]. Flexibility, mass customization, consistency, and efficiency are all gaining ground as a result of Industry 4.0. The importance of maintenance in Industry 4.0, as well as the potential for new quality control and production development, is highlighted. Present maintenance measures are trending toward "remote maintenance" and the appealing prospect of "self-maintenance." It has been highlighted that the human element is also a significant factor in achieving various goals such as flexibility and efficiency [5].

Several channel impairments, attenuation, random fluctuations, and shadowing impacts, as well as non-line-of-sight, occur in industrial scenarios where devices interact freely and in a stable and efficient manner with one another using wireless technology. For channel model evaluation, performance estimation techniques such as acquired signal power, packet loss likelihood, RMS (Root Mean Square) error, and RMS delay spread are examined. This research may be used to satisfy the requirements for adequate IIoT channel modelling [4]. Industry 4.0 places a high value on industrial automation. Typical tools are transformed into intelligent objects capable of sensing, acting, and behaving in a smart world. Smart factory and cloud manufacturing are all terms used to describe different types of smart manufacturing systems. Intelligent production is enabled by cyber-physical systems, cloud computing, big data analytics, and information and communications technology (ICT) [6]. The NICE Nanogrid was used to build laboratory-based IIoT-driven software and controls framework for the Shenhua community as part of a NICE smart device initiative. The aim is to create an open architecture-based Industrial Smart Energy Consortium to bring together industrial partners, academic institutions, module suppliers, equipment vendors, and other interested parties to investigate and apply to a test-bed-centric open laboratory template and platform for next-generation energy-oriented smart industry applications [1]. The rapid hardware proliferation of a complicated smart sensor communication system was accomplished using a high-level system design. The GRAFCET is being used to design the system's structural actions as well as internal functional units. A sequential-concurrent hybrid discrete event scheme is used to describe the actions of every unit. For a smart grid sensor and an intelligent gateway module, high-level synthesis rules were used to generate VHSIC hardware description language- (VHDL-) target efficient hardware. Ultimately, to incorporate all smart system components into a complicated embedded system, these embedded systems controllers are generated automatically, and a hardware circuitry is synthesized [15].

The biogeography-based optimization and the Markov decision process are combined in a hybrid model. The MDP is used to determine the likelihood of selecting a radio access tool, which serves as an input to the BBO process. As a result, in a heterogeneous network, the BBO uses the described multipoint algorithm to decide the best RAT [16]. The security of the IoT layered architecture, as well as the standards used for communication, networking, and maintenance, was investigated. Security specifications for IoT were established, as well as current threats, risks, and state-of-the-art

Table 1: Existing approaches in the area.

No.	Title	Year	Ref.
1	Maintenance Transformation through Industry 4.0 Technologies: A Systematic Literature Review	2020	[5]
2	Attacks on the Industrial Internet of Things–Development of a Multilayer Taxonomy	2020	[25]
3	Comparative Link-Level Analysis and Performance Estimation of Channel Models for IIoT (Industrial-IoT) Wireless Communications	2020	[4]
4	Research on the Optimization of IIoT Data Processing Latency	2020	[18]
5	Deep Robust Cramer Shoup Delay Optimized Fully Homomorphic for IIOT Secured Transmission in Cloud Computing	2020	[10]
6	Intelligent Equipment Design Assisted by Cognitive Internet of Things and Industrial Big Data	2020	[26]
7	Big Data and Disruptive Computing Platforms Braced Internet of Things: Facets and Trends	2020	[27]
8	Industrial Internet of Things Based Ransomware Detection Using Stacked Variational Neural Network	2019	[22]
9	The Need for Deterministic Virtualization in the Industrial Internet of Things	2019	[23]
10	A Fog Computing Based Approach to DDoS Mitigation in IIoT Systems	2019	[13]
11	A Hybrid Intelligent Model for Network Selection in the Industrial Internet of Things	2019	[16]
12	Towards Distributed IoT/Cloud-Based Fault Detection and Maintenance in Industrial Automation	2019	[11]
13	Flexible Manufacturing Systems Using IIoT in the Automotive Sector	2019	[9]
14	Internet of Things (IoT) and Big Data Analytics in Healthcare	2019	[28]
15	An Efficient Intrusion Detection Model for Edge System in Brownfield Industrial Internet of Things	2019	[20]
16	A Microservice Architecture for the Industrial Internet-Of-Things	2018	[21]
17	Communication Systems for Industry 4.0 and the IIoT	2018	[12]
18	Implementing Industry 4.0 in Discrete Manufacturing: Options and Drawbacks	2018	[14]
19	Enabling Industrial Internet of Things (IIoT) towards an Emerging Smart Energy System	2018	[1]
20	Industry 4.0 – A Glimpse	2018	[3]
21	Parallel Computing and Network Analytics for Fast Industrial Internet-of-Things (IIoT) Machine Information Processing and Condition Monitoring	2018	[24]
22	Avoiding the Internet of Insecure Industrial Things	2018	[2]
23	IoT Security: Review, Blockchain Solutions, and Open Challenges	2018	[17]
24	Industry 4.0 Implications in Logistics: An Overview	2017	[7]
25	Intelligent Manufacturing in the Context of Industry 4.0: A Review	2017	[6]
26	High-Level Modeling and Synthesis of Smart Sensor Networks for Industrial Internet of Things	2017	[15]
27	A Large-Scale Web QoS Prediction Scheme for the Industrial Internet of Things Based on a Kernel Machine Learning Algorithm	2016	[19]
28	Industrial Internet of Things Monitoring Solution for Advanced Predictive Maintenance Applications	2016	[8]

solutions. It was discovered that blockchain could be a major driver in resolving many IoT security issues [17]. To address high latency when processing industrial data on cloud, an Industrial Internet of Things cloud-fog hybrid network architecture is proposed. Framework uses edge devices like routers and switches to build a fog computing layer between the cloud server and the production equipment in the production equipment region. A distributed computing approach for multiple devices is proposed because the computing power of the edge devices in fog computing is very low. The constrained particle swarm optimization load balancing algorithm based on the simulated annealing approach is used to achieve the goal of minimizing task processing delays (SAPSO-LB) [18]. The kernel least mean square (KLMS) algorithm is used to suggest a data-driven scheme for predicting missing QoS values for the IIoT. For each known QoS entry, the Pearson correlation coefficient (PCC) is first applied during the data prediction process to find related QoS values from similar service users and web service objects. The secret relationships between all known QoS data and corresponding QoS data with the highest similarities are then investigated using KLMS [19]. The NGS-PlantOne system is a cutting-edge Industrial IoT (IIoT) solution that allows ubiquitous tracking of machinery and equipment via rechargeable IoT sensing devices, allowing for the creation of advanced planned maintenance applications in the IIoT scenario. The generated IIoT solution's reliability results are promising [8]. Flexible manufacturing systems (FMS) allow all systems to function via IIoT. The willingness of automakers to adapt processes in order to achieve ultimate versatility would help them maintain profitability in an increasingly global marketplace. The key elements to effective implementation of smart manufacturing for the automotive sector were investigated, as well as flexible manufacturing systems that use IIoT [9].

For privacy protection, Robust Cramer Shoup Delay Optimized Fully Homomorphic (RCS-DOFH) is proposed. To decode privately induced signals and their weighted variables, the RCSD protocol is used. The encoding is then carried out using DOFHE method. After that, the encoding territory's activation role is calculated [10]. A software architecture is proposed which incorporates concepts from the IoT environment, various industrial automation domains, and modern IT and cloud architectures. Application patterns are used to gain a better understanding of the problem domain and to create a consistent software architecture throughout the development of the software architecture. The software architecture's resulting lightweight and modular nature, combined with support for modern development approaches (containerization, CI, CD), makes it ideal for deployment on cloud, fog, and edge devices in industrial and nonindustrial IoT environments [20]. In the industrial

automation market, the Internet of Things (IoT) and cyber-physical system (CPS) principles have been introduced. Due to tighter dependability and real-time constraints, the IoT concept cannot be completely adopted in IIoT environments [21]. A detection model based on a fully connected neural network and a stacked Variational Autoencoder (VAE) was proposed which can learn the latent structure of device activities and reveal ransomware actions [22]. Real-time virtualization is widely acknowledged as a primary enabler of fog computing and the IIoT. There are some requirements that any hypervisor that claims to be a deterministic virtualization solution for IIoT must meet. Current studies on real-time virtualization were examined to explain the trade-off between versatility and deterministic execution. Hypervisors that fulfil all of the criteria for deterministic virtualization have been defined as a requirement [23].

3. Methodology

The following sections briefly show the methodology of the paper.

3.1. Existing Approaches in the Area. Industry 4.0, which includes Internet of Things, Industrial Internet, Smart Manufacturing, and Cloud-based Production, is still a futuristic yet practical term. Industry 4.0 refers to the strict incorporation of humans into the production process in order to achieve quality improvement, an emphasis on value-added operations, and the avoidance of waste. An overview of Industry 4.0 was provided, as well as an understanding of the nine pillars of Industry 4.0 and their implementations, in addition to identifying the problems and issues that arise when implementing Industry 4.0 [3]. A new approach to large-scale IIoT system information processing, network modelling, fault detection, and fault diagnosis has been suggested. A dynamic warping algorithm was first implemented to describe the distinction of system signatures (e.g., power profiles during operations). Second, a stochastic network embedding algorithm was created for constructing a large-scale network of IIoT machines, with the dissimilarity between machine signatures maintained in the network node-to-node distance. A parallel computing method was designed for effective system modelling of large-scale IIoT-enabled devices, using the power of multiple processors. According to experimental results, the proposed technique efficiently and effectively characterizes signature differences in both cycle-to-cycle and machine-to-machine dimensions. [24]. A multilayer taxonomy was used to distinguish similarities and differences between attacks on the IIoT by incorporating IT protection, IIoT, and risk mitigation into a multidisciplinary approach. The study also encourages organizations and their value-creation networks to identify, record, and share events [25]. Admiring the transition from offline to online infrastructure, handling temporal aspects of protection, resolving the identified gaps for best practice, and dealing with infrastructural complexity are all security issues posed by IoT. Noncompliance risks can arise as a result of smart energy systems. The study focuses on how to escape these dangers [2]. Table 1 shows some of the existing approaches proposed by the researchers in the field.

Table 2: Characteristics and features of IoH devices.

Characteristics of IoH devices	References
Efficiency	[29–34]
Reliability	[29, 33, 35, 36]
Security	[29–31, 36–38]
Privacy	[29, 35, 37, 39]
Real time	[29, 30, 40]
Intelligent	[30, 32, 34, 35, 41]
High sensing	[30, 32, 36]
Availability	[29, 31, 41]
Interoperability	[29, 42]

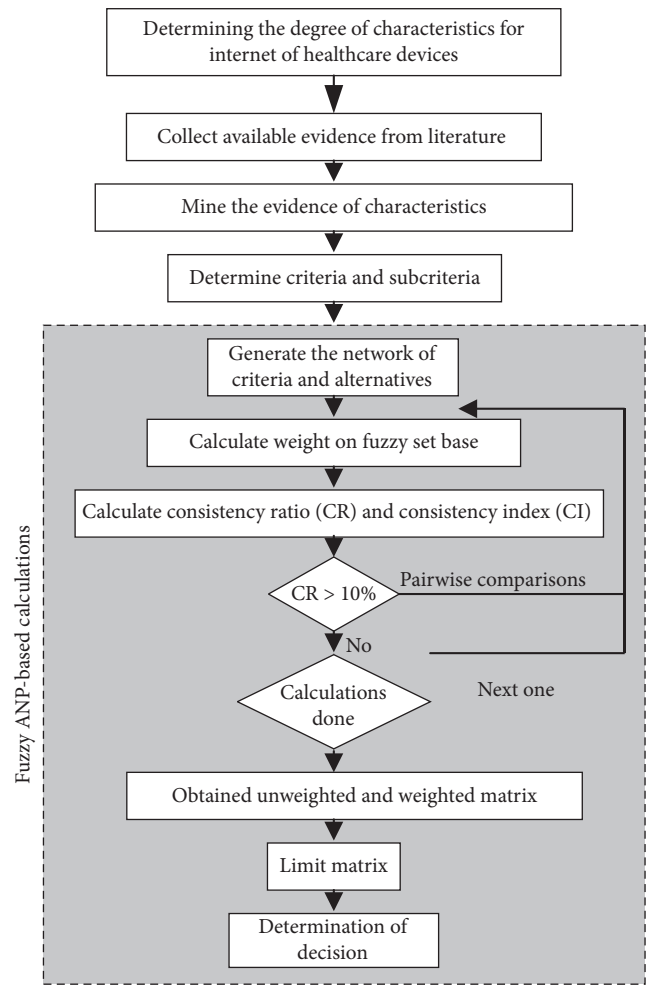


Figure 1: Process of the proposed research for determination of characteristics of IoH devices.

3.2. Defining the Characteristics of IoH Devices. The existing literature in the area of research was studied for identifying various features and characteristics of the IoH for determination of degree of characteristics of devices. More features were identified, but these features' analysis was quite tricky, so the features were then reduced and the significant features were

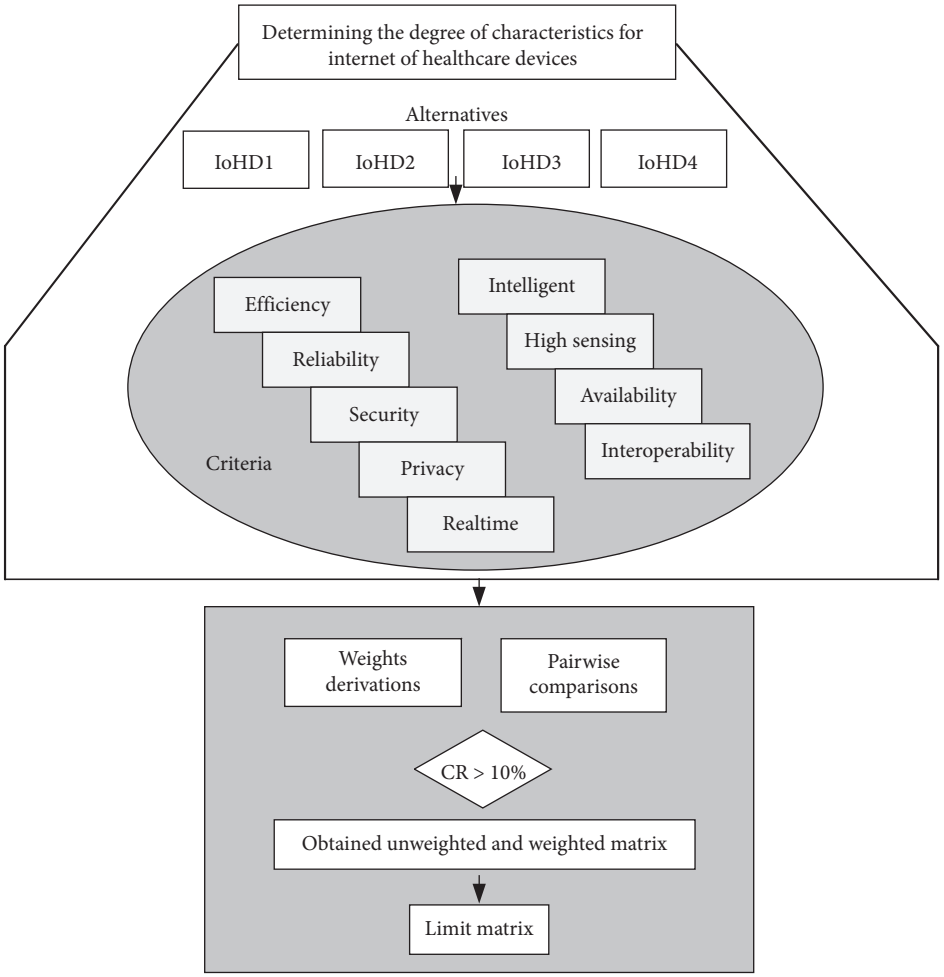


Figure 2: Network of the criteria, alternatives, and the application of fuzzy ANP.

selected. Table 2 shows the common features and characteristics of the IoH devices along with the references given.

3.3. Process of the Fuzzy ANP for Determining the Characteristics of IoH Devices. Fuzzy ANP is the combination of fuzzy and ANP which are used for uncertain, vague, and complex situation where decision-making is difficult. The proposed method has used the network of elements and clusters and fuzzy set of membership range from $1'$ to $9'$ domain. Fuzzy set is described by membership function [43]. It is extremely beneficial where decision-making is about uncertainty and vagueness. Here the decision can be made in linguistics variable instead of numbers; on the other side, the ANP is a more general form of analytic hierarchy process. In fuzzy ANP, triangular membership functions are used to represent opinions and the triangular membership function has (l, m, u) parameters which are defined as $l \leq m \leq u$. In the proposed method, triangular fuzzy numbers from $1'$ to $9'$ are used for pairwise comparisons of the selection process. Figure 1 depicts the proposed research process for determination of characteristics of IoH devices.

Figure 2 briefly shows the network of the criteria, alternatives, and the application of fuzzy ANP. Four different alternatives were presented for the process of research and nine criteria were presented. The whole process of the experiments was followed for conducting the proposed process.

Once the criteria, alternatives, and their network of elements are drawn, the process of fuzzy ANP was started and various comparison measures have been done. Table 3 gives the representation of values assigned to the Internet of healthcare device (IoHD) 1.

After assignment of the values to different features of the IoHD, the process of comparisons was started. Table 4 depicts the value obtained of Table 3 after comparison process.

The same process was applied on the remaining IoHD. The comparisons process for each feature was done. Table 5 describes the value assignment to the available features of IoHD.

Table 3: Value assignment to IoHD1.

	Efficiency	Reliability	Security	Privacy	Real time	Intelligent	High sensing	Availability	Interoperability
Efficiency	1	[2, 4]	[2, 4]	[8, 10]	[3, 5]	[4, 6]	[6, 8]	[8, 10]	[8, 10]
Reliability	[1/2, 1/4]	1	[1, 3]	[2, 4]	[1, 3]	[2, 4]	[4, 6]	[2, 4]	[8, 10]
Security	[1/2, 1/4]	[1, 1/3]	1	[1, 3]	[1, 3]	[1, 3]	[1, 3]	[2, 4]	[6, 8]
Privacy	[1/8, 1/10]	[1/2, 1/4]	[1, 1/3]	1	[1, 3]	[2, 4]	[1, 3]	[1, 3]	[4, 6]
Real time	[1/3, 1/5]	[1, 1/3]	[1, 1/3]	[1, 1/3]	1	[1, 3]	[1, 3]	[1, 3]	[2, 4]
Intelligent	[1/4, 1/6]	[1/2, 1/4]	[1, 1/3]	[1/2, 1/4]	[1, 1/3]	1	[3, 5]	[2, 4]	[1, 3]
High sensing	[1/6, 1/8]	[1/4, 1/6]	[1, 1/3]	[1, 1/3]	[1, 1/3]	[1/3, 1/5]	1	[4, 6]	[1, 3]
Availability	[1/8, 1/10]	[1/2, 1/4]	[1/2, 1/4]	[1, 1/3]	[1, 1/3]	[1/2, 1/4]	[1/4, 1/6]	1	[1, 3]
Interoperability	[1/8, 1/10]	[1/8, 1/10]	[1/6, 1/8]	[1/4, 1/6]	[1/2, 1/4]	[1, 1/3]	[1, 1/3]	[1, 1/3]	1

Table 4: Comparison process of Table 3.

	Efficiency	Reliability	Security	Privacy	Real time	Intelligent	High sensing	Availability	Interoperability	EV
Efficiency	0.386	0.475	0.354	0.528	0.312	0.293	0.295	0.316	0.225	0.354
Reliability	0.129	0.158	0.236	0.176	0.156	0.176	0.211	0.105	0.225	0.175
Security	0.129	0.079	0.118	0.117	0.156	0.117	0.084	0.105	0.175	0.120
Privacy	0.043	0.053	0.059	0.059	0.156	0.176	0.084	0.070	0.125	0.092
Real time	0.096	0.079	0.059	0.029	0.078	0.117	0.084	0.070	0.075	0.077
Intelligent	0.077	0.053	0.059	0.020	0.039	0.059	0.169	0.105	0.050	0.070
High sensing	0.055	0.032	0.059	0.029	0.039	0.015	0.042	0.175	0.050	0.055
Availability	0.043	0.053	0.039	0.029	0.039	0.020	0.008	0.035	0.050	0.035
Interoperability	0.043	0.018	0.017	0.012	0.026	0.029	0.021	0.018	0.025	0.023

Table 5: Values assignment to the features.

	IoHD1	IoHD2	IoHD3	IoHD4
IoHD1	1	[4, 6]	[3, 5]	[1, 3]
IoHD2	[1/4, 1/6]	1	[1/3, 1/5]	[1/5, 1/7]
IoHD3	[1/3, 1/5]	[3, 5]	1	[1/2, 1/4]
IoHD4	[1, 1/3]	[5, 7]	[2, 4]	1

Table 6: Comparisons of Table 5.

	IoHD1	IoHD2	IoHD3	IoHD4	E.V
IoHD1	0.513	0.313	0.485	0.571	0.470
IoHD2	0.103	0.063	0.030	0.048	0.061
IoHD3	0.128	0.250	0.121	0.095	0.149
IoHD4	0.256	0.375	0.364	0.286	0.320

The same process of comparisons was done and the calculated values are obtained in Table 6.

4. Results and Discussion

The process of fuzzy ANP was done in order to achieve efficient results of the available IoHD for their effective usage and to determine the characteristics of the IoH devices. Initially, all the steps of the process were followed. After the comparisons process of all the features and the available IoHD, the values were combined into a single table called the unweighted matrix. Table 7 depicts the unweighted super matrix.

The unweighted super matrix was converted into the weighted super matrix. The purpose of this matrix was to obtain the limit matrix. Table 8 shows the weighted super matrix.

Table 8 was converted into limit matrix for the purpose of decision. This matrix was obtained through taking the power of the given matrix. The limit matrix is shown in Table 9. This matrix was achieved for the purpose of decision-making.

On the basis of the limit matrix obtained, the decisions regarding the available IoHDs are made. Figure 3 briefly represents the details of ranking of the available IoHDs.

Table 7: Unweighted super matrix.

		Features									Alternatives			
		Efficiency	Reliability	Security	Privacy	Real time	Intelligent	High sensing	Availability	Interoperability	IoHD1	IoHD2	IoHD3	IoHD4
Features	Efficiency	0.000	0.299	0.368	0.379	0.359	0.380	0.379	0.284	0.377	0.354	0.315	0.348	0.309
	Reliability	0.351	0.000	0.173	0.179	0.158	0.176	0.156	0.197	0.147	0.175	0.193	0.199	0.205
	Security	0.186	0.190	0.000	0.118	0.122	0.121	0.124	0.136	0.134	0.120	0.127	0.100	0.155
	Privacy	0.127	0.133	0.119	0.000	0.105	0.102	0.105	0.114	0.129	0.092	0.102	0.100	0.092
	Real time	0.104	0.110	0.103	0.098	0.000	0.076	0.086	0.099	0.082	0.077	0.076	0.074	0.073
	Intelligent	0.077	0.085	0.077	0.076	0.094	0.000	0.065	0.077	0.060	0.070	0.086	0.085	0.082
	High sensing	0.074	0.083	0.064	0.064	0.067	0.064	0.000	0.059	0.038	0.055	0.043	0.040	0.034
	Availability	0.049	0.065	0.065	0.055	0.056	0.044	0.045	0.000	0.033	0.035	0.037	0.032	0.032
	Interoperability	0.032	0.034	0.032	0.032	0.039	0.037	0.041	0.034	0.000	0.023	0.020	0.023	0.019
Alternatives	IoHD1	0.322	0.277	0.606	0.519	0.470	0.130	0.044	0.093	0.558	0.000	0.000	0.000	0.000
	IoHD2	0.137	0.161	0.110	0.201	0.061	0.582	0.186	0.349	0.122	0.000	0.000	0.000	0.000
	IoHD3	0.093	0.096	0.232	0.079	0.149	0.169	0.293	0.097	0.263	0.000	0.000	0.000	0.000
	IoHD4	0.448	0.466	0.052	0.201	0.320	0.118	0.477	0.461	0.057	0.000	0.000	0.000	0.000

Table 8: Weighted super matrix.

		Features									Alternatives			
		Efficiency	Reliability	Security	Privacy	Real time	Intelligent	High sensing	Availability	Interoperability	IoHD1	IoHD2	IoHD3	IoHD4
Features	Efficiency	0.000	0.150	0.184	0.189	0.180	0.190	0.189	0.142	0.188	0.354	0.315	0.348	0.309
	Reliability	0.176	0.000	0.086	0.089	0.079	0.088	0.078	0.098	0.074	0.175	0.193	0.199	0.205
	Security	0.093	0.095	0.000	0.059	0.061	0.060	0.062	0.068	0.067	0.120	0.127	0.100	0.155
	Privacy	0.064	0.066	0.059	0.000	0.052	0.051	0.053	0.057	0.065	0.092	0.102	0.100	0.092
	Real time	0.052	0.055	0.052	0.049	0.000	0.038	0.043	0.050	0.041	0.077	0.076	0.074	0.073
	Intelligent	0.039	0.042	0.039	0.038	0.047	0.000	0.033	0.038	0.030	0.070	0.086	0.085	0.082
	High sensing	0.037	0.042	0.032	0.032	0.034	0.032	0.000	0.029	0.019	0.055	0.043	0.040	0.034
	Availability	0.024	0.033	0.032	0.027	0.028	0.022	0.022	0.000	0.017	0.035	0.037	0.032	0.032
	Interoperability	0.016	0.017	0.016	0.016	0.019	0.019	0.020	0.017	0.000	0.023	0.020	0.023	0.019
Alternatives	IoHD1	0.161	0.139	0.303	0.260	0.235	0.065	0.022	0.047	0.279	0.000	0.000	0.000	0.000
	IoHD2	0.068	0.081	0.055	0.100	0.030	0.291	0.093	0.174	0.061	0.000	0.000	0.000	0.000
	IoHD3	0.046	0.048	0.116	0.039	0.074	0.084	0.147	0.049	0.132	0.000	0.000	0.000	0.000
	IoHD4	0.224	0.233	0.026	0.100	0.160	0.059	0.239	0.230	0.028	0.000	0.000	0.000	0.000

Table 9: Limit matrix.

		Features									Alternatives			
		Efficiency	Reliability	Security	Privacy	Real time	Intelligent	High sensing	Availability	Interoperability	IoHD1	IoHD2	IoHD3	IoHD4
Features	Efficiency	0.193	0.193	0.193	0.193	0.193	0.193	0.193	0.193	0.193	0.193	0.193	0.193	0.193
	Reliability	0.128	0.128	0.128	0.127	0.127	0.127	0.128	0.127	0.128	0.128	0.127	0.128	0.128
	Security	0.089	0.090	0.089	0.089	0.089	0.089	0.090	0.089	0.090	0.090	0.089	0.089	0.090
	Privacy	0.068	0.068	0.068	0.068	0.068	0.068	0.068	0.068	0.068	0.068	0.068	0.068	0.068
	Real time	0.056	0.056	0.056	0.056	0.056	0.056	0.056	0.056	0.056	0.056	0.056	0.056	0.056
	Intelligent	0.051	0.051	0.051	0.051	0.051	0.051	0.051	0.051	0.051	0.051	0.051	0.051	0.051
	High sensing	0.037	0.037	0.037	0.037	0.037	0.037	0.037	0.037	0.037	0.037	0.037	0.037	0.037
	Availability	0.029	0.029	0.029	0.029	0.029	0.029	0.029	0.029	0.029	0.029	0.029	0.029	0.029
	Interoperability	0.018	0.018	0.018	0.018	0.018	0.018	0.018	0.018	0.018	0.018	0.018	0.018	0.018
Alternatives	IoHD1	0.117	0.117	0.117	0.117	0.117	0.117	0.117	0.117	0.117	0.117	0.117	0.117	0.117
	IoHD2	0.061	0.061	0.061	0.061	0.061	0.061	0.061	0.061	0.061	0.061	0.061	0.061	0.061
	IoHD3	0.046	0.046	0.046	0.046	0.046	0.046	0.046	0.046	0.046	0.046	0.046	0.046	0.046
	IoHD4	0.110	0.110	0.110	0.110	0.110	0.110	0.110	0.110	0.110	0.110	0.110	0.110	0.110

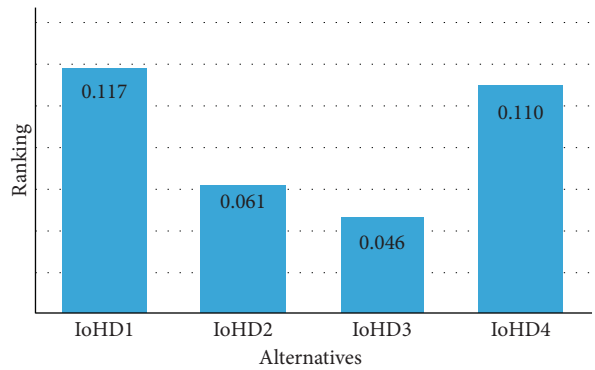


Figure 3: Ranking of the available alternatives (IoHDs).

From the figure, it is shown that IoHD1 is the best choice among the available alternatives. This is followed by IoHD4.

5. Conclusions

With the revolution in digital technology and Internet, every organization is adopting digital things to carry out their day-to-day activities. IoT is a perception used to link various devices over the Internet for increasing the production and quality of service, deliver a huge amount of data in seconds, and automate the processes of daily life. IoT implementation in the industrial sector has transformed the typical industrial setup into smart and intellectual setup. With the "smart" and "Intelligent" abilities of IoT devices such as sensors and with the collaboration of humans and computers, physical processes can be monitored and, based on the received data, optimal decisions can ultimately be taken. Industry 4.0 is increasing flexibility, mass customization, quality, and productivity. The Internet of Things associates various heterogeneous devices, and unfortunate determination of the degree of characteristics of IoH devices may affect the effectiveness of services. Keeping in view this issue, the proposed research is carried out with the use of decision support system and application of fuzzy analytic network process was considered for optimum determination of the degree of characteristics of IoH devices based on exclusive properties which would significantly upturn the efficiency of industry and its endeavors.

Data Availability

No data were used to support this study.

Conflicts of Interest

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

References

[1] D. Zhang, C. C. Chan, and G. Y. Zhou, "Enabling industrial internet of things (IIoT) towards an emerging smart energy system," *Global energy interconnection*, vol. 1, no. 1, pp. 39–47, 2018.

[2] L. Urquhart and D. McAuley, "Avoiding the internet of insecure industrial things," *Computer Law & Security Report*, vol. 34, no. 3, pp. 450–466, 2018.

[3] S. Vaidya, P. Ambad, and S. Bhosle, "Industry 4.0-a glimpse," *Procedia manufacturing*, vol. 20, pp. 233–238, 2018.

[4] M. Mahbub, "Comparative link-level analysis and performance estimation of channel models for IIoT (industrial-IoT) wireless communications," *Internet of Things*, vol. 12, Article ID 100315, 2020.

[5] L. Silvestri, A. Forcina, V. Introna, A. Santolamazza, and V. Cesarotti, "Maintenance transformation through industry 4.0 technologies: a systematic literature review," *Computers in Industry*, vol. 123, p. 103335, 2020.

[6] R. Y. Zhong, X. Xu, E. Klotz, and S. T. Newman, "Intelligent manufacturing in the context of industry 4.0: a review," *Engineering*, vol. 3, no. 5, pp. 616–630, 2017.

[7] L. Barreto, A. Amaral, and T. Pereira, "Industry 4.0 implications in logistics: an overview," *Procedia Manufacturing*, vol. 13, pp. 1245–1252, 2017.

[8] F. Civerchia, S. Bocchino, C. Salvadori, E. Rossi, L. Maggiani, and M. Petracca, "Industrial internet of things monitoring solution for advanced predictive maintenance applications," *Journal of Industrial Information Integration*, vol. 7, pp. 4–12, 2017.

[9] C. Cronin, A. Conway, and J. Walsh, "Flexible manufacturing systems using IIoT in the automotive sector," *Procedia Manufacturing*, vol. 38, pp. 1652–1659, 2019.

[10] Q. Li, Y. Yue, and Z. Wang, "Deep Robust Cramer Shoup delay optimized fully homomorphic for IIOT secured transmission in cloud computing," *Computer Communications*, vol. 161, pp. 10–18, 2020.

[11] A. Xenakis, A. Karageorgos, E. Lallas, A. E. Chis, and H. González-Vélez, "Towards distributed IoT/cloud based fault detection and maintenance in industrial automation," *Procedia Computer Science*, vol. 151, pp. 683–690, 2019.

[12] F. Zezulka, P. Marcon, Z. Bradac, J. Arm, T. Benesl, and I. Vesely, "Communication systems for industry 4.0 and the IIoT," *IFAC-PapersOnLine*, vol. 51, no. 6, pp. 150–155, 2018.

[13] L. Zhou, H. Guo, and G. Deng, "A fog computing based approach to DDoS mitigation in IIoT systems," *Computers & Security*, vol. 85, pp. 51–62, 2019.

[14] J. Arm, F. Zezulka, Z. Bradac et al., "Implementing industry 4.0 in discrete manufacturing: options and drawbacks," *IFAC-PapersOnLine*, vol. 51, no. 6, pp. 473–478, 2018.

[15] F. Al-Obaidy, F. Yazdani, and F. A. Mohammadi, "Intelligent testing for arduino UNO based on thermal image," *Computers & Electrical Engineering*, vol. 58, pp. 88–100, 2017.

[16] S. Goudarzi, M. H. Anisi, A. H. Abdullah, J. Lloret, S. A. Soleymani, and W. H. Hassan, "A hybrid intelligent model for network selection in the industrial internet of things," *Applied Soft Computing*, vol. 74, pp. 529–546, 2019.

[17] M. A. Khan and K. Salah, "IoT security: review, blockchain solutions, and open challenges," *Future Generation Computer Systems*, vol. 82, pp. 395–411, 2018.

[18] W. Liu, G. Huang, A. Zheng, and J. Liu, "Research on the optimization of IIoT data processing latency," *Computer Communications*, vol. 151, pp. 290–298, 2020.

[19] X. Luo, J. Liu, D. Zhang, and X. Chang, "A large-scale web QoS prediction scheme for the industrial internet of things based on a kernel machine learning algorithm," *Computer Networks*, vol. 101, pp. 81–89, 2016.

[20] M. Al-Hawawreh, E. Sitnikova, and F. den Hartog, "An efficient intrusion detection model for edge system in brownfield industrial internet of things," in *Proceedings of the*

3rd International Conference on Big Data and Internet of Things, pp. 83–87, Melbourn, Australia, August 2019.

[21] J. Dobaj, J. Iber, M. Krisper, and C. Kreiner, "A microservice architecture for the industrial internet-of-things," in *Proceedings of the 23rd European Conference on Pattern Languages of Programs*, pp. 1–15, Irsee, Germany, July 2018.

[22] M. Al-Hawawreh and E. Sitnikova, "Industrial internet of things based ransomware detection using stacked variational neural network," in *Proceedings of the 3rd International Conference on Big Data and Internet of Things*, pp. 126–130, Melbourn, Australia, August 2019.

[23] J. Ruh and W. Steiner, "The need for deterministic virtualization in the industrial internet of things," in *Proceedings of the Workshop on Fog Computing and the IoT*, pp. 26–30, Montreal, Canada, April 2019.

[24] C. Kan, H. Yang, and S. Kumara, "Parallel computing and network analytics for fast Industrial internet-of-things (IIoT) machine information processing and condition monitoring," *Journal of Manufacturing Systems*, vol. 46, pp. 282–293, 2018.

[25] S. Berger, O. Bürger, and M. Röglinger, "Attacks on the industrial internet of things-development of a multi-layer taxonomy," *Computers & Security*, vol. 93, Article ID 101790, 2020.

[26] J. Wan, J. Li, Q. Hua, A. Celesti, and Z. Wang, "Intelligent equipment design assisted by cognitive internet of things and industrial big data," *Neural Computing & Applications*, vol. 32, no. 9, pp. 4463–4472, 2020.

[27] N. Jayakumar and D. P. Joshi, "Big data & disruptive computing platforms braced internet of things: facets & trends," in *Studies in Systems, Decision and Control, Internet of Things, Smart Computing and Technology: A Roadmap Ahead*, pp. 119–150, Springer, Berlin, Germany, 2020.

[28] J. Chanchaichujit, A. Tan, F. Meng, and S. Eaimkhong, "Internet of things (IoT) and big data analytics in healthcare," in *Healthcare 4.0*, pp. 17–36, Springer, Berlin, Germany, 2019.

[29] E. Sisinni, A. Saifullah, S. Han, U. Jennehag, and M. Gidlund, "Industrial internet of things: challenges, opportunities, and directions," *IEEE Transactions on Industrial Informatics*, vol. 14, no. 11, pp. 4724–4734, 2018.

[30] H. Boyes, B. Hallaq, J. Cunningham, and T. Watson, "The industrial internet of things (IIoT): an analysis framework," *Computers in Industry*, vol. 101, pp. 1–12, 2018.

[31] J. H. Park, *Advances in Future Internet and the Industrial Internet of Things*, Multidisciplinary Digital Publishing Institute, Basel, Switzerland, 2019.

[32] Y. Chen, G. Lee, L. Shu, and N. Crespi, "Industrial internet of things-based collaborative sensing intelligence: framework and research challenges," *Sensors*, vol. 16, no. 2, p. 215, 2016.

[33] A. Liu, Q. Zhang, Z. Li, Y.-j. Choi, J. Li, and N. Komuro, "A green and reliable communication modeling for industrial internet of things," *Computers & Electrical Engineering*, vol. 58, pp. 364–381, 2017.

[34] H. Xu, W. Yu, D. Griffith, and N. Golmie, "A survey on industrial internet of things: a cyber-physical systems perspective," *IEEE Access*, vol. 6, pp. 78238–78259, 2018.

[35] W. Z. Khan, M. H. Rehman, H. M. Zangoti, M. K. Afzal, N. Armi, and K. Salah, "Industrial internet of things: recent advances, enabling technologies and open challenges," *Computers & Electrical Engineering*, vol. 81, Article ID 106522, 2020.

[36] H. Ezz El-Din and D. H. Manjaiah, "Internet of nano things and industrial internet of things," in *Internet of Things: Novel Advances and Envisioned Applications*, pp. 109–123, Springer, Berlin, Germany, 2017.

[37] H. Mouratidis and V. Diamantopoulou, "A security analysis method for industrial internet of things," *IEEE Transactions on Industrial Informatics*, vol. 14, no. 9, pp. 4093–4100, 2018.

[38] Q. Yan, W. Huang, X. Luo, Q. Gong, and F. R. Yu, "A multi-level DDoS mitigation framework for the industrial internet of things," *IEEE Communications Magazine*, vol. 56, no. 2, pp. 30–36, 2018.

[39] C. Yin, J. Xi, R. Sun, and J. Wang, "Location privacy protection based on differential privacy strategy for big data in industrial internet of things," *IEEE Transactions on Industrial Informatics*, vol. 14, no. 8, pp. 3628–3636, 2017.

[40] W. Chen, "Intelligent manufacturing production line data monitoring system for industrial internet of things," *Computer Communications*, vol. 151, pp. 31–41, 2020.

[41] N. Muthukumar, S. Srinivasan, K. Ramkumar, D. Pal, J. Vain, and S. Ramaswamy, "A model-based approach for design and verification of industrial internet of things," *Future Generation Computer Systems*, vol. 95, pp. 354–363, 2019.

[42] A.-H. Muna, 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.

[43] Z. Ayağ, R. G. ÖzdemİR, An intelligent approach to ERP software selection through fuzzy ANP," *International Journal of Production Research*, vol. 45, no. 10, pp. 2169–2194, 2007.

Hindawi
Scientific Programming
Volume 2021, Article ID 9999368, 9 pages
https://doi.org/10.1155/2021/9999368

Research Article

Improved 3D U-Net for COVID-19 Chest CT Image Segmentation

Ruiyong Zheng, Yongguo Zheng, and Changlei Dong-Ye

Collage of Computer Science and Engineering, Shandong University of Science and Technology, Qingdao, Shandong 266590, China

Correspondence should be addressed to Yongguo Zheng; skd991317@sdust.edu.cn

Received 5 April 2021; Accepted 10 June 2021; Published 25 June 2021

Academic Editor: Shah Nazir

Coronavirus disease 2019 (COVID-19) has spread rapidly worldwide. The rapid and accurate automatic segmentation of COVID-19 infected areas using chest computed tomography (CT) scans is critical for assessing disease progression. However, infected areas have irregular sizes and shapes. Furthermore, there are large differences between image features. We propose a convolutional neural network, named 3D CU-Net, to automatically identify COVID-19 infected areas from 3D chest CT images by extracting rich features and fusing multiscale global information. 3D CU-Net is based on the architecture of 3D U-Net. We propose an attention mechanism for 3D CU-Net to achieve local cross-channel information interaction in an encoder to enhance different levels of the feature representation. At the end of the encoder, we design a pyramid fusion module with expanded convolutions to fuse multiscale context information from high-level features. The Tversky loss is used to resolve the problems of the irregular size and uneven distribution of lesions. Experimental results show that 3D CU-Net achieves excellent segmentation performance, with Dice similarity coefficients of 96.3% and 77.8% in the lung and COVID-19 infected areas, respectively. 3D CU-Net has high potential to be used for diagnosing COVID-19.

1. Introduction

Coronavirus disease 2019 (COVID-19) has rapidly spread worldwide since its outbreak in December 2019 [1, 2]. In March 2020, the World Health Organization declared COVID-19 as a global pandemic [3].

The reverse transcription polymerase chain reaction (RT-PCR) test is the standard for COVID-19 detection. However, this test has a high false negative rate, and it cannot accurately detect the initial infection. Hence, infected patients cannot be diagnosed on time [4, 5]. Compared with the RT-PCR test, chest computed tomography (CT) provides higher sensitivity in the diagnosis of COVID-19; therefore, it can be used as one of the main clinical detection methods [6, 7].

The chest CT scans of patients with COVID-19 show characteristic imaging features, such as ground-glass opacity and occasional consolidation plaques in the lungs [8–10], which are considerably useful for diagnosing COVID-19 and evaluating the severity of a patient's condition. However, owing to a significant increase in the number of patients, it has become quite challenging to use chest CT scans for COVID-19 detection because of the large workload and experience requirements for doctors.

Numerous deep learning methods have been used to segment and quantitatively analyse infected areas in chest CT scans [11–15]. Li et al. [11] proposed an automatic neural network architecture to detect COVID-19 from chest CT scans and distinguish it from other types of pneumonia and lung diseases. Fan et al. [12] used a Siamese convolutional neural network to assay COVID-19 and automatically evaluate the severity of lung diseases. Gao et al. [13] improved the detection capacity of the network for small lesions and improved the interpretability of the network using a lesion attention module with a dual-branch combination network and an attention mechanism. Paluru et al. [14] proposed Anam-Net to address anomalies in COVID-19 chest CT images. A lightweight convolutional neural network was embedded in Anam-Net, which contained significantly fewer parameters compared to U-Net. The use of Anam-Net in embedded systems demonstrated its applicability to point-of-care platforms. Yan et al. [15] proposed a

high-accuracy network (COVID-SegNet) to segment COVID-19 lesions from chest CT images. COVID-SegNet used multiscale feature fusion and enhanced features to segment lung and COVID-19 lesions accurately and automatically. Although these methods play an important role in the diagnosis and analysis of COVID-19, they are based on CT slices. These methods frequently neglect the correlation between continuous CT slices and cannot fully utilise the spatial information of CT scans.

It is challenging to automatically segment the lesions of COVID-19 pneumonia because of the complexity of CT spatial imaging, the difficulty of marking infected areas, and the difference between medical image characteristics. First, infections may have different characteristic appearances, such as ground-glass opacity and consolidation plaques. Second, lesions have irregular shapes and fuzzy boundaries, and a few lesions have a lower contrast compared to surrounding areas. Third, it is tedious and time consuming to artificially mark pulmonary infection, and it is frequently influenced by doctors' knowledge and clinical experience of lesions [9, 10, 12, 16].

We propose a deep learning method, named 3D CU-Net, to improve the segmentation performance of the neural network models for COVID-19. In addition, we propose a new feature encoding module (residual channel attention, Res_CA) for 3D CU-Net. In the feature extraction stage, the channel attention mechanism of local cross-channel information interaction is used to recalibrate the feature weight of global information and enhance the performance of feature representation. We propose a pyramid fusion module with multiscale global information interaction in the bottom encoder, which enhances the performance of the network by fusing the feature information of different scales and improving the performance of the network for lesion area segmentation.

2. Related Work

2.1. U-Net Structure. U-Net [17] was proposed by Ronneberger et al. in 2015 for medical cell segmentation. It consists of a contraction path to obtain context information and an expansion path to recover a feature map. As high-level and low-level semantic information has the same importance in image segmentation, U-Net combines the high-definition features of an encoder with the advanced semantic features of a decoder stage to help restore the details of a target and obtain an accurate output.

2.2. Variants of U-Net. Numerous methods based on U-Net have achieved better results in different medical image segmentation tasks by integrating the new design concepts of networks. Oktay et al. [18] added an attention mechanism based on U-Net for targets with different shapes and sizes and used an attention gate to highlight the salient features of a skip connection. Xiao et al. proposed a model, named Res U-Net [19], with a weighted attentional mechanism to deal with extreme changes in the ocular vascular background. Feng et al. proposed CPFNet [20], which improved the segmentation performance by utilising two pyramid modules to fuse multiscale context information.

Wang et al. proposed a new cross-channel information interaction method, named ECA-Net [21], to recalibrate features. They prevented the adverse effect of dimensionality reduction in SE-Net on channel attention. However, most U-shaped networks use only abstract features, neglect certain details, and cannot effectively use multiscale context information [20].

3. Proposed Method

3.1. Network Overview. We propose an automatic segmentation model, named 3D CU-Net, for COVID-19 lesions. The model is based on the 3D U-Net architecture, as shown in Figure 1. The network structure of 3D CU-Net is composed of a feature encoding module (Res_CA) with an attention mechanism, a pyramid dilated convolution module (PDS block) for extracting and fusing multiscale information at different resolutions, and a feature decoding module for segmentation. A fixed-size 3D slice extracted from a 3D CT image is used as the input of the network. The predicted segmentation result is obtained after a series of upsampling and downsampling operations for feature encoding and decoding. The model can ensure continuity between CT images and retain a certain amount of interlayer information. Thus, the 3D input contains more contextual information compared to a 2D image.

In the feature encoding part, an efficient channel attention mechanism [21] is used to reallocate feature weights under the guidance of global information, and residual networks are used to mitigate problems such as gradient vanishing. Global average pooling is used to obtain multiscale global information under different receptive fields to enhance the feature representation in the PDS module, thereby improving the segmentation performance of the network for the irregular shapes and sizes of lesions. Finally, segmentation results are obtained by a feature decoding module, which includes two consecutive $23 \times 3 \times 3$ convolutions and a residual connection with a $1 \times 1 \times 1$ convolution.

3.2. Feature Encoding Block. As shown in Figure 2, the feature encoding module mainly consists of the following two parts.

3.2.1. Feature Extraction. In each encoding module, except for the bottom encoder, two continuous $3 \times 3 \times 3$ convolutions are used to extract deeper feature information. This expands the receptive field, extracts more feature information, improves the complexity of the network, and reduces the amount of calculation and number of parameters. After each $3 \times 3 \times 3$ convolution, we add the ReLU activation function and batch normalisation to alleviate the problem of gradient disappearance and increase the speed of network learning.

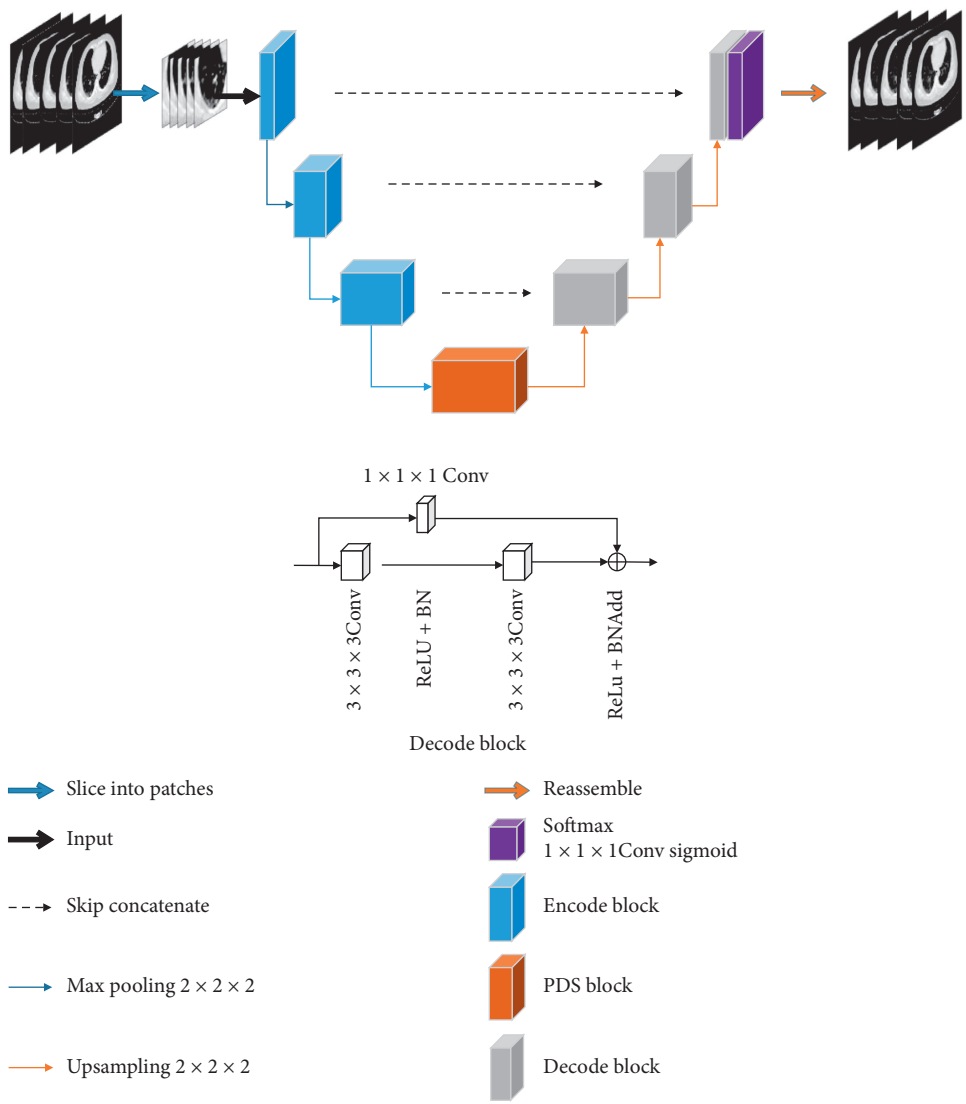


Figure 1: Overview of 3D CU-Net.

3.2.2. Feature Calibration Block. We introduce a channel attention mechanism to obtain representative features and highlight useful information. According to the correlation between adjacent channels, cross-channel interactive fusion methods are used to recalibrate the weights of the extracted features. Cross-channel information communication can effectively prevent the influence of the reduction in dimensions on channel attention and enhance the feature representation of lesion areas.

3.3. Pyramid Fusion Module for Dilated Convolutional Global Information Interaction. Multiscale context information helps improve the performance of semantic segmentation. Thus, we propose a pyramid fusion module that converts low-scale global information into high-scale features. As shown in Figure 3, a residual block is used to deepen the network and extract feature information. Then, a parallel expanded convolution with expanded sizes of 1, 2, and 4 is used to obtain the multiscale information of advanced features. Next, according to the correlation between feature channel information at different scales, global average pooling is used to obtain the global channel features and their weights at different scales. Thus, the global information obtained in a small receptive field is used to enhance the feature expression ability of a large receptive field. Finally, the features at different scales are fused by stitching.

In the last part of this module, we connect the multiscale feature information that has been recalibrated with feature weights, normalise it using a $1 \times 1 \times 1$ convolution, and then fuse it with the original advanced features.

3.4. Feature Decoding Block. As shown in the decoding block in Figure 1, two $3 \times 3 \times 3$ convolutions and a residual connection with a $1 \times 1 \times 1$ convolution are applied to the feature map after the series connection, and a feature map is obtained with the same size as that of the original input image.

3.5. Loss Function. In the medical image segmentation task (lesion detection), the high imbalance of the training data leads

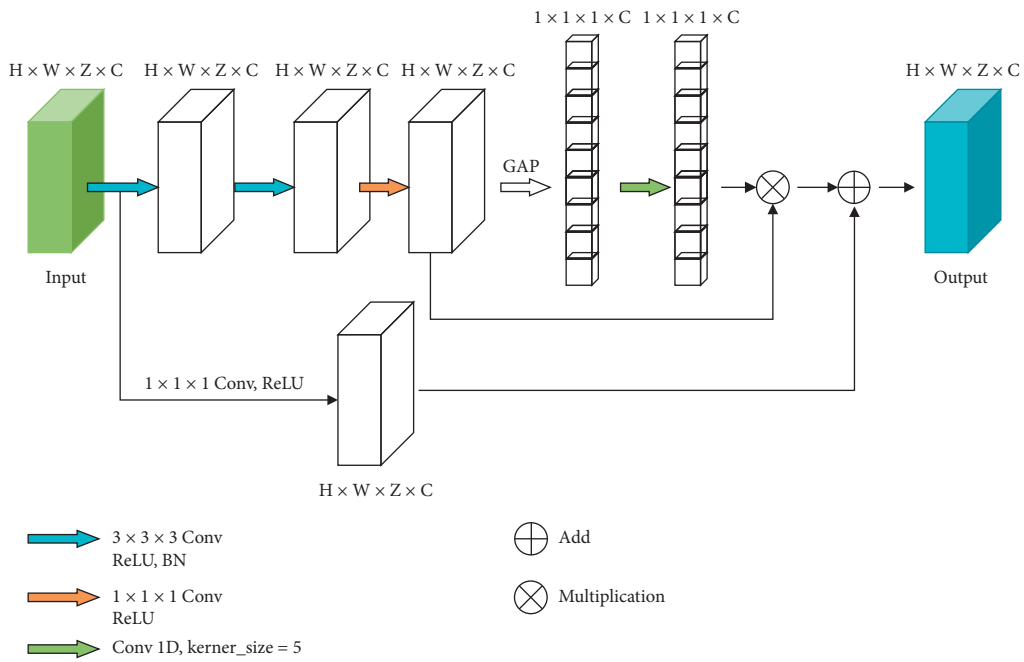


Figure 2: Network structure of the feature encoding block.

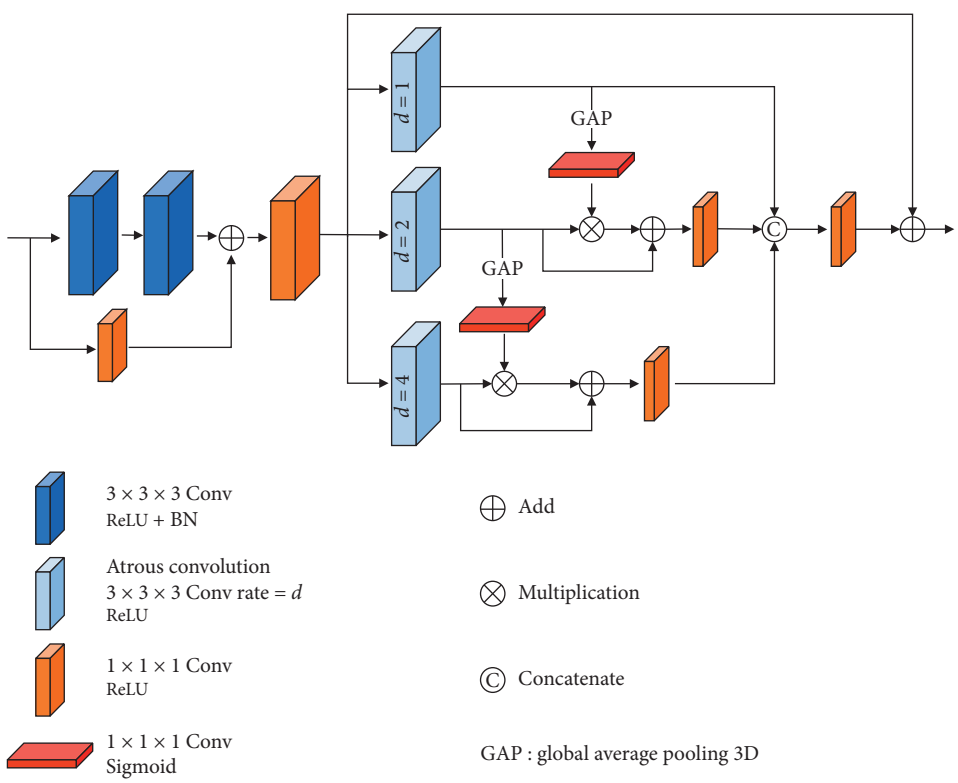


Figure 3: Convolutional global information interaction fusion network with pyramid expansion.

to high precision and low recall. Lesion areas account for only 1.12% of the total area in our data. We employ the combination of categorical cross-entropy loss and the Tversky loss as the loss function of segmentation to resolve the problem of unbalanced data category distribution and improve the performance of network generalisation. The loss function is expressed as

Table 1: Comparison of results obtained using 3D CU-Net and 3D U-Net.

	Lung_left			Lung_right			Lesion		
	DSC	Sens	Spec	DSC	Sens	Spec	DSC	Sens	Spec
3D U-Net	0.925	0.943	0.997	0.948	0.959	0.997	0.705	0.707	0.999
3D CU-Net Ours_1: $\alpha = 0.5$, $\beta = 0.5$	0.966	0.968	0.998	0.964	0.975	0.997	0.778	0.738	0.999
3D CU-Net Ours_2: $\alpha = 0.3$, $\beta = 0.7$	0.960	0.969	0.998	0.963	0.966	0.998	0.771	0.837	0.998

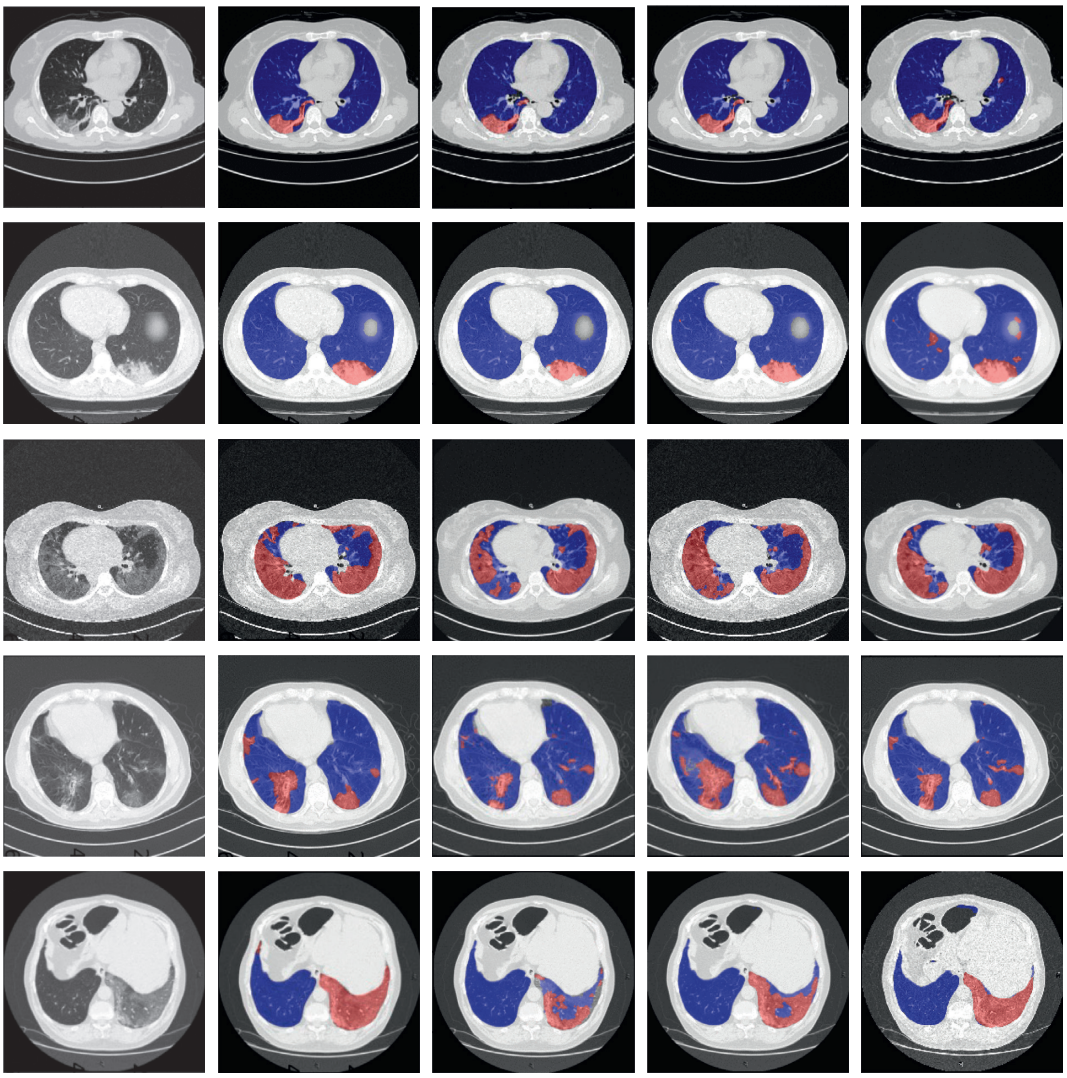


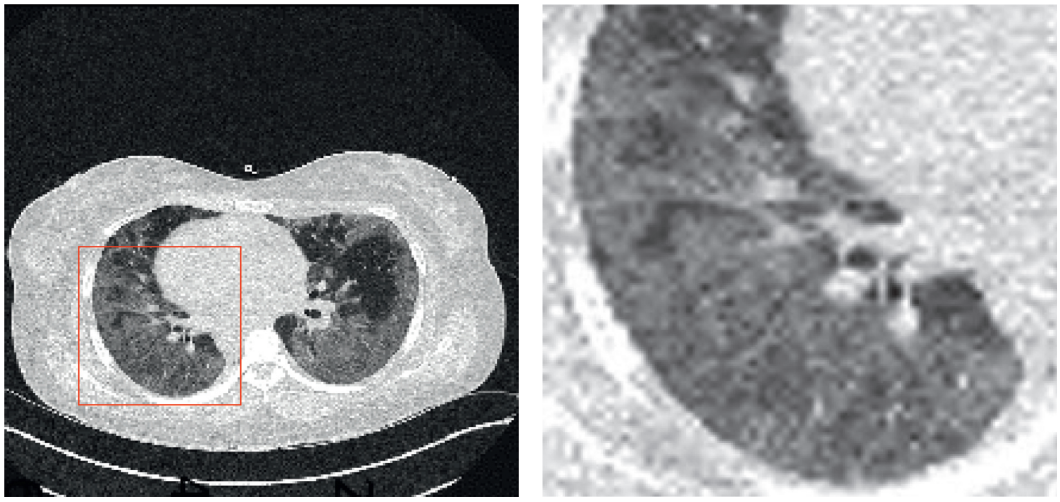
Figure 4: Visualisation of segmentation results.

$$\begin{aligned} \mathrm{Loss}_{\mathrm{Total}} &= \mathrm{Loss}_{\mathrm{CE}} + \mathrm{Loss}_{\mathrm{Tversky}}, \\ \mathrm{Loss}_{\mathrm{CE}} &= -\frac{1}{N}\sum_{i=1}^{N} y_i \log f_i(x), \\ \mathrm{Loss}_{\mathrm{Tversky}} &= N - \sum_{i=1}^{N} \frac{\mathrm{TP}(c)}{\mathrm{TP}(c) + \alpha \mathrm{FN}(c) + \beta \mathrm{FP}(c)}, \end{aligned} \tag{1}$$

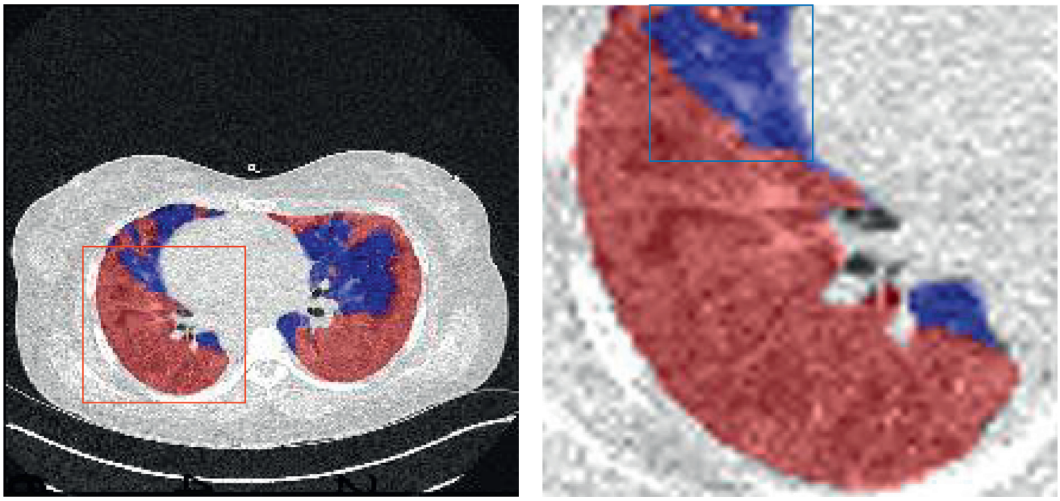
where x is the input value, y_i is the true label corresponding to category i, and $f(x_i)$ is the model output value. TP, FN, and FP represent true positive, false negative, and false positive, respectively.

4. Experiments and Results

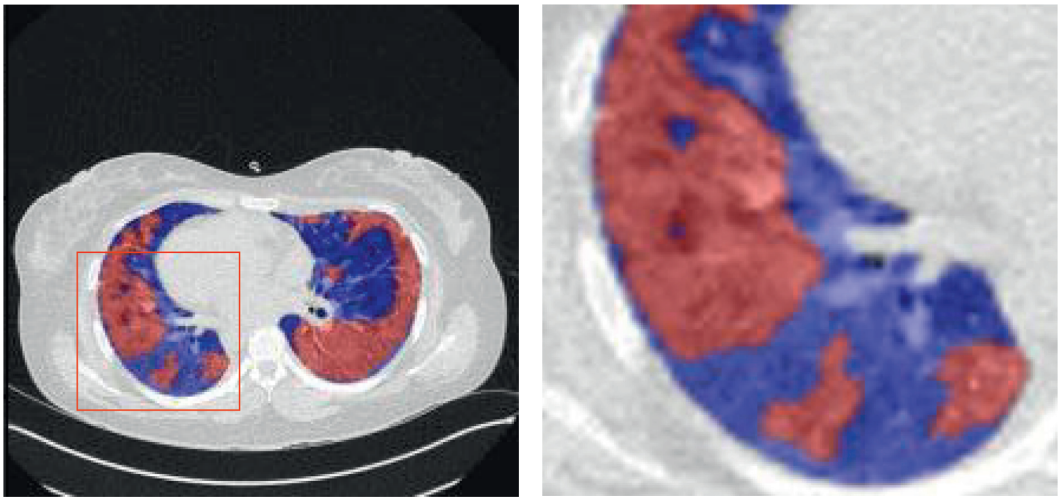
4.1. Experimental Data. We train and evaluate 3D CU-Net using the open COVID-19 CT dataset provided by Jun et al.



(a)

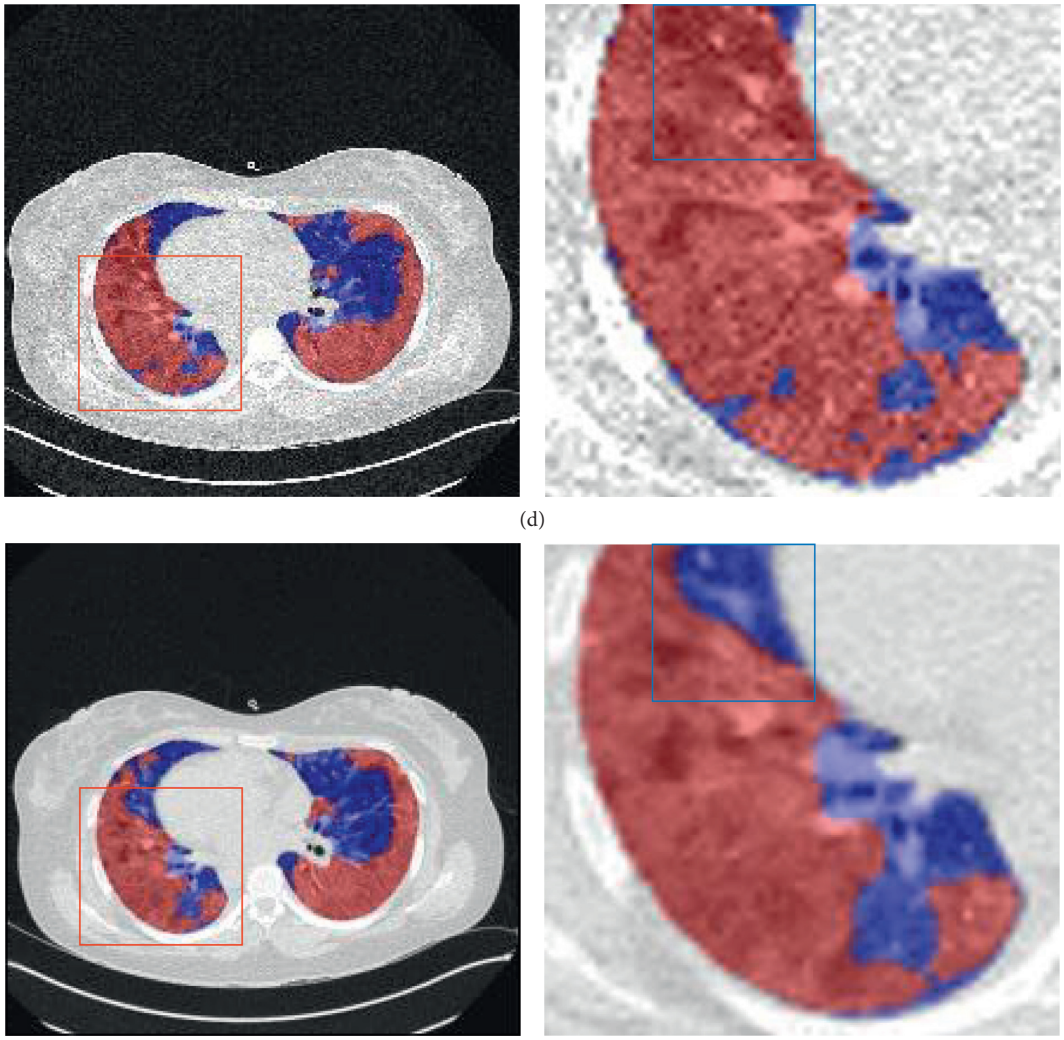


(b)



(c)

Figure 5: Continued.



(d)

(e)

Figure 5: Local details of COVID-19 segmentation results.

[22]. In addition, MosMedData [23] provided by Forrest is used as an independent test dataset to further verify the performance of the model. The COVID-19 CT dataset consists of the chest CT scans of 20 COVID-19 patients validated with annotations by a senior radiologist, in which the left lung, right lung, and lesion areas are annotated. The dataset contains 3250 CT slices with sizes of 630×630, 512×512, and 401×630. The lesion area accounts for only 2.12% of the CT slice area, and the slices with lesion markers account for 52.86% of the total slices. MosMedData was provided by Moscow Municipal Hospital. It consists of the chest CT scans of 50 confirmed COVID-19 patients, with lesion areas annotated by a few experts.

Ma et al. [24] proposed a COVID-19 infection segmentation benchmark based on 3D nnU-Net. The Dice coefficients of the COVID-19 CT dataset and MosMedData are 0.673 and 0.588, respectively.

We normalise the training data and pixel values between [0, 1] by considering −250 and 1250 as thresholds. In addition, 3D CT images are resampled at a fixed isomorphic resolution to normalise them into the same voxel spacing. We use random elastic deformation, random rotation, random scaling, Gaussian noise, and other common medical image data enhancement methods to enhance the training data and prevent the overfitting problem caused by a small amount of training data.

4.2. Experimental Details and Evaluation Metrics. 3D CU-Net is compared with standard 3D U-Net [25] in terms of segmentation results, and the performance of 3D CU-Net is further analysed using MosMedData.

We build an operating environment on a Linux server. The NVIDIA Tesla P100 GPU is used, and the TensorFlow 2.0 deep learning framework is adopted. The installation environment comprises cuda10.0, cudnn7.6.5, python3.6, opencv, gcc, etc. In the fitting process, we set the batch size as 2. We use Adam optimisation with an initial learning rate of 0.001 and a minimum learning rate of 0.00001. We reduce the learning rate by 0.1 times when loss does not decrease in

Table 2: Comparison of results on the basis of MosMedData.

		Lesion	
	DSC	Sens	Spec
3D U-Net	0.609	0.584	0.999
3D CU-Net Ours_1: $\alpha = 0.5$, $\beta = 0.5$	0.664	0.651	0.999
3D CU-Net Ours_2: $\alpha = 0.3$, $\beta = 0.7$	0.668	0.734	0.998

15 epochs. The training process ends when loss does not decrease in 50 epochs.

We employ 5-fold cross-validation for model fitting. Sixteen sets of CT images are used for fitting and the remaining four sets for validation. After each fitting process, the model is evaluated using the validation data.

In addition, after fitting, we utilise three widely used metrics in medical image analysis to evaluate the segmentation performance of the model for the left and right lungs and COVID-19 lesion areas. These metrics are the Dice similarity coefficient (DSC), sensitivity (Sens), and specificity (Spec).

4.3. Experimental Results. As shown in Table 1, the DSC is 77.8% and Sens is 73.8% for 3D CU-Net, compared with 3D U-Net, and the DSC and Sens increased by 7.3% and 3.1%, respectively. By adjusting the parameters of the Tversky loss ($\alpha = 0.3$ and $\beta = 0.7$), Sens for 3D CU-Net increases to 83.7% with few losses of DSC. In addition, the accuracy of overall lung segmentation improves.

Figure 4 shows the segmentation results obtained using 3D CU-Net and 3D U-Net for five different slices. The images from left to right are the original CT image, ground truth, segmentation results of 3D U-Net, results of 3D CU-Net, and results of 3D CU-Net with the Tversky loss parameters as $\alpha = 0.3$ and $\beta = 0.7$.

Figure 5 shows the local details of the CT image slice segmentation results shown in the third row of Figure 4. In the first column of Figure 5, rows (a)–(e) show the original CT image, ground truth, segmentation result of 3D U-Net, segmentation result of 3D CU-Net, and segmentation result of 3D CU-Net with the Tversky loss parameters as $\alpha = 0.3$ and $\beta = 0.7$, respectively. The second column shows the details of the area enclosed by the red box in the first column. 3D U-Net shows poor segmentation performance, and a large infected area is not identified, as shown in row (c). In contrast, 3D CU-Net provides better segmentation performance, and most infected areas are accurately identified, as shown in row (d). The area enclosed by the blue box in the second column of Figure 5 shows that setting $\alpha = 0.3$ and $\beta = 0.7$ effectively reduces the false positive rate of 3D CU-Net and improves the sensitivity of infection region segmentation.

Furthermore, we compare the performance of the model in terms of infected area segmentation on the basis of MosMedData. As shown in Table 2, the performance of 3D CU-Net is better than that of 3D U-Net, with a 5.9% improvement in the DSC and a 15% increase in Sens.

The experiments performed using the COVID-19 CT dataset and MosMedData show that 3D CU-Net provides excellent segmentation performance. For the left lung, right lung, and lesion areas, the DSC is 0.960, 0.963, and 0.771, Sens is 0.969, 0.966, and 0.837, and Spec is 0.998, 0.998, and 0.998, respectively.

It has great potential in evaluating COVID-19 infection. The above results suggest that the 3D CU-Net model has good performance in COVID-19 lesion segmentation.

5. Conclusion

We proposed a deep learning segmentation network (3D CU-Net) for detecting COVID-19 pulmonary infection. The proposed network was based on 3D U-Net. An attention mechanism was introduced for channel features in the encoding stage to enhance the representation ability of features. The full utilisation of the multiscale global information of high-level features extracted from the bottom encoder improved the accuracy of COVID-19 detection. The proposed network has high potential to be used for diagnosing COVID-19.

However, 3D CU-Net has certain limitations. Its accuracy must be improved for the irregular shapes and different sizes of lesions. In addition, the segmentation performance can be improved via further research and by utilising high-quality medical imaging data.

Data Availability

The labeled datasets used to support the findings of this study are publicly available.

Conflicts of Interest

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

Acknowledgments

The authors would like to thank Editage (http://www.editage.com) for English language editing. This work was supported by the 2016 Excellent Teaching Team Construction Plan of Shandong University of Science and Technology.

References

[1] H. Lu, C. W. Stratton, and Y. W. Tang, "Outbreak of pneumonia of unknown etiology in Wuhan, China: the mystery and the miracle," *Journal of Medical Virology*, vol. 92, no. 4, pp. 401-402, 2020.

[2] C. Wang, P. W. Horby, F. G. Hayden, and G. F. Gao, "A novel coronavirus outbreak of global health concern," *The Lancet*, vol. 395, no. 10223, pp. 470–473, 2020.

[3] World Health Organization (WHO), *WHO Director-General's Opening Remarks at the media Briefing on COVID-19 11 March 2020*, WHO, Geneva, Switzerland, 2020, https://www.who.int/dg/speeches/detail/who-director-general-s-opening-remarks-at-the-media-briefing-on-covid-19. 11-march-2020.

[4] Y. Fang, H. Zhang, J. Xie et al., "Sensitivity of chest CT for COVID-19: comparison to RT-PCR," *Radiology*, vol. 296, no. 2, pp. E115–E117, 2020.

[5] T. Ai, Z. Yang, H. Hou et al., "Correlation of chest CT and RT-PCR testing for coronavirus disease 2019 (COVID-19) in China: a report of 1014 cases," *Radiology*, vol. 296, no. 2, pp. E32–E40, 2020.

[6] X. Xie, Z. Zhong, W. Zhao, C. Zheng, F. Wang, and J. Liu, "Chest CT for typical coronavirus disease 2019 (COVID-19) pneumonia: relationship to negative RT-PCR testing," *Radiology*, vol. 296, no. 2, pp. E41–E45, 2020.

[7] N. Zhu, D. Zhang, W. Wang et al., "A novel coronavirus from patients with pneumonia in China, 2019," *New England Journal of Medicine*, vol. 382, no. 8, pp. 727–733, 2020.

[8] J. Lei, J. Li, X. Li, and X. Qi, "CT imaging of the 2019 novel coronavirus (2019-nCoV) pneumonia," *Radiology*, vol. 295, no. 1, p. 18, 2020.

[9] M.-Y. Ng, E. Y. P. Lee, J. Yang et al., "Imaging profile of the COVID-19 infection: radiologic findings and literature review," *Radiology: Cardiothoracic Imaging*, vol. 2, no. 1, Article ID e200034, 2020.

[10] M. Chung, A. Bernheim, X. Mei et al., "CT imaging features of 2019 novel coronavirus (2019-nCoV)," *Radiology*, vol. 295, no. 1, pp. 202–207, 2020.

[11] M. D. Li, N. T. Arun, M. Gidwani et al., "Automated assessment and tracking of COVID-19 pulmonary disease severity on chest radiographs using convolutional siamese neural networks," *Radiology: Artificial Intelligence*, vol. 2, no. 4, Article ID e200079, 2020.

[12] D.-P. Fan, T. Zhou, G.-P. Ji et al., "Inf-net: automatic COVID-19 lung infection segmentation from CT images," *IEEE Transactions on Medical Imaging*, vol. 39, no. 8, pp. 2626–2637, 2020.

[13] K. Gao, J. Su, Z. Jiang et al., "Dual-branch combination network (DCN): towards accurate diagnosis and lesion segmentation of COVID-19 using CT images," *Medical Image Analysis*, vol. 67, Article ID 101836, 2021.

[14] N. Paluru, A. Dayal, H. B. Jenssen et al., "ANAM-net: anamorphic depth embedding-based lightweight CNN for segmentation of anomalies in COVID-19 chest CT images," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 32, no. 3, pp. 932–946, 2021.

[15] Q. Yan, B. Wang, D. Gong et al., "COVID-19 chest CT image segmentation network by multi-scale fusion and enhancement operations," *IEEE Transactions on Big Data*, vol. 7, no. 1, pp. 13–24, 2021.

[16] L. Li, L. Qin, Z. Xu et al., "Using artificial intelligence to detect COVID-19 and community-acquired pneumonia based on pulmonary CT: evaluation of the diagnostic accuracy," *Radiology*, vol. 296, no. 2, pp. E65–E71, 2020.

[17] O. Ronneberger, P. Fischer, and T. Brox, "U-net: convolutional networks for biomedical image segmentation," in *Lecture Notes in Computer Science*, pp. 234–241, Springer, Cham, Switzerland, 2015.

[18] O. Oktay, J. Schlemper, L. L. Folgoc et al., "Attention u-net: Learning where to look for the pancreas," 2018, https://arxiv.org/abs/1804.03999.

[19] X. Xiao, S. Lian, Z. Luo et al., "Weighted res-unet for high-quality retina vessel segmentation," in *Proceedings of the 9th International Conference on Information Technology in Medicine and Education (ITME)*, pp. 327–331, IEEE, Hangzhou, China, October 2018.

[20] S. Feng, H. Zhao, F. Shi et al., "CPFNet: context pyramid fusion network for medical image segmentation," *IEEE Transactions on Medical Imaging*, vol. 39, no. 10, pp. 3008–3018, 2020.

[21] Q. Wang, B. Wu, P. Zhu, P. Li, W. Zuo, and Q. Hu, "ECA-Net: efficient channel attention for deep convolutional neural networks," in *Proceedings of theCVF Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2020.

[22] M. Jun, G. Cheng, W. Yixin et al., "Covid-19 ct lung and infection segmentation dataset [DB/OL]. Zenodo," 2020, https://zenodo.org/record/3757476#.YLov8vkzaUk.

[23] S. P. Morozov, A. E. Andreychenko, I. A. Blokhin et al., "MosMedData: data set of 1110 chest CT scans performed during the COVID-19 epidemic," *Digital Diagnostics*, vol. 1, no. 1, pp. 49–59, 2020.

[24] J. Ma, Y. Wang, X. An et al., "Towards efficient covid-19 ct annotation: a benchmark for lung and infection segmentation," 2020, https://arxiv.org/abs/2004.12537.

[25] D. Müller, I. S. Rey, and F. Kramer, "Automated chest CT image segmentation of COVID-19 lung infection based on 3D U-Net," 2020, https://arxiv.org/abs/2007.04774.

Hindawi
Scientific Programming
Volume 2021, Article ID 3137083, 7 pages
https://doi.org/10.1155/2021/3137083

Research Article

Concrete Crack Detection Algorithm Based on Deep Residual Neural Networks

Xiuying Meng

Henan Vocational College of Water Conservancy and Environment, Zhengzhou 450000, China

Correspondence should be addressed to Xiuying Meng; mengxiuyinghb@163.com

Received 29 April 2021; Revised 22 May 2021; Accepted 26 May 2021; Published 24 June 2021

Academic Editor: Shah Nazir

Crack is the early expression form of the concrete pavement disease. Early discovery and treatment of it can play an important role in the maintenance of the pavement. With ongoing advancements in computer hardware technology, continual optimization of deep learning algorithms, as compared to standard digital image processing algorithms, utilizing automation of crack detection technology has a deep learning algorithm that is more exact. As a result of the benefits of greater robustness, the study of concrete pavement crack picture has become popular. In view of the poor effect and weak generalization ability of traditional image processing technology on image segmentation of concrete cracks, this paper studies the image segmentation algorithm of concrete cracks based on convolutional neural network and designs an end-to-end segmentation model based on ResNet101. It integrates more low-level features, which make the fracture segmentation results more refined and closer to the practical application scenarios. Compared with other methods, the algorithm in this paper has achieved higher detection accuracy and generalization ability.

1. Introduction

Concrete structure [1] is the basic element of constructing urban environment, which needs to be maintained systematically. Crack [2] is a common damage form of concrete structure, which is one of the important indexes to evaluate the safety of concrete structure [3, 4]. For bridges [5–7], roads [8, 9], and other engineering structures [10, 40], it is necessary to carry out regular management and maintenance; crack detection [11–14] is an important part of management and maintenance. At present, the crack detection is mainly artificial detection, which has a lot of problems, such as in the crack detection of the bridge, the staff need to observe the cracks on the bottom and side of the bridge by building scaffolding, which requires a lot of preparation work and low detection efficiency. The safety of staff working in high places is also difficult to be guaranteed. There are blind spots in the manual detection process, and the inspection effect is often affected by the professional accomplishment and experience level of the detection personnel, and the detection accuracy is low. Due to these shortcomings of manual detection methods, a crack damage detection method based on crack image is developed [15–18].

The traditional detection of crack damage is based on pictures. In the early stage, the detection method based on morphology is often low and its detection accuracy is often low and is greatly affected by light and noise. In the later stage, methods such as seed points and tensor voting are used. Better detection accuracy however calculations are frequently complex, and computation efficiency is low. For crack pictures with complex backgrounds, false detections and missed detections may occur, so it is necessary to develop other concrete crack detection methods.

In recent years, with the continuous improvement of Internet technology [19] and the support of related hardware, artificial intelligence technology with deep learning [20] as the core has developed rapidly. Convolutional neural network is a feed-forward neural network that applies deep learning. It has significant advantages such as high parallelism, good robustness, and strong generalization ability. Due to the characteristics of weight sharing, the

convolutional neural network has a much smaller amount of calculation than the BP neural network, which reduces the burden on the computer and does not require manual feature extraction. These all make it have a wide range of applications in the field of computer vision, especially in image classification and detection. When there are enough samples, the classifier of the trained convolutional neural network can accurately identify concrete damage. Nowadays, there are many convolutional neural network models with more efficient training and more accurate classification every year. The detection of concrete cracks through convolutional neural networks helps to reduce the subjective influence of people and is more efficient and economical. It is an excellent solution to the problem of crack detection. Following are the main innovative points of this paper:

To propose an improved fully convolutional neural network model based on ResNet101, using ResNet101 as the backbone network for feature extraction and then fusing the extracted features of different scales with the up-sampling and restoring feature maps. Because it combines shallow position information and high-level semantic information, it makes the image edge detail segmentation more refined.

To use the concrete crack detection method based on deep residual neural network proposed in this paper is a nondestructive detection technology, which has urgent needs and extremely high application value in the field.

The rest of this paper is organized as follows: Section 2 shows the related work of the paper. Section 3 discusses the methodology of the paper. The experiments and results are shown in Section 4. Section 5 shows the conclusion of the paper.

2. Related Work

In recent years, convolutional neural networks have been widely used in the field of civil engineering. For cracks and other structural damage problems, scholars usually use classified convolutional neural networks to locate the damaged parts and use semantic segmentation networks to describe the damaged parts at pixel level to achieve the description of the damage.

Classical convolutional neural networks include AlexNet [21] and GoogLeNet [22] which have specific requirements on the size of the input image. When using these networks, the image to be tested needs to be divided into small samples of specific sizes. The network is trained to classify these samples, determine the position of the crack sample on the picture to be tested, and realize the location of the crack in the picture.

Dorafshan et al. [23] compared the detection effects of common edge detectors and deep convolutional neural networks on concrete cracks. In this article, six edge detection methods, Roberts, Prewitt, Sobel, Laplacian of Gaussian, Butterworth, and Gaussian, are used to detect concrete crack images. The final binary image produces residual noise, and the best method can only detect cracks with a width greater than 0.1 mm. When using the deep convolutional neural network (DCNN) AlexNet, the correct rate of labeled samples reached 99%. The network pair in the migration learning mode can correctly detect 86% of the crack samples and can detect cracks with a width greater than 0.04 mm, while the network in the fully trained mode can also detect cracks greater than 0.08 mm. In terms of time, if the training process of the network is not considered, the calculation time of DCNN is also shorter than the most effective edge detection algorithm. These all show the superiority of using convolutional neural network to detect concrete cracks.

Kim and Cho [24] used concrete pictures on the Internet to construct a dataset, which included five categories, such as cracks and plants. AlexNet is trained by transfer learning, that is, the network parameters of the pretrained AlexNet are fine-tuned by using the existing dataset of concrete images. In the test process, two kinds of sliding windows with 50% overlap rate were used to scan the image so that the cracks appearing on the edge of window 1 appeared in the central area of window 2, so as to avoid the cracks being missed due to appearing on the edge of window, and the detection effect was enhanced by adjusting the threshold value of classification layer.

The classification network locates the cracks based on the crack samples and cannot accurately describe the shape of the cracks. Semantic segmentation is to classify each pixel in the image. Therefore, when the semantic segmentation network is applied to crack detection, the precise location and shape description of the crack can be realized at the same time.

Ye et al. [25] used full convolutional networks (FCN) to identify cracks in concrete bridge images at the pixel level and compared them with the results of edge detection methods. The results show that the edge detection method is greatly affected by noise, and when there is enough training data collected from the real situation, the full convolutional network (FCN) can eliminate a lot of noise interference, the performance is reliable, and the position of the crack in the image is displayed, and the ability of the path, its detection, has stronger robustness. Du et al. [26] proposed a new model based on the improvement of the original DeepLab v2 to adapt to the particularity of crack damage detection. Compared with the previous method, this method can obtain very high-precision output, and a new method of marking cracks is proposed, which is beneficial to the measurement of crack length and width. Hoskere et al. [27] proposed a structural damage detection method based on multiscale pixel-level deep convolutional neural network [28–31]. This method utilizes two networks [32], derived from VGG19 and ResNet45, for structural damage classification and semantic segmentation, respectively. By parallel, the two networks can simultaneously process multiple types of damage classification and pixel-level segmentation. Due to the high accuracy of damage detection in the classification network [33, 34], the segmentation network only needs to perform semantic segmentation [35] on the images judged as damage and does not need to identify the nondamage parts, which reduces the probability of false detection. Hoskere constructed a dataset of 1695 images from 250 different

structures, including six types of damage. On this dataset, the applicability of the method to civil infrastructure was verified, and the damage classification was achieved at the pixel level.

3. Methodology

Some scholars have proposed a variety of methods based on digital image processing to detect cracks on the surface of concrete structures, but there may be many problems such as uneven illumination intensity and distortion in the actual collected images, and the traditional image processing algorithms cannot solve these problems well. Due to its powerful feature extraction ability [36], deep learning technologies [37–39] can better overcome the interference brought by external environmental factors, and it has better detection accuracy compared with traditional methods. Therefore, this paper proposes a concrete crack detection algorithm based on deep residual neural network to achieve pixel-level segmentation detection of concrete crack images. Next in this section, the proposed network model for crack segmentation and the whole training process are described in detail.

In the maintenance of the surface of concrete structures, cracks are one of the most common and serious diseases and are the early features of many other diseases, as shown in Figure 1.

3.1. Segmentation Detection Algorithm Based on Deep Residual Neural Network. At present, most of the semantic segmentation network models based on deep learning adopt encoder-decoder network architecture, which is the mainstream structure of semantic segmentation network models due to its simple structure, strong scalability, and good segmentation effect. In this paper, the encode-decoder architecture is still adopted, and by combining the advantages of FCN network model and U-Net network model, an improved full convolutional neural network (CrackNet model for short) is proposed. CrackNet also contains the narrowing path for collecting image features and the expanding path for locating and restoring features. The two are symmetrical structures and constitute a U-shaped network. The deeper and more generalizable ResNet101 network is used in the CrackNet encoder. Compared with VGG16 network, ResNet101 network has fewer parameters and can extract deeper features. In the decoder, the advanced semantic features encoded are decoded and restored to the size of the original image by bilinear interpolation and convolution operation. Another important feature of CrackNet is that it splices and integrates the features of different scales extracted in the process of down-sampling with the feature images restored from up-sampling, which further enhances the segmentation effect of crack edge details. Figure 2 is the network structure diagram of this paper.

3.2. Model Training. The training process of our model mainly includes the selection of parameters such as fracture dataset annotation, loss function, optimization algorithm, and learning rate. Finally, this experiment completed the network model training under the Keras deep learning framework.

3.2.1. Data Labeling and Data Augmentation. Application scenarios of crack detection in this paper are mainly surface structures composed of concrete such as bridges and roads, but currently there is no standard crack dataset specifically used for semantic segmentation. In order to obtain enough concrete crack images to train the model, part of crack datasets was found in some papers, and a large number of crack images were supplemented by manual photography. Before training the network model, these collected crack images need to be annotated first. In this paper, the professional semantic segmentation and annotation software Labelme is used to annotate the crack images. There is a significant difference in the target detection task dataset of annotation and semantic segmentation, as in the mark of target detection. You only need to paint the rectangular box in the target and the appropriate categories, whereas annotation and semantic segmentation are more difficult. In particular, this slender aim requires cracking along the margins of each point to indicate until the contour map is complete. It forms an enclosed space. Figure 3 shows the process of labeling crack images with Labelme software.

For semantic segmentation tasks, a large number of training samples is the key factor to improve the segmentation effect of network models, but sometimes it is not easy to obtain a large number of datasets. An effective way to improve performance is to enlarge the existing data and expand the diversity of the existing data. This article uses Augmentor, a very useful Python data enhancement tool, to augment datasets. Its operation is very simple. According to the original image and label, the program performs rotation, stretch, zoom-in, and zoom-out of the labeled image with a certain probability to complete the expansion of the dataset. After augmenting the data by Augmentor library, this paper finally obtained 3000 fracture datasets, among which 2500 data were selected as the training set, and the remaining 500 were used as the test set.

3.2.2. Loss Function. The cross-direction loss function is one of the most widely used loss functions in deep learning. This article adopts the cross-direction loss function, but it is different from the cross-direction loss of the traditional image classification network in that the loss function in this paper needs to predict and evaluate each pixel and then calculate the pixel average. If the crossover loss is greater, the corresponding gradient will be greater and the training speed will be faster. Since this crack image segmentation task belongs to a two-class classification problem, the calculation equation of the cross-function is as follows:

$$L = y \log y' + (1 - y)\log(1 - y'), \tag{1}$$

where y is the truly marked crack value, y' is the predicted value of the model, L is the cross-child loss calculated according

Figure 1: Cracks on the concrete surface.

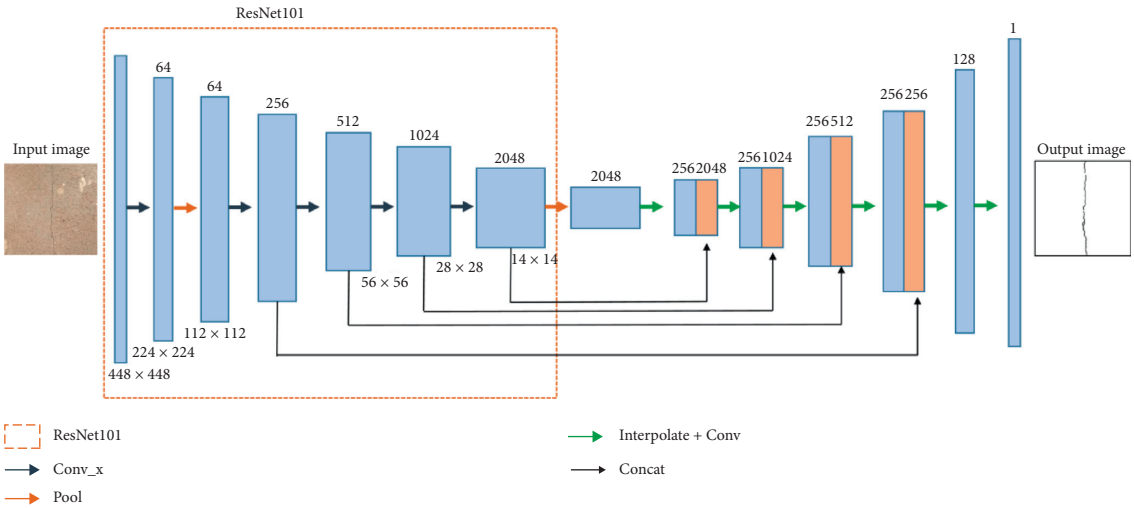


Figure 2: Network structure of the proposed model.

to the cross-child loss function, and the overall loss of the image is equal to the average loss value of all pixel points.

4. Experiments and Results

4.1. Experimental Setup. All the experiments in this section were performed on a Lenovo workstation with Nvidia GeForce GTX 1080 Ti and PyCharm installed. In this paper, the previously produced fracture dataset is divided into 2500 training sets and 500 test sets. Since the color information is of little significance to the concrete image and will take up too much storage space, in order to reduce this redundant information, this paper converts the image of three-channel concrete cracks captured into single-channel grayscale image. In order to maximize the use of computer memory without causing memory overflow, the batch-size of this experiment is 4, the size of the input image is 448×448, the loss function is the cross-saturated loss function, the optimization algorithm is the gradient descent method of driving quantity, the learning rate is 0.001, and the momentum coefficient Y is 0.9. The number of iterations is 100 epochs.

4.2. Evaluation Method. Precision and recall are two basic quantitative evaluation indexes in semantic segmentation. FP represents the total number of pixels that were wrongly judged as cracks in the background, and TP represents the total number of pixels that were correctly extracted from the cracks. FN represents the total number of pixels that belong

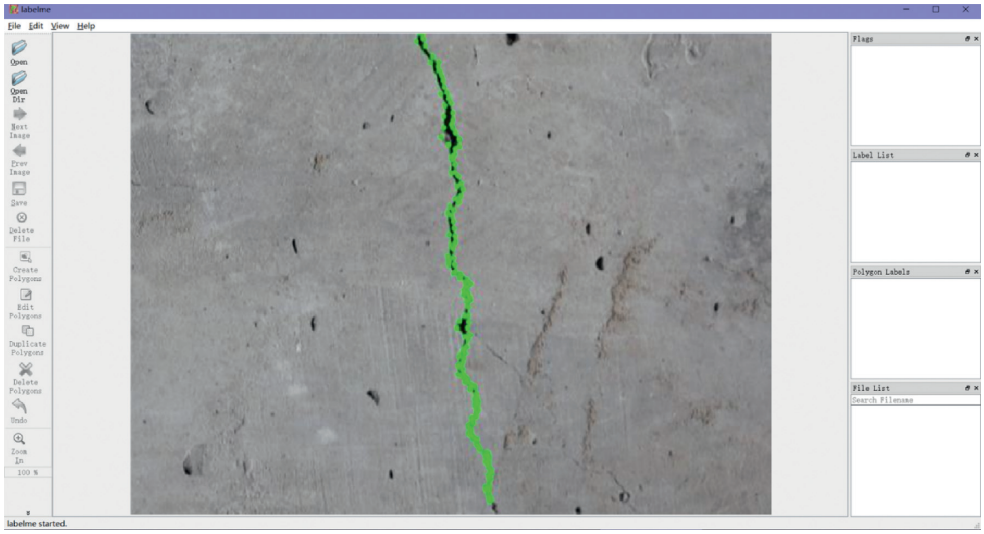
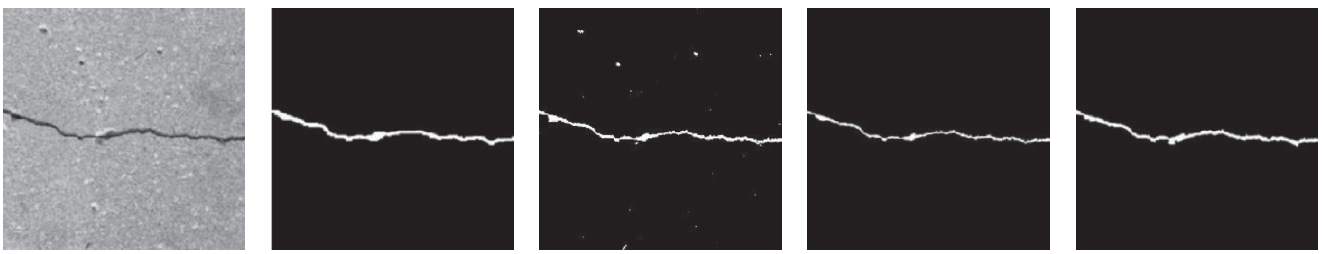
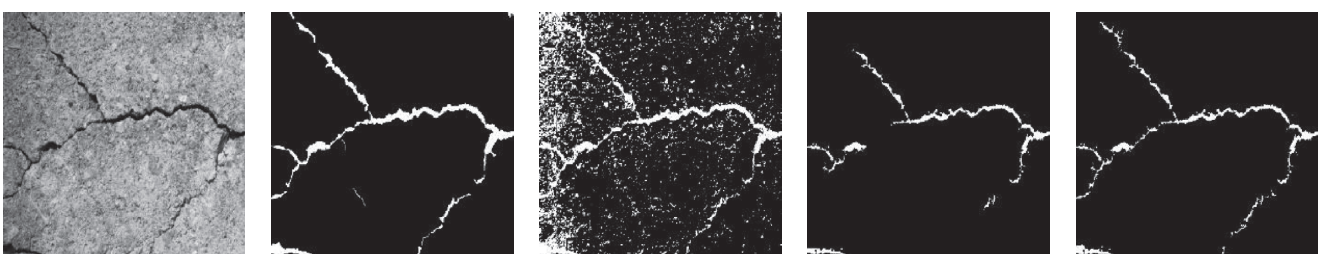


Figure 3: Data labeling.



(a) (b) (c) (d) (e)

Figure 4: Visualization results of segmentation detection. (a) Input image. (b) Ground truth. (c) FCN-8s. (d) FCN-32s. (e) Ours.



(a) (b) (c) (d) (e)

Figure 5: Visualization results of segmentation detection. (a) Input image. (b) Ground truth. (c) FCN-8s. (d) FCN-32s. (e) Ours.

to the crack area but are misjudged as the background. The calculation equations of fracture accuracy index P and recall rate index R are as follows:

$$P = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FP}}, \tag{2}$$

$$R = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FN}}. \tag{3}$$

4.3. Comparative Experiments. This paper conducts a comparative experiment with the classic semantic segmentation model FCN. Figures 4 and 5 show the visualization results of the comparative experiment.

Table 1: Comparison results of crack segmentation algorithms.

	FCN		Ours	
	P	R	P	R
Sample 1	0.9261	0.9366	0.9452	0.9525
Sample 2	0.9312	0.9425	0.9569	0.9655

It can be seen from Figures 4 and 5 that although the FCN segmentation method can overcome isolated noise points, the effect of wider crack width is acceptable, but because of the low-level features of its fusion, some cracks are not as detailed as others, and the noise extract is not as effectively split. The segmentation algorithm in this article can still maintain a relatively stable and accurate extraction

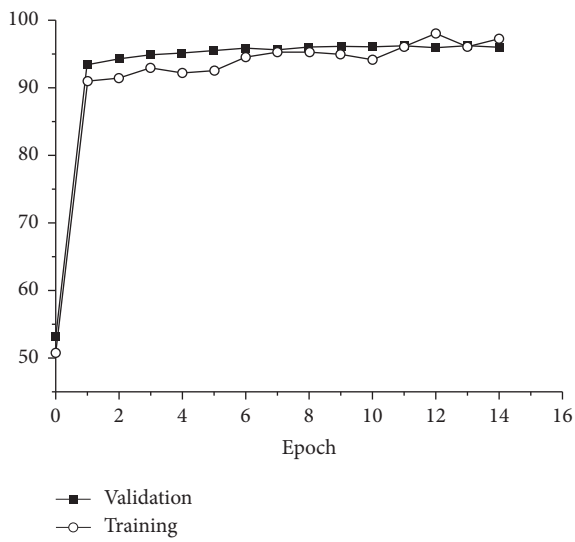


Figure 6: *P* curve during training.

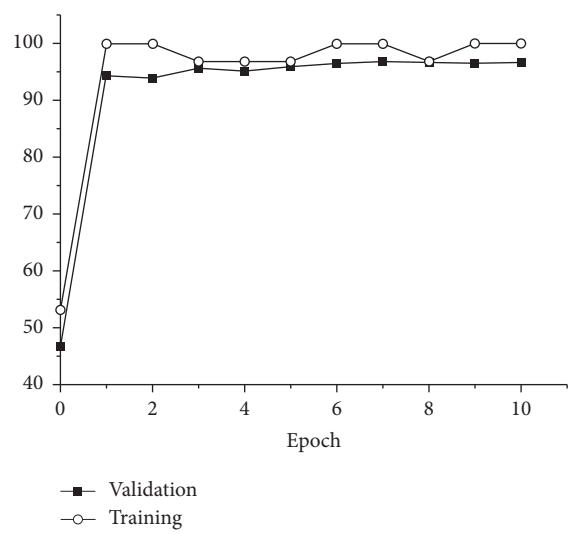


Figure 7: *R* curve during training.

for different types of cracks. It can be clearly seen from Table 1 that the algorithm in this paper is superior to the other two methods in all indicators. In addition, Figures 6 and 7 also show the *P* and *R* curves of the algorithm in the training process.

5. Conclusion

In this paper, an image segmentation algorithm for concrete cracks based on convolutional neural network is studied, and an end-to-end segmentation model based on ResNet101 is designed. In order to better train the network model, Labelme software was used to manually annotate the collected real scene fracture dataset and expand the data. Then, Keras deep learning framework was used to train the model, and the extracted feature map was visually displayed. Finally, the proposed algorithm is compared with the FCN method, and the experimental results show that the proposed segmentation model has better performance in segmentation accuracy and generalization ability.

Data Availability

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

Conflicts of Interest

The author does not have any possible conflicts of interest.

References

[1] S. W. Tang, Y. Yao, C. Andrade, and Z. J. Li, "Recent durability studies on concrete structure," *Cement and Concrete Research*, vol. 78, pp. 143–154, 2015.

[2] S. Kabir, "Imaging-based detection of AAR induced map-crack damage in concrete structure," *NDT & E International*, vol. 43, no. 6, pp. 461–469, 2010.

[3] G. Peiying, D. Chang, and T. Lei, "Determination of local damage probability in concrete structure," *Procedia Engineering*, vol. 28, pp. 489–493, 2012.

[4] L. Yang and H. Jianguo, "Comparison of safety level of Chinese and foreign crack control of concrete members," *Building Structure*, vol. 2, 2011.

[5] P. L. Fuhr and D. R. Huston, "Corrosion detection in reinforced concrete roadways and bridges via embedded fiber optic sensors," *Smart Materials and Structures*, vol. 7, no. 2, pp. 217–228, 1998.

[6] P. L. Fuhr and D. R. Huston, "Corrosion detection in reinforced concrete roadways and bridges via embedded fiber optic sensors," *Smart Materials and Structures*, vol. 7, no. 2, pp. 217–228, 1998.

[7] S. Yehia, O. Abudayyeh, S. Nabulsi, and I. Abdelqader, "Detection of common defects in concrete bridge decks using nondestructive evaluation techniques," *Journal of Bridge Engineering*, vol. 12, no. 2, pp. 215–225, 2007.

[8] H. Maeda, Y. Sekimoto, T. Seto, T. Kashiyama, and H. Omata, "Road damage detection and classification using deep neural networks with smartphone images," *Computer-Aided Civil and Infrastructure Engineering*, vol. 33, no. 12, pp. 1127–1141, 2018.

[9] M. Foedisch and A. Takeuchi, "Adaptive real-time road detection using neural networks," in *Proceedings. The 7th International IEEE Conference on Intelligent Transportation Systems (IEEE Cat. No. 04TH8749)*, pp. 167–172, IEEE, Washington, WA, USA, October 2004.

[10] M. Z. Naser, R. A. Hawileh, and J. A. Abdalla, "Fiber-reinforced polymer composites in strengthening reinforced concrete structures: a critical review," *Engineering Structures*, vol. 198, Article ID 109542, 2019.

[11] C. V. Dung and L. D. Anh, "Autonomous concrete crack detection using deep fully convolutional neural network," *Automation in Construction*, vol. 99, pp. 52–58, 2019.

[12] A. Mohan and S. Poobal, "Crack detection using image processing: a critical review and analysis," *Alexandria Engineering Journal*, vol. 57, no. 2, pp. 787–798, 2018.

[13] Q. Zou, Z. Zhang, Q. Li, X. Qi, Q. Wang, and S. Wang, "Deepcrack: learning hierarchical convolutional features for crack detection," *IEEE Transactions on Image Processing*, vol. 28, no. 3, pp. 1498–1512, 2018.

[14] X. Yang, H. Li, Y. Yu, X. Luo, T. Huang, and X. Yang, "Automatic pixel-level crack detection and measurement using

fully convolutional network," *Computer-Aided Civil and Infrastructure Engineering*, vol. 33, no. 12, pp. 1090–1109, 2018.

[15] T. Nishikawa, J. Yoshida, T. Sugiyama, and Y. Fujino, "Concrete crack detection by multiple sequential image filtering," *Computer-Aided Civil and Infrastructure Engineering*, vol. 27, no. 1, pp. 29–47, 2012.

[16] H. G. Moon and J. H. Kim, "Intelligent crack detecting algorithm on the concrete crack image using neural network," in *Proceedings of the 28th ISARC*, pp. 1461–1467, Seoul, South Korea, June 2011.

[17] B. Kim and S. Cho, "Image-based concrete crack assessment using mask and region-based convolutional neural network," *Structural Control and Health Monitoring*, vol. 26, no. 8, p. e2381, 2019.

[18] Y. Ren, J. Huang, Z. Hong et al., "Image-based concrete crack detection in tunnels using deep fully convolutional networks," *Construction and Building Materials*, vol. 234, Article ID 117367, 2020.

[19] J. Chen, C. Du, Y. Zhang, P. Han, and W. Wei, "A clustering-based coverage path planning method for autonomous heterogeneous UAVs," *IEEE Transactions on Intelligent Transportation Systems*, vol. 99, pp. 1–11, 2021.

[20] X. Zhang, Y. Yang, Z. Li, X. Ning, Y. Qin, and W. Cai, "An improved encoder-decoder network based on strip pool method applied to segmentation of farmland vacancy field," *Entropy*, vol. 23, no. 4, p. 435, 2021.

[21] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," *Advances in Neural Information Processing Systems*, vol. 25, pp. 1097–1105, 2012.

[22] 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, Golden, CO, USA, June 2015.

[23] S. Dorafshan, R. J. Thomas, and M. Maguire, "Comparison of deep convolutional neural networks and edge detectors for image-based crack detection in concrete," *Construction and Building Materials*, vol. 186, pp. 1031–1045, 2018.

[24] B. Kim and S. Cho, "Automated crack detection from large volume of concrete images using deep learning," in *Proceedings of the 7th World Conference on Structural Control and Monitoring*, pp. 22–25, Qingdao, China, July 2018.

[25] X.-W. Ye, T. Jin, and P.-Y. Chen, "Structural crack detection using deep learning-based fully convolutional networks," *Advances in Structural Engineering*, vol. 22, no. 16, pp. 3412–3419, 2019.

[26] Z. Du, J. Yang, C. Ou, and T. Zhang, "Smallholder crop area mapped with a semantic segmentation deep learning method," *Remote Sensing*, vol. 11, no. 7, p. 888, 2019.

[27] V. Hoskere, Y. Narazaki, T. Hoang, and B. Spencer, "Vision-based structural inspection using multiscale deep convolutional neural networks," 2018, https://arxiv.org/abs/1805.01055.

[28] L. Nkenyereye, B. A. Tama, and S. Lim, "A stacking-based deep neural network approach for effective network anomaly detection," *CMC-Computers Materials & Continua*, vol. 66, no. 2, pp. 2217–2227, 2021.

[29] L. Zhang, X. Wang, X. Dong, L. Sun, W. Cai, and X. Ning, "Finger vein image enhancement based on guided tri-Gaussian filters," *ASP Transactions on Pattern Recognition and Intelligent Systems*, vol. 1, no. 1, pp. 17–23, 2021.

[30] K. Khan, J. Ali, K. Ahmad et al., "3D head pose estimation through facial features and deep convolutional neural networks," *Computers, Materials & Continua*, vol. 66, no. 2, pp. 1757–1770, 2021.

[31] G. Krishna Durbhaka, B. Selvaraj, M. Mittal, T. Saba, A. Rehman, and L. Mohan Goyal, "Swarm-LSTM: condition monitoring of gearbox fault diagnosis based on hybrid LSTM deep neural network optimized by swarm intelligence algorithms," *Computers, Materials & Continua*, vol. 66, no. 2, pp. 2041–2059, 2021.

[32] Q. Liu, L. Cheng, A. L. Jia, and C. Liu, "Deep reinforcement learning for communication flow control in wireless mesh networks," *IEEE Network*, vol. 35, no. 2, pp. 112–119, 2021.

[33] W. Cai and Z. Wei, "Remote sensing image classification based on a cross-attention mechanism and graph convolution," *IEEE Geoscience and Remote Sensing Letters*, In Press, 2020.

[34] R. Liu, "Multiscale dense cross-attention mechanism with covariance pooling for hyperspectral image scene classification," *Mobile Information Systems*, vol. 2021, Article ID 9962057, 15 pages, 2021.

[35] X. Zhang, Y. Yang, Z. Li, X. Ning, Y. Qin, and W. Cai, "An improved encoder-decoder network based on strip pool method applied to segmentation of farmland vacancy field," *Entropy*, vol. 23, no. 4, p. 435, 2021.

[36] 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.

[37] Y. Tong, L. Yu, S. Li, J. Liu, H. Qin, and W. Li, "Polynomial fitting algorithm based on neural network," *ASP Transactions on Pattern Recognition and Intelligent Systems*, vol. 1, no. 1, pp. 32–39, 2021.

[38] J. Zhang, J. Sun, J. Wang, and X.-G. Yue, "Visual object tracking based on residual network and cascaded correlation filters," *Journal of Ambient Intelligence and Humanized Computing*, vol. 20, 2020.

[39] X. Ning, Y. Wang, W. Tian, L. Liu, and W. Cai, "A biomimetic covering learning method based on principle of homology continuity," *ASP Transactions on Pattern Recognition and Intelligent Systems*, vol. 1, no. 1, pp. 9–16, 2021.

[40] Z. Luo, "Application and development of electronic computers in aero engine design and manufacture," *ASP Transactions on Computers*, vol. 1, no. 1, pp. 6–11, 2021.

Hindawi
Scientific Programming
Volume 2021, Article ID 1215065, 8 pages
https://doi.org/10.1155/2021/1215065

Research Article

Trampoline Motion Decomposition Method Based on Deep Learning Image Recognition

Yushan Liu,[1] Huijuan Dong,[2] and Liang Wang[3]

[1]*Institute of Physical Education., North Minzu University, Yinchuan, Ningxia 750021, China*
[2]*Hebei Sport University, Shijiazhuang, Hebei 050021, China*
[3]*Baoding Vocational and Technical College, Baoding 071000, China*

Correspondence should be addressed to Liang Wang; wangliang201904@163.com

Received 21 April 2021; Revised 21 May 2021; Accepted 4 June 2021; Published 23 June 2021

Academic Editor: Shah Nazir

The automatic segmentation and classification of an unknown motion data stream based on given motion classes constitute an important research problem with applications in computer vision, animation, healthcare, and sports sciences. In this paper, the scenario of trampoline motions is considered, where an athlete performs a routine consisting of sequence of jumps that belong to predefined motion classes such as somersaults. The purpose of this study was to make theoretical discussions on the turning starting time and starting technique of trampoline somersault based on image recognition and point out that the appropriate turning starting time of trampoline somersault is the event when the spring net of the trampoline recovers and applies force to the human body, and the overturning start exists in the latter half of the take-off action. It is considered that how to obtain the ideal full reaction force of the net facing the human body is the flip starting technique. This work analyzes the key steps and events for trampoline somersaults and the application of artificial intelligence for the recognition of actions in the healthcare and sports fields. The effectiveness of the proposed study is shown through experimental results. The study can facilitate the process of recognition of trampoline somersault.

1. Introduction

In recent years, the integration of artificial intelligence with sports field is gradually increasing [1–3]. The use of artificial intelligence to simulate the real sports training has gradually become the key technology to improve the training level of athletes and increase the teaching efficiency of school physical education. It has become a key research topic in sports training and teaching to correct the wrong movements and determine the correct movements [4].

There are great differences between computer image recognition technology and human image recognition. Human recognition is to compare and analyze different images by naked eyes and obtain the corresponding information elements directly. Moreover, the number of images that can be recognized is relatively small, and with the continuous extension of recognition working time, the accuracy of recognition will be affected [5–8]. However, image recognition technology can solve this problem well, because it will first store a large number of images and then obtain the most intuitive data information through the arrangement and comparative analysis. When people need to identify the results, they can also directly obtain reference. At the same time, the process of image recognition by naked eyes is often very short and cannot save image information for long time. However, image recognition technology has a very long processing flow and can save data for a long time. When the relevant personnel need the image data information, they can obtain it directly from the computer system [9–11].

Worldwide, the training of the trampoline is still in the stage of exploration and experience accumulation. At this stage, according to the performance characteristics of the trampoline and the requirements of competition rules, this paper examines the key link of trampoline somersault, the turning start time and its action technology, and finds out

the corresponding theoretical basis. It has positive practical significance for the athletes to train the trampoline somersault and improve the technical level of the action. Based on the analysis of the image data of the world's elite athletes, performance characteristics, and relevant provisions of the trampoline competition rules, combined with the research results of gymnastics and acrobatic somersault, we theoretically discussed the turning start time and action technology of trampoline somersault. The proposed research can be helpful for the future training of trampoline somersault and development of trampoline. However, the application of artificial intelligence especially computer vision in the field of sports action recognition is more in the experimental state, lacking systematic and clear technical guidance. The innovations of the proposed work are as follows:

(i) To construct a perfect application framework in the field of sports action correction.

(ii) To analyze the key steps, and realization ways of artificial intelligence simulation in the sports action correction field in detail.

(iii) To provide technical guidance for the extensive development of artificial intelligence technology in the field of physical education teaching and sports training.

2. Application of Artificial Intelligence and Virtual Reality Technology in Trampoline Action Decomposition Image Recognition

Using artificial intelligence-based systems, users can restore the vision, hearing, touch, force, and motion perception in the virtual environment to the greatest extent. The integration of artificial intelligence technology and modern physical education teaching can visualize the three-dimensional dynamic information observed by computer, which is conducive to the observer and athletes to objectively evaluate the sports action and, finally, optimize and perfect it. After fusion, it can help these sports effectively avoid sports risks, reduce training costs, and eliminate the impact of the weather. In a word, the emergence of artificial intelligence technology can break the limitation of traditional sports on movement correction, simulation training, and other hardware conditions. Moreover, it can optimize the sports and training experience of sports participants at the lowest cost.

At present, there are some researches on the integration of artificial intelligence technology and the sports field. Ma et al. combined the multiperception, interactivity, immersion, and other characteristics of artificial intelligence with some activities in physical education teaching and summarized the application strategies of sports action comparison and remote interaction technology [12, 13]; Xi analyzed the convenience of artificial intelligence for sports teaching field and summarized the defects and improvement schemes of artificial intelligence in the field of sports teaching [14]; Wu further conceived the application methods of artificial intelligence in specific sports such as football, diving, and trampoline [15]. However, these studies mainly focus on the macro direction and strategy of the application of artificial intelligence and physical education teaching. In addition, they do not conduct in-depth excavation and research on the specific application framework and technology.

2.1. Characteristics of Trampoline Somersault under the Requirements of Rules. Trampoline is a recreational and sport exercise device. Compared with gymnastics and acrobatic somersault, the spring coefficient of trampoline somersault is far greater than that of the tumbling board. Therefore, the time of jumping in the air is longer and the turning space is larger. It is much easier to complete the same difficult somersault on a trampoline than on gymnastics and acrobatic flip board. However, with the development of trampoline, the long time and large space of tumbling have brought trampoline athletes not only plentiful flipping space-time conditions but also the challenge of how to complete more difficult movements. In the competition rules, there are more difficult points in the final score of a set of movements. It is an indisputable fact that the movement is characterized by high difficulty and complex technique.

2.2. Characteristics of Artificial Intelligence Technology. Thanks to the rapid development of modern sensor technology, the multiperception of artificial intelligence technology has gradually developed from the original vision and jerk perception to a comprehensive perception system with multiple senses such as vision, hearing, touch, movement, and force and can provide users with more comprehensive and real perception environment. Similarly, artificial intelligence enables users to simulate different application scenarios through professional equipment according to their own needs, so that users can communicate, contact, and even create simulation scenarios more naturally and truly. Artificial intelligence can simulate and develop scenarios that are difficult or impossible to achieve in real-life applications. Various scenes can fully activate the human brain's imagination of the outside world and expand the imagination space of human beings. Artificial intelligence equipment can simulate the real scene of the outside world to the greatest extent and help users to truly perceive the virtual world and get familiarity with the environment [16].

2.2.1. Artificial Intelligence-Based Image Recognition Mechanism. Image recognition is the capability of a computer system to identify people, places, objects, and actions in images. It uses computer vision techniques with artificial intelligence and machine learning algorithms to recognize images through a camera system. While human and other animal brains recognize actions and objects with comfort, computers-based image recognition requires machine learning algorithms such as deep learning algorithms. Deep learning algorithms can function by use of comparative 3D models, appearances from different angles using edge detection, or components. Image recognition algorithms are

often trained on millions of prelabeled pictures with guided computer learning. Image recognition works by detecting salient sections called features, which are the areas inside the image that contain the most distinct information about the object or image. The most informative features in an image are selected and localized, while the less informative points are ignored. For this purpose, a classification algorithm is used that takes an image as input and outputs what the image contains. Mostly supervised learning algorithms are used that are trained to learn the differences between image classes. For example, if the goal of an image recognition system is to detect a cat, the image recognition algorithm needs to be trained with hundreds of images of cats and hundreds of images of backgrounds that do not contain any cats [17–19].

2.2.2. Analysis of the Technical Principle of Image Recognition Technology in Artificial Intelligence. In artificial intelligence, the technical principle of image recognition technology is to use a computer to process pictures and then extract the discriminant features of the pictures for further processing. Image recognition technology is not complex in the technical principle. If a person is regarded as a computer, so every time people look at each other can be said to have completed an image recognition process, and then people, according to the impression of others in their brain, analyze, search, and memorize other people's information. Computer image recognition process is using the same principle. The only difference is that, in artificial intelligence-based image recognition technology, the image on the computer provides information dependent on the number. According to the principle of image recognition technology, artificial intelligence is combined with a computer algorithm, so that the information of the image can be extracted and analyzed. In the image with large amount of information, the recognition efficiency of image recognition technology may be reduced due to redundant information. Therefore, while analyzing images on computer, we should also find a better, simpler, and more convenient information extraction method to make image recognition simpler and faster.

2.2.3. Analysis of the Composition Principle of Image Recognition Technology in Artificial Intelligence. Pattern recognition is mainly used in image processing technology to process different types of plane pictures and three-dimensional pictures, to analyze pictures and objects. Image recognition technology is mainly based on pattern recognition and development of three-dimensional images. In the field of modern medicine and architecture, the use of this technology is more prominent such as to analyze the three-dimensional structure of the human body and examine the health status of patients. In the field of architecture, the traditional building structure is mainly based on the plan. With the development of modern science and technology, architects can use pattern recognition to make the architectural plan stereoscopic and make the architectural image clearer and more understandable. At the same time, it meets the needs of people not only in the application of image recognition but also in the application of image recognition technology. In our daily life, people's eyes are the best pattern recognizer, so the ordinary pattern recognition technology is doomed to be unable to meet the needs of people's daily life, but after combining with artificial intelligence, pattern recognition has achieved better performance. For example, the analysis and imaging of complex stereo images, artificial intelligence, and pattern recognition can help us process such images more conveniently and efficiently.

2.2.4. Research on the Process of Image Recognition Technology in Artificial Intelligence. Human brain image processing is similar to computer image processing. In the computer image processing, the information is obtained first and then converted into a computer readable form and the relevant programs and tools are used to process information. The second step is to store the image. After the image is stored, we can use the image for simple processing, sharpen the important information of the image, and make the image information more prominent. The third step is to extract the relevant information of the image and further process the image according to the keywords and needs of people. On the computer, there are many software programs and methods for processing picture information. People can choose the tools and methods according to their own needs, and at the same time, they can also choose the region and characteristics of the picture for processing. In the development of the computer, the most prominent image processing should be in photoshop (PS), which is a relatively fast and stable software for image processing and also has a very large development prospect. In the PS image processing, you can operate the image arbitrarily as long as you select the processing area, and you can carry out relevant operations on the image. Then the processing process and method of image processing technology in artificial intelligence are very similar to those of the computer processing process. The image processing technology in artificial intelligence also needs to have pictures before the subsequent work of pictures can be carried out. The difference is that the image processing technology in artificial intelligence is more intelligent than the computer processing method. For the technical realization, the image processing technology in artificial intelligence is more convenient, and the operation is more convenient and simpler. It is believed that, in the later stage of the development of science and technology, the image processing technology in artificial intelligence will develop into a processing technology which is more suitable for public use.

Technology has changed the complexity of computer image processing and can also achieve the same versatility, which is very convenient for our daily application [20, 21]. Therefore, a comprehensive analysis of the image processing technology in artificial intelligence [22–26] will be better recognition of technology after the traditional image processing technology [27–32].

3. Simulation Experiment and Result Analysis

After getting the 3D reconstruction model provided by the system, the participants can directly see the virtual policy results of their actions. Through communication with the coaches and the comparison of excellent athletes' action completion, this paper analyzes the unreasonable points in their movements, to correct them in future training. Besides, using the established and perfect system to expand the function, an automatic comparison interface can be established to automatically compare the movement changes in the process of a training cycle and the action comparison between itself and other excellent athletes and assist manual analysis of action essentials. First of all, we can collect the technical action images of elite athletes in related fields for three-dimensional reconstruction, to obtain the complete three-dimensional pose data of such athletes. Then, according to the three-dimensional posture data extraction technology provided in this paper, the effective characteristic information of three-dimensional posture data of excellent athletes is quantified. Thirdly, the information obtained from mining is taken as the system standard. In the last, the standard is established. The quasi model database is used to enrich the standard action model in the system database. In this technical condition, the more the number of excellent athletes' technical movements entered is, the more it can help the later users learn from more excellent movements. After comparing with the actions in the database, the system will score the user's action scores according to the scoring system in the competition, so that the users can obtain the evaluation results more directly.

Firstly, the preprocessed image is transformed by scale-invariant feature transformation, and the feature points whose position, scale, and rotation are invariant are detected by the operator. Secondly, the feature points of the detected image are matched by nearest neighbor matching algorithm, and the matching measure is the Euclidean distance between the feature points. Thirdly, to eliminate the error matching caused by the nearest matching two secondly, RANSAC algorithm is used to enhance the consistency of random samples. Fourthly, SFM method is used to filter the data from two-dimensional information to obtain the three-dimensional point cloud structure. Then the matching point pairs between multiple images are used to estimate the parameters of the topological camera and obtain the final three-dimensional structure.

For the application of virtual sports scene simulation, the system application focuses on the interaction between the virtual scene and the real scene of athletes. The immersion of the virtual scene is the key to determine whether the participants can get a sense of reality. Therefore, the application of the virtual scene simulation direction mainly lies in the rendering of 3D reconstruction. Therefore, designers should strengthen the comprehensive analysis and quantification of a certain kind of motion and establish a simulation physical engine based on a huge database. In the process of using the system, the 3D reconstruction technology is used to accurately obtain the technical action data of the user and input it into the simulation physical engine of the system, and the system will provide feedback on the action of the moving personnel according to the set program.

A complete set of trampoline movements is characterized by its movement and floating upward. The rhythmic connection and transformation between the athletes' movements include two feet take-off, back sigh, sitting, and bouncing. As shown in Figure 1, there will be no long pause and intermediate jump in the middle of the whole set of movements. All kinds of forwarding and backward somersaults and turns as well as some nonrotation movements of athletes coordinate with each other to form a complete trampoline action. Team trampoline events require athletes to show a complete set of movements to show the beauty of body posture and the accuracy of technical movements, including the ideal height and control of the body. In a set of actions, not only can there be repeated actions, but also they cannot be interrupted. If there is an interruption, the action after the interruption will not be given points. In the actual competition, the score of trampoline action is mainly divided into skill score, difficulty score, and synchronous score. Among them, the skill score is mainly determined according to the athlete's action completion, while the difficulty score is mainly calculated by the number of turnover cycles of the movement, and the synchronous score is the score of the synchronization of the routine performed by two athletes. As a result, gymnastics trampoline has become a special sports competition, and the development of a competitive sports event on the field is an inevitable result.

The trampoline set consists of 10 single movements. A complete set of technical links should be composed of preparation posture, prejump, last prejump, take-off, completion of overturning action, falling, net touching rebound jump (at the same time, it is the take-off of the second action). In the same way, when the last action is completed, it includes the following links: falling, net touching rebound, jumping, buffer stops, etc., as shown in Figure 2.

From this, we can see that there is an overlapping link in the technical link between each action, that is, the touch net rebound jump link. From the technical aspects of a single action, the task of net contact rebound jump is to buffer the rebound strength. But from the connection movement technical link, this link's task becomes through the lower limb muscle, and the joint's strength flexes and stretches, causes the net surface to produce the maximum distortion, and thus obtains the ideal take-off effect. Generally speaking, it is the end of the previous action and the beginning of the next action. This overlapping part is the main feature and key part of trampoline technique. From this point of view, in the practice of a complete set of movements, we must grasp every link of net contact rebound, so that the buffer of every passive net touch rebound becomes an active and powerful take-off, which is the key problem of an active connection. In the transition from single-action practice to complete set of actions practice, many athletes are often thrown "flying around" by rebound force because they do not grasp this link well and they do not take-off actively. This has a great relationship with the fact that the athletes do not timely change the passive buffer of net rebound take-off into the main dynamic take-off. To solve this problem, the best way is to

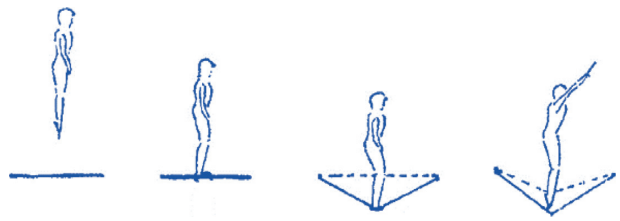


Figure 1: Dynamic work diagram of trampoline.

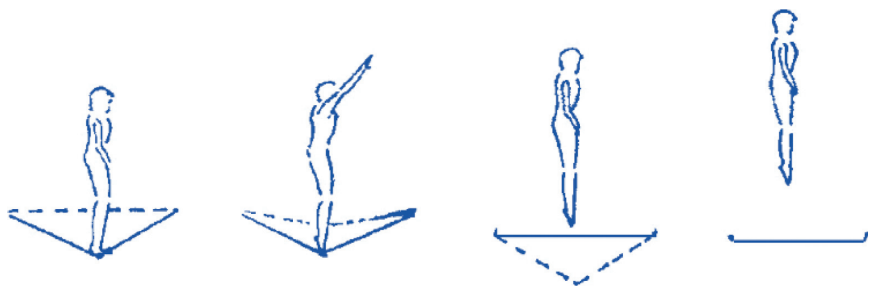


Figure 2: Continuous prejump of trampoline.

regard the net touch rebound jump as a part of the actual action in the single-action training. Only when the action is completed can the touch net rebound link be effectively turned into a powerful take-off again. On this basis, it will be easier to practice in sections, that is, to divide the 10 movements in the set into several small segments and practice them separately. Finally, it will be much easier to carry out the whole set of exercises, as shown in Figure 3.

Besides, through the analysis, we also found that there is an inevitable internal relationship between the movements. We can observe the interaction between these movements in training. To sum up, there are two forms of this connection and influence: one is that when a certain action moves forward and backward (longitudinally), it will affect the air trajectory and landing point of the next action, thus affecting the stability of the action. Second, the left and right (lateral) displacement of a certain action will directly affect the stability of the next action. Sometimes the displacement may not be obvious and can be easily ignored, but it has a great influence on the next action, which is enough to cause the failure of the action. This kind of connection and influence is often seen in the usual training.

The reason for the forward and backward displacement is that the athlete is eager to do the action, resulting in forward and backward movement of upper body during taking off. The cause of the left and right displacement may be due to the strength differences between the legs when taking off and the length difference of the legs. Therefore, special attention should be paid to this connection in training. If we only look for the reason from the obvious wrong action itself and do not consider the influence of the previous action, the effect will be very unsatisfactory, such as forward somersault 2 weeks, twist 180 degrees, and somersault 2 weeks. The athlete has a slight forward displacement when doing the front somersault, but the backward somersault obviously moves backward and even falls on the

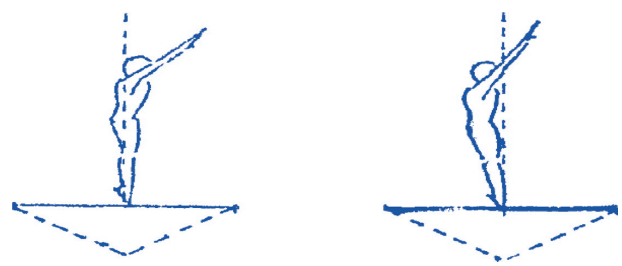


Figure 3: Schematic diagram of the last jump of trampoline taking off and doing a somersault.

safety platform. The problem of two weeks of front somersault has appeared. At this time, if the coach only emphasizes the backward somersault and does not consider the internal relationship between the movement connections, the improvement effect will be very unsatisfactory. If the front somersault is improved, try to make the front somersault not having a slight forward displacement, and the problem of backward somersault moving back can be easily solved. The influence of left and right displacement is also the same.

3.1. Decomposition and Recognition Process of Trampoline Action Based on Image. The combined morphological operation can eliminate part of the background in the video image, and the morphological features can be well preserved to obtain the silhouette image of the human body, which is very similar to the background subtraction technology. Then the operation formula of combined morphology can be expressed as follows:

$$G(x, y) = F(x, y) \cdot B(x, y) - F(x, y), \tag{1}$$

where $G(x, Y_1)$ represents the image processed by the combined morphological operation, $f(x, Y_1)$ represents a

frame image in the original video, and $G(x, Y_1)$ represents structural elements. Through (1), the area in the original image that is darker than the background and smaller than the size of the result element can be removed. The appropriate structural elements can be selected and the remaining background image can be obtained through the closing operation so that the target extraction can be completed by subtracting it with the original image.

In the video image of human trampoline action, one frame of the image cannot fully describe an action. Generally, it is necessary to extract multiple image features to display human action completely. Due to the difference of action rate, the frame number of each video image may be different even for the same action. To deal with these two kinds of rate changes, considering the characteristics of rate changes, the gray features of each edge image in the same time window are accumulated into the same image, and the features are extracted by using the accumulated edge image to represent the human trampoline action.

The operation flow of cumulative edge image is as follows: a frame of video image processed by morphological gradient operation is represented by $C(x, Y_1)$, the edge detection image obtained by using edge detection operator on $C(x, Y_1)$ is represented by $E(x, Y_1)$, which is a binary image, the edge image obtained by multiplying $C(x, Y_1)$ and $E(x, Y_1)$ on each pixel is $I(x, Y_1)$, and the gray image is represented by $I(x, Y_1)$. It means that for the gray information s on the edge points if the pixels are outside the edge, the gray value is 0; the cumulative edge image is represented by HF, y, 1, and its scale is consistent with the size of $C(x, y, t)$ and inning purpose of obtaining $H(x, y, t)$ is to accumulate all $I(x, Y_1)$ in a video image on a certain time window.

For $H\ (x, y, T_1)$ initialization, all pixels are set to 0, and the time condition is $t = 0$; based on edge detection, edge image $e(x, y)$ can be obtained on the first frame of the morphological gradient image $g(x, Y_1)$ of the video time window; the gray image $I(x, y)$ is obtained by multiplying gradient image $g\ (x, y)$ and edge image $e(x, y)$; cumulative previous $I(x, Y_1)$ and previous frame are compared on all pixels edge image $H(x, y, T_1)$; the gray value of the pixel with larger gray value will be taken as the new value of $H(x, y, T_1)$ action at the edge detection step until the whole image is complete.

The image operation is completed.

The main idea of this method is to compress the information in the video sequence into a frame image to express the action. The information contained in the accumulated edge image is huge. The formula of accumulating edge image at the point FX, Y_1 can be described as follows:

$$\begin{aligned} I(x, y) &= G(x, y)E(x, y), \\ H(x, y, t) &= \max(H(x, y, t-1), I(x, y)). \end{aligned} \tag{2}$$

The cumulative edge image is to multiply the binary tattoo $e(x, Y_1)$ and the morphological gradient image $c(x, Y_1)$ at each pixel to obtain the edge image $I(x, Y_1)$ with gray information and then accumulate all the edge images into one image, not every binary image into one image. 0 and 1 are the only two gray values of the binary image $e(x, Y_1)$. If the pixel value of the binary image $e(x, y)$corresponding to the edge image $I(x, y)$ is 1, then the gray value range at this point is more than that of the binary image.

If the edge image is accumulated for the target image, the image information center already contains more edge information of the image. So it is not necessary to extract the edge features, and the directional gradient histogram can be directly solved at each point of the accumulated edge image.

The calculation of the grid-based directional gradient histogram is to solve the directional gradient of all points in the cumulative edge image. The cumulative edge image is divided into Ie spatial grids, and the histogram vector on each grid is calculated. One of the scale feature vectors is extracted and used as the action feature. The local shape of the target is counted, and then the feature vector of the cumulative edge image is obtained.

3.2. Trampoline Movement Recognition Based on Dynamic Time Warping. Action expression has temporal persistence; that is, action can be a set of static actions in a certain period. The movement process of the human body can reflect the changing trend of movement through the change of joint angle curve, which can be called the joint angle time series. The motion characteristics can be defined as follows:

$$\text{Action\_feature} = \{A_1, A_2, \ldots, A_M\}, \tag{3}$$

where the time series of a certain joint angle is represented by line vector a; the row vector with the number of motion features is represented by a, and the range is $1 \le m \le 16I$. If the time-varying one-dimensional signal can be understood, it evolves into the classification problem of time-varying feature data by simple action recognition. From the past data, we can see that when the examinee freely displays martial arts motions, the same movement, it is hard to rule out the potential of being similar to a because of distinct waveforms and amplitudes Therefore, action recognition is realized by comparing the similarity of time series; that is, Wushu action decomposition judgment is realized by comparing the distance between different length vectors.

The comparison of similarity between curves is the focus of the time series change trend. Because there are uncertainties in the video feedback system and testers, which will lead to data deviation and fluctuation, the following formula is used to smooth the sequence:

$$x_i = \frac{x_1 + x_2 + \cdots + x_n + x_{n+1}}{n}. \tag{4}$$

RT is an integer greater than 0. To sum up, this paper extracts the feature vector of martial arts action features in video images by accumulating edge images and then calculates the time series of martial arts actions by using the

dynamic time warping theory. After the martial arts actions to be identified are matched with the reference time series samples, the process of Wushu action decomposition and recognition is completed.

4. Conclusion

The assimilation of artificial intelligence in the field of sports is progressively increasing. The use of AI to simulate the real sports training has regularly become the key technology to improve the training level of athletes and increase the teaching efficiency of school physical education. Human action recognition technology is widely used in healthcare and sports fields. This work proposes an effective recognition method based on image recognition for trampoline decomposition, using morphological gradient operation based on the cumulative edge image. The contour edge image features are extracted and accumulated into an image. The directional gradient histogram of the image is calculated to obtain the image feature vector. The dynamic time warping method is used to calculate the joint action change sequence of different martial arts actions to form a reference sample set. Finally, the test sample is compared with the reference sample to realize the image recognition of the trampoline decomposition action analysis. It is concluded that the proposed method will be beneficial for analyzing and improving the trampoline somersaults of sports individuals. Results of the experiments reveal the effectiveness of the study.

Data Availability

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

Conflicts of Interest

The authors declare no conflicts of interest.

Acknowledgments

This work was supported by "the Fundamental Research Funds for the Central Universities", North Minzu University (2020XYSTY04).

References

[1] L. Rong and Y. Xu, "Research on image feature extraction algorithm based on visual information," *ElectronicDesign Engineering*, vol. 24, no. 9, pp. 188–190, 2016.

[2] L. Xu and D. Liu, "Visual image recognition simulation of bus dangerous action," *Computer Simulation*, vol. 32, no. 6, pp. 150–153, 2015.

[3] L. Fei, *Research on Human Motion Recognition Method in RGB-D Sequence Images*, Shenyang University of Aeronautics and Astronautics, Shenyang, China, 2017.

[4] Y. Liu, *Human Motion Recognition is Based on 3D Skeleton Fragment Representation Metric and Manifold Segmentation*, University of Science and Technology of China, Hefei, China, 2017.

[5] Li Yang, *Motion Recognition and Application Research Based on Depth Image*, National University of Defense Science and Technology, Changsha, China, 2016.

[6] J. Yang, *Human Continuous Motion Recognition Based on LSTM*, Xi'an University of Technology, Xi'an, China, 2020.

[7] L. Gu, *Research on Human Motion Recognition Based on Skeleton Information*, Inner Mongolia University of Science and Technology, Baotou, China, 2020.

[8] F. Lin, *3D Human Skeleton Motion Recognition Based on Joint Local Disturbance Learning*, University of Science and Technology of China, Hefei, China, 2020.

[9] Z. Wang, *Research on Human Motion Recognition Based on Spatiotemporal Convolution Neural Network*, University of Electronic Science and Technology, Chengdu, China, 2020.

[10] L. Shi, *Deep Motion Recognition in Natural Human-Computer Interaction for Web3D Engine*, University of Electronic Science and Technology, Chengdu, China, 2020.

[11] G. Y. Shen, *Research on 3D Motion Recognition and Interaction Model*, University of Electronic Science and Technology, Chengdu, China, 2020.

[12] Y. Bai, "Research on student action recognition method based on video," *Instrument Users*, vol. 27, no. 1, pp. 10–12, 2020.

[13] Y. Ma, Q. Han, R. Liang, J. Hou, and Y. Wang, "Research on the identification and evaluation of rhythmic gymnastics movements based on deep learning," in *Proceedings of the Abstracts of the 11th National Congress of Sports Sciences*, Chinese Society of Sports Sciences, Nanjing, China, 2019.

[14] Xi long, *Research and Application of Human Motion Recognition Based on Deep Learning*, Xi'an University of Technology, Xi'an, China, 2019.

[15] Y. Wu, *Online Human Motion Analysis Based on Deep Learning*, Jinan University, Guangzhou, China, 2019.

[16] M. Lei, *Research on Human Motion Recognition and Interaction Technology in a Virtual Environment*, North University of Technology, Beijing, China, 2019.

[17] L. Jin, *Research on Motion Capture Algorithm Based on the 3D Static Model*, North University of Technology, Beijing, China, 2019.

[18] W. Ye, *Research and Application of Motion Detection Algorithm in Motion Video Technology*, North University of Technology, Beijing, China, 2019.

[19] L. Rui, *Motion Recognition and Gesture Estimation in Image and Depth Map*, Zhejiang University, Hangzhou, China, 2019.

[20] Z. Luo, "Application and development of electronic computers in aero engine design and manufacture," *ASP Transactions on Computers*, vol. 1, no. 1, pp. 6–11, 2021.

[21] L. Liang, Q. Yin, and C. Shi, "Exploring proper names online and its application in English teaching in university," *ASP Transactions on Computers*, vol. 1, no. 1, pp. 24–29, 2021.

[22] R. Liu, X. Ning, W. Cai, and G. Li, "Multiscale dense cross-attention mechanism with covariance pooling for hyperspectral image scene classification," *Mobile Information Systems*, vol. 2021, Article ID 9962057, 15 pages, 2021.

[23] Y. Tong, L. Yu, S. Li, J. Liu, H. Qin, and W. Li, "Polynomial fitting algorithm based on neural network," *ASP Transactions on Pattern Recognition and Intelligent Systems*, vol. 1, no. 1, pp. 32–39, 2021.

[24] X. Zhang, Y. Yang, Z. Li, X. Ning, Y. Qin, and W. Cai, "An improved encoder-decoder network based on strip pool method applied to segmentation of farmland vacancy field," *Entropy*, vol. 23, no. 4, p. 435, 2021.

[25] X. Ning, F. Nan, S. Xu, L. Yu, and L. Zhang, "Multi-view frontal face image generation: a survey," *Concurrency and*

Computation: Practice and Experience, 2020, In press, Article ID e6147.

[26] Q. Liu, L. Cheng, A. L. Jia, and C. Liu, "Deep reinforcement learning for communication flow control in wireless mesh networks," *IEEE Network*, vol. 35, no. 2, pp. 112–119, 2021.

[27] X. Ji, *Research on Human Motion Recognition Method Based on Depth Image Sequence*, University of Chinese Academy of Sciences, Shenzhen Institute of Advanced Technology, Chinese Academy of Sciences, Beijing, China, 2018.

[28] Li Fei, *Human Motion Recognition Based on Feature Fusion*, Xi'an University of Technology, Xi'an, China, 2018.

[29] Y. Jiang, X. Gu, D. Wu et al., "A novel negative-transfer-resistant fuzzy clustering model with a shared cross-domain transfer latent space and its application to brain CT image segmentation," *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, vol. 18, no. 1, pp. 40–52, 2020.

[30] W. Cai, Z. Wei, R. Liu, Y. Zhuang, Y. Wang, and X. Ning, "Remote sensing image recognition based on multi-attention residual fusion networks," *ASP Transactions on Pattern Recognition and Intelligent Systems*, vol. 1, no. 1, pp. 1–8, 2021.

[31] J. Zhang, J. Sun, J. Wang, and X.-G. Yue, "Visual object tracking based on residual network and cascaded correlation filters," *Journal of Ambient Intelligence and Humanized Computing*, 2020, In press.

[32] L. Zhang, X. Wang, X. Dong, L. Sun, W. Cai, and X. Ning, "Finger vein image enhancement based on guided Tri-Gaussian filters," *ASP Transactions on Pattern Recognition and Intelligent Systems*, vol. 1, no. 1, pp. 17–23, 2021.

Hindawi
Scientific Programming
Volume 2021, Article ID 4102294, 11 pages
https://doi.org/10.1155/2021/4102294

Research Article

Football Player Posture Detection Method Combining Foreground Detection and Neural Networks

Xin Hu

School of Football, Xi'an Physical Education University, Xi'an 710068, Shaanxi, China

Correspondence should be addressed to Xin Hu; huxinsport@163.com

Received 29 April 2021; Revised 30 May 2021; Accepted 5 June 2021; Published 22 June 2021

Academic Editor: Shah Nazir

In recent years, with the rapid development of artificial intelligence, information technology, intelligent digital video surveillance systems, real-time sports competition playback, and other technologies have emerged one after another, making the advantages of deep learning-based football posture detection tasks become more obvious. Related models and methods have been applied to the research field of sports posture estimation and have achieved great improvement, surpassing the traditional football posture estimation method based on manual design features in one fell swoop. In addition, the application of video foreground detection has developed rapidly and has great application value in sports analysis. Therefore, this paper proposes a novel football motion detection approach combining foreground detection and deep learning for real-time detection of football player posture. The main task of foreground target detection is to extract the interesting foreground target in the real monitoring scene and use it as the target of interest for subsequent analysis. Then, we propose a triple DetectNet detection framework based on deep learning technology, which can quickly and robustly realize the three-dimensional pose estimation of multiperson motion. For input, the triple DetectNet framework uses three neural networks and is executed in three stages; the first stage is to use the DetectNet (DN) network to detect the bounding box of each person separately, the second stage uses the 2DPoseNet (2DPN) network to estimate each of the corresponding two-dimensional poses of the individual, and the third stage uses the 3DPoseNet (3DPN) network to obtain the 3D pose of the person. This paper also conducted experiments on four datasets, and the results proved the superiority and success of this algorithm.

1. Introduction

In recent years, with the rapid development of artificial intelligence technology, the emergence of intelligent digital video surveillance systems, real-time sports competition playback, and other technologies have made the advantages of computer vision tasks based on deep learning more and more obvious, and it has high flexibility and openness, which indicates the development direction of intelligent image processing. With the development of deep learning, related models and methods have been applied to the field of football player posture estimation and have achieved great improvement, surpassing traditional posture estimation methods based on hand-designed features in one fell swoop. The football player posture estimation based on the deep learning method has made breakthroughs in various aspects. The problem of moving target detection and tracking based on deep learning has developed in the development of science and technology and engineering applications and has a certain research and application foundation in the fields of football competition, football training, and artificial intelligence research. As far as video playback and monitoring are concerned, if we use deep learning-based related motion detection and tracking technology, we can perfectly assist the competitive team or coach team in completing the task. Since 2009, Hinton et al. [1] published important research work on deep belief networks; deep learning [2–7] has become a new direction of machine learning, and it has been used in the handling of many problems in the field of artificial intelligence. So far, deep learning-related frameworks have developed into a variety of targeted methods, such as convolutional neural networks, deep neural networks,

recurrent neural networks, and deep belief networks, which have been successfully applied in various fields of computer and obtained excellent results. As a result, artificial intelligence has been fully developed.

Football player posture estimation is a basic problem in computer vision. It is the basis of multiperson posture estimation, behavior recognition, and sports goal analysis. It can be widely used in many meaningful fields, such as sports competition and behavior analysis. The goal of soccer player single pose estimation is to find the coordinates of each joint point of the player from the image or video containing a single soccer player. Affected by the shooting angle, scene, lighting, wearing, etc., the estimation of the single athlete's posture in the image is facing arduous challenges. However, with the opening of pose estimation data sets and the development of computer hardware and deep learning technology [8–12], various networks such as convolutional neural networks based on deep learning have gradually penetrated into various research fields of computer vision, and single athlete pose estimation has also been achieved. Since the methods used for pose estimation are mainly based on deep convolutional neural networks, a small number of methods are based on generative adversarial networks. At the same time, based on single-person pose estimation, we have developed multiperson pose estimation, which is accurate from the input RGB video, detects the target person, predicts the 2D key points, and finally predicts the accurate 3D key points through our network to get the final multiperson pose. Based on the existing hardware, we can fully realize real-time multi-motion target analysis.

In addition, the application of video foreground detection is also very extensive. The feature is that in competitive sports, its complex dynamic scenes, color camouflage, illumination changes, and static foreground have brought various difficulties and problems to video foreground detection. Foreground target detection algorithms are mainly divided into three categories: detection algorithms based on target modeling, foreground detection algorithms based on background modeling, and detection algorithms combined with deep learning. Arghavan et al. [13] optimized, improved, and, based on the motion information, extracted the feature points corresponding to the moving target from the next frame. Then, the number of moving objects in each frame is determined according to the motion information and location, and then the k-means algorithm is used for clustering. The moving target is clustered using the feature vector composed of pixel intensity, motion amplitude, motion direction, and feature point position. This algorithm has high accuracy in determining the number of moving targets, but it cannot detect contour information well. Therefore, detection algorithms based on target modeling are more inclined to detect speed and can detect the position of moving targets, but they cannot effectively detect target contours and easily lose information. The basic principle of the detection algorithm based on background modeling is to compare the current frame image information with the established background model and extract the difference area as the foreground target. But with the development of artificial intelligence, it is a new research idea to integrate deep learning ideas with foreground detection algorithms.

Based on the above observations, this paper proposes a novel soccer motion detection method combining foreground detection and deep learning for real-time detection of football players' posture. The main task of foreground target detection is to extract the interesting foreground targets in the real monitoring scene and use them as interesting targets for subsequent analysis. Then, we proposed a triple DetectNet detection framework based on deep learning technology, which can quickly and robustly realize the three-dimensional pose estimation of multiperson motion. For input, the triple DetectNet framework uses three neural networks and is executed in three stages: the first stage is to use the DetectNet (DN) network to detect the bounding box of each person separately; the second stage uses the 2DPoseNet (2DPN) network to estimate each the corresponding two-dimensional pose of the individual and the third stage uses the 3DPoseNet (3DPN) network to obtain people's 3D pose. Following are the main contributions points of this paper:

(i) To propose a foreground detection method based on confidence weighted fusion and visual attention, which is used to solve the problem of color camouflage and static foreground in the background subtraction method.

(ii) To use target design based on DetectNet detection selects the corresponding 2DPoseNet (2DPN) network framework for two-dimensional pose prediction.

(iii) To present a novel triple DetectNet detection framework to achieve the prediction from two-dimensional pose to three-dimensional pose and obtain better results than existing advanced methods in training.

(iv) To conduct experiments on three data sets, and the results proved the superiority of this algorithm.

The paper is structured as follows: section 2 represents the related work to the proposed research. The research methodology of the proposed study is given in Section 3, with details of the approach used. The experiments and results are given in Section 4. The conclusion of the paper is presented in Section 5.

2. Related Work

Target detection and tracking based on deep learning: first is the problem of target detection, that is, determining the position of the target object in the image or scene, which is generally determined by the bounding box of the object. In response to such problems, Sande et al. [14] proposed that RCNN uses a selection domain method to obtain local candidate regions that may have detection targets in the image, and then input these candidate regions into the convolutional neural network to obtain their features and connect the classifiers Go to the feature map to determine

whether the corresponding area belongs to the target to be detected, and finally perform regression on the calibration frame to correct the position of the prediction frame, but RCNN has the problem of repeated calculation. He et al. [15] introduced the spatial pyramid pooling layer into CNN and proposed SPPnet, which reduced the CNN network's limitation on the input image size and improved the accuracy. Based on the idea of SPPnet, Girshick et al. [16] also proposed a one-stage method Fast-RCNN, based on an adaptive pooling method, mapping the final candidate region to the feature map of the last convolutional layer of the convolutional network. The benefits are obvious. Only one feature extraction is required to perform the detection task, which greatly improves the detection speed. However, Fast-RCNN also has this defect; that is, it takes a long time to extract feature candidate frames; Zhou et al. [17] propose a new detection method whose performance is greatly improved over the traditional one-stage and two-stage frameworks, especially based on real-time conditions; at the same rate, its accuracy should be based on deep learning motion. Human target recognition and pose estimation are much higher than YOLOV3 [18], and compared to YOLOV3, it can recognize small objects significantly better. From the above introduction, we know that the two-stage detector first detects all potential object positions (i.e., candidate areas) for each category to judge, and this process consumes a lot of time and storage space, which means it is not applicable for real-time object detection.

Foreground target detection mainly serves advanced video analysis technology in intelligent video surveillance. The main task is to extract interesting foreground targets in real surveillance scenes. Beaugendre et al. [19] proposed a random block background modeling (RBBM) algorithm. The update of the background model is divided into multiple random blocks evenly distributed on the image and in time so as to quickly update the video background image. This makes it possible to save a lot of calculation time when processing high-definition video images while sacrificing certain detection accuracy requirements. Subsequently, Beaugendre et al. [20] proposed a mixed block background modeling (MBBM) algorithm based on the spatiotemporal update. The background model is updated by carefully selecting blocks in linear and pseudorandom order and updating the block part of the corresponding model. The two-block selection sequence ensures that each block will be updated. This algorithm is combined with the adaptive block propagation background subtraction method (ABPBGS) [21] for foreground detection. Savas et al. [22] proposed an algorithm based on block matching to estimate the motion in each frame and generated the corresponding motion field to serve the foreground target detection, which showed a better detection effect in the crowded situation. In order to overcome the shortcomings of convolutional neural networks in foreground detection, Dimitrios et al. [23] proposed an end-to-end 3D convolutional neural network model and used it for foreground target detection. The model can track changes in time, so there is no need to retain and update the model.

3. Methodology

3.1. Confidence Weighted Fusion and Visual Attention. For the foreground processing of football players' competitive video, this paper proposes a foreground detection method based on confidence weighted fusion and visual attention to solve the problem of color camouflage and static foreground in the background subtraction method. The algorithm first builds a model through color features, LBSP texture features, and the corresponding confidence, and then calculates the sum of the confidence of valid samples for weighted fusion for pixel classification and builds a visual attention mechanism to determine the static foreground. Finally, the model samples are updated with the minimum sample confidence strategy, and a two-dimensional confidence update strategy and an adaptive weight update strategy are constructed to update the corresponding confidence and weight.

3.1.1. Pixel Classification for Sample Consistency. The pixel classification method based on sample consistency is based on statistics to determine whether the current pixel is the foreground by comparing the current frame with the samples in the model. In the actual activity of pixel classification, two different pixel matching methods are shown according to the difference of the background model.

(1) Color dimension pixel classification: In a background model composed of individual color features such as the ViBe algorithm and the PBAS algorithm, the samples in the model are composed of n color values, as shown in the following equation:

$$B(x) = \{v_1(x), v_2(x), \ldots, v_n(x)\}. \tag{1}$$

Therefore, only the pixel classification of the color dimension is required in the foreground detection stage. The pixel classification method is shown in Figure 1. The pixel point $I_t(x)$ of the current frame is taken as the center, and the color value of the sample in the model in a circle with a radius of $R(x)$ is compared, and the number of samples in the circle is recorded. If it is greater than or equal to the minimum matching number min, the current pixel can be judged to be the background; otherwise, it is the foreground. The equation is as follows:

$$F_t(x) = \begin{cases} 1, & \text{if} \#\{\text{dist}(I_t(x), B_n(x)) < R(x), \forall n\} < \min, \\ 0, & \text{otherwise}, \end{cases} \tag{2}$$

where $F_t(x) = 1$ indicates that the current pixel is the foreground; otherwise it is recorded as the background. $\#\{\ldots\}$ represents the number of samples whose Euclidean distance between the observation and the sample is less than the given distance threshold. min is the minimum number of matches that satisfy the condition.

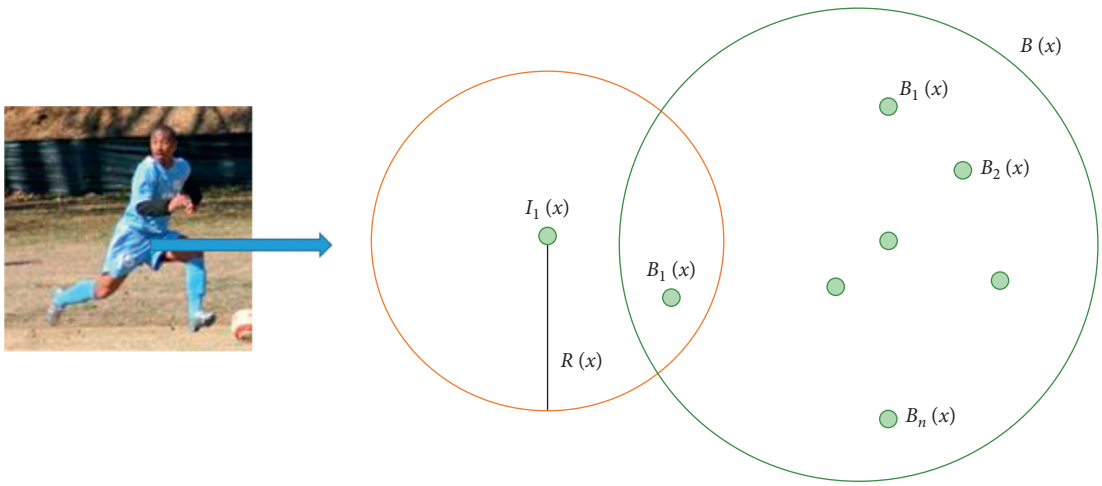


Figure 1: Example of color dimension pixel classification.

(2) Joint verification: Different from the background model composed of a single color value, the pixel-level model constructed by combining the color value and the texture value (taking the LBSP texture feature value as an example) uses the joint verification method of color dimension and texture dimension when performing foreground detection. For pixel classification, the representative ones are LOBSTER algorithm, SuB SENSE algorithm, etc. As shown in Figure 2, the pixel classification of the color dimension is first performed during pixel classification. If it is preliminarily determined as the foreground in the color dimension, the pixel classification of the texture dimension is performed; otherwise, the pixel classification activity of the texture dimension is not performed. If the logical AND operation value is 1, it means that a match is obtained; otherwise, it is not considered a match.

$$\left(\operatorname{dist}\left(v_t(x), v_i(x)\right) < R_c(x)\right) \&\& \left(\operatorname{dist}\left(\operatorname{lbsp}_t(x), \operatorname{lbsp}_i(x)\right) < R_{L1}(x)\right). \tag{3}$$

As can be seen from Figure 2, the pixel classification of the texture dimension is consistent with the pixel classification of the color dimension. The difference is that the pixel classification of the texture dimension uses the Hamming distance as a measure of similarity distance. Predict and determine whether the current pixel is the foreground or background by judging whether the number of matches obtained is less than the minimum matching value min. Taking the color feature and LBSP texture feature to construct the background model as an example, the pseudocode implementation process based on the double verification method of color level and texture level is shown in Algorithm 1.

3.2. Pixel Classification Algorithm Based on Confidence Weighted Fusion. In the model initialization stage, the background model $B(x)$ is established by obtaining the pixel information of the previous frame. The model is composed of N modules, and the equation is as follows:

$$B(x) = \{B_1(x), B_2(x), \ldots, B_i(x), \ldots, B_n(x)\}. \tag{4}$$

Different from the pixel-level model of regular sample consistency, the template $B_i(x)$ consists of color value v_i, LBSP texture feature value $\mathrm{LBSP}_i(x)$, color level confidence level $C_i^1(x)$, and texture level confidence center $C_i^2(x)$ composition. The calculation equation is as follows:

$$B_i(x) = \left\{v_i, \mathrm{LBSP}_i(x), C_i^1(x), C_i^2(x)\right\}. \tag{5}$$

In the foreground segmentation, that is, when pixel classification is performed, the samples whose distance between the current pixel $I_t(x)$ and the sample in the model is less than the given distance value $R(x)$ are recorded as strong correlation samples, and the number is obtained n. Mark the color confidence $C_i^1(x)$ and texture confidence $C_i^2(x)$ corresponding to the strong correlation samples, denoted as $t_i^1(x)$ and $t_i^2(x)$, respectively. The calculation equation is as follows:

$$t_i^m(x) = \begin{cases} C_i^m(x), & \operatorname{dist}\left(I_{\mathrm{t}}(x), B_{\mathrm{i}}(x)\right) < R(x), \\ 0, & \text{otherwise}, \end{cases} \tag{6}$$

where m takes the value 1 or 2, corresponding to the color dimension and texture dimension. Euclidean distance is used for color dimension judgment, and Hamming distance is used for texture dimension judgment. Then, the color confidence and texture confidence of the strong correlation samples are, respectively, summed, and then weighted and

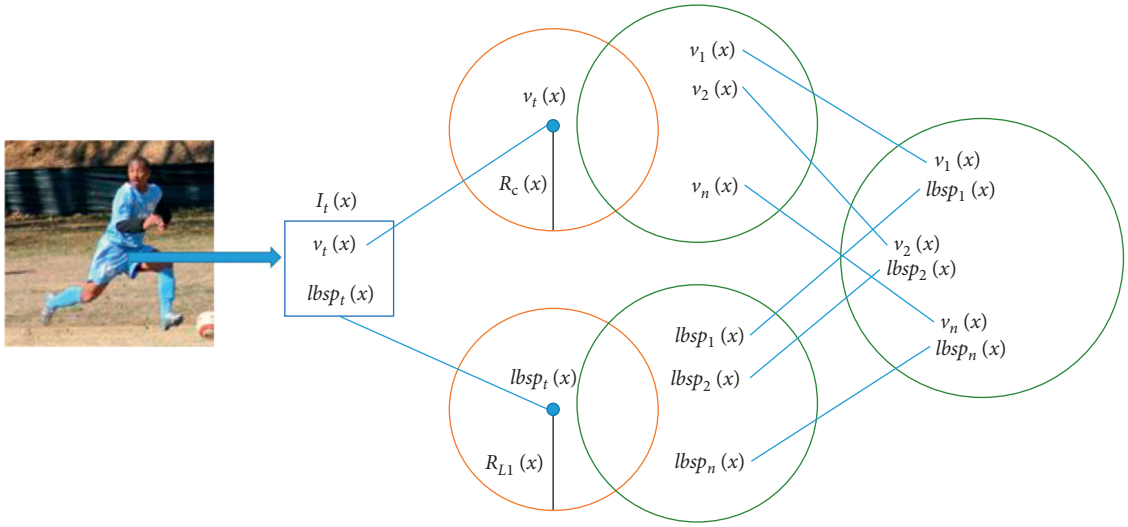


Figure 2: Example of joint verification pixel classification.

```
Input: current input frame pixel I_t(x)
Output: the foreground/background label of I_t(x)
initialize related variable (nCounts = 0, i = 0)
while nCounts < min && i < N
colorDist = dist(I_t(x), v_i(x))
if colorDist ≥ R_color
        goto wrongMatch;
lbspDist = dist(lbsp_t(x), LBSP_i(x))
if lbspDist ≥ R_color
        goto wrongMatch;
nCounts++;
   wrongMatch;
   i++;
   if nCounts < min
   I_t(x) is foreground
   else
   I_t(x) is background
```

Algorithm 1: Pixel classification for sample consistency.

summed. If it is less than the minimum threshold min, it is judged as the foreground; otherwise, it is the background.

$$F(x) = \begin{cases} 1, & \lambda_1(x)\sum_{i=1}^{n} t_i^1(x) + \lambda_2(x)\sum_{i=1}^{p} t_i^2(x) \le \min, \\ 0, & \text{otherwise.} \end{cases} \tag{7}$$

3.3. DetectNet. This section describes the specific implementation of DetectNet. Due to the outdated technology and network framework of the network architecture provided by the original text [17], it cannot meet the requirements of the overall framework of the system in this article. It is necessary to reimplement CenterNet, including training and testing of the network architecture. The input image is $I \in R^{W*H*3}$, and the target output is the heat map of generating key points $\widehat{Y} \in [0,1]^{(W/R)*(H/R)*C}$ where R is the transformation scale and C is the number of keypoint output feature channels (that is, the number of categories), and three basic frameworks are needed to generate the heat map: ResNet (including ResNet101, ResNet18, etc.), stacked hourglass network, and deep layer aggregation (DLA). As shown in Figure 3, (a) hourglass network, we use it just like in CornerNet; (b) ResNet and transposed convolution. Add a 3×3 deformable convolutional layer before each upsampling layer. Specifically, first, use deformable convolution to change the channel and then use convolution to upsample the feature map (the two steps are, respectively, in $32 \longrightarrow 16$. This article shows these two steps together as $16 \longrightarrow 8$ and $8 \longrightarrow 4$ dotted arrows); (c) use the original DLA-34 for semantic segmentation; (d) about DLA-34. Add more skip

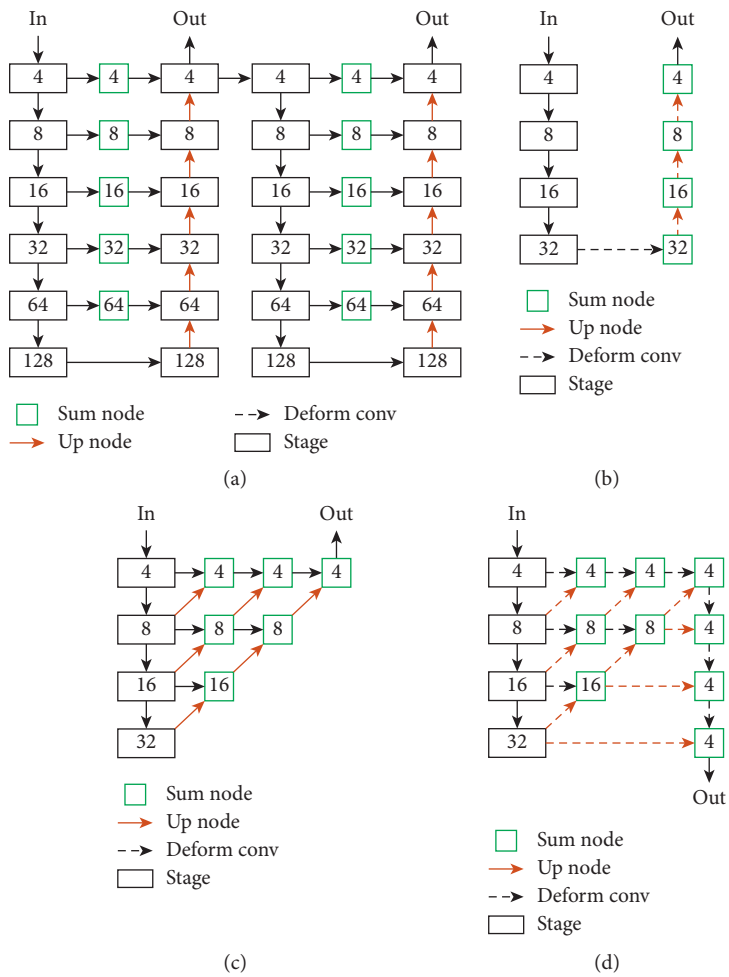


Figure 3: Proposed DetectNet.

connections at the bottom layer and upgrade each convolutional layer into the upsampling stage to a deformable convolutional layer. Combining the performance of each network given in the paper and our test results, taking into account the needs of this paper, realize the ResNet-based DetectNet network.

The objective function of the training is a pixel-level logistic regression. For the key point p of each category c, first, calculate a low-resolution equivalent value $\tilde{P} = [P/R]$ and then calculate the key on the heat map through the Gaussian kernel. The point is as follows:

$$Y_{xyc} = \exp\left(-\frac{\left(x-\breve{p}_x\right)^2+\left(y-\breve{p}_y\right)^2}{2\sigma_p^2}\right), \tag{8}$$

$$L_k = \frac{1}{N}\sum_p \begin{cases} \left(1-\widehat{Y}_{\mathrm{xyc}}\right)^{\alpha} * \log\left(\widehat{Y}_{\mathrm{xyc}}\right) \text{if } Y_{\mathrm{xyc}} = 1 \\ \left(1-\widehat{Y}_{\mathrm{xyc}}\right)^{\beta} * \left(\widehat{Y}_{\mathrm{xyc}}\right)^{\alpha} \\ \log\left(\widehat{Y}_{\mathrm{xyc}}\right) \end{cases}, \tag{9}$$

where α and β are the hyperparameters of focal loss, respectively, and N is the number of key points in the image.

When $Y_{xyc} = 1$, if $\widehat{Y}_{xyc}$ is close to 1, it means that this is very easy to detect point, and the corresponding $(1-\widehat{Y}_{\mathrm{xyc}})^{\alpha}$ is very low.

When $Y_{xyc} = 1$, if $\widehat{Y}_{xyc}$ is close to 0, then it means that this is a point that is not easy to detect. That is, the center point has not been learned, so the proportion of training should be increased, and the corresponding $(1-\widehat{Y}_{\mathrm{xyc}})^{\alpha}$ is very big.

4. Experiments and Results

4.1. Experimental Environment. Since the experiment in this article needs to train a deep neural network [24, 25], the scale is large, the structure is more complex, and the calculation scale is large. The programming language used is Python, the version is 3.6, the deep learning framework used is Keras2.1.5, and the IDE for program deployment is Pycharm, and all experiments are conducted in the same environment. All our experiments have been conducted on a desktop PC with an Intel Core i7-8700 processor and an NVIDIA GeForce GTX 1080ti GPU.

4.2. Datasets Preprocessing

4.2.1. 2D Pose Estimation Datasets. The training and verification of the 2D Pose network need to use some public data sets. Compared with the 3D human pose estimation, since the pictures we need can be obtained directly from the Internet, such as YouTube, Flickr, etc., the 2D human pose data set is relatively rich. In addition, these data sets have high-quality pictures while providing the two-dimensional coordinates of the skeleton of the characters in their images. These data sets not only provide image data containing the human body but also provide the annotation information corresponding to the pose of the human body, which is convenient for training your own 2D pose estimation network and has a unified standard evaluation and comparison of the performance of each network, generally using mAP as the evaluation standards; the following are some commonly used public data sets.

(1) MPII dataset, used as a benchmark for evaluating hinged human skeleton: The data set has about 25,000 high-quality images in total, of which more than 40,000 high-quality images containing people are annotated with two-dimensional skeleton coordinates, and the number of skeleton key points is 16. These images are derived from human daily life, collected by the classifier. There are about 410 different human activities, and they contain their corresponding two-dimensional skeleton coordinate labels; at the same time, most of these images are extracted from online videos, such as YouTube, so it also provides a lot of frame images without label information, which is very useful in the testing and training of some methods; at the same time, its test set contains the label information of the three-dimensional torso and the direction after the body is occluded.

(2) MSCOCO dataset, the full name is Miscrpspft COCO Dataset: we have already introduced the detection data set, so this data set is very large and has a wide range of uses. It can be used for detection and tracking, segmentation, and two-dimensional pose estimation. Here we only use the key point information of the human body's two-dimensional skeleton for two-dimensional pose estimation. As far as the MSCOCO data set contains the joint point information of the human body skeleton, this part includes more than 200,000 pieces of various types derived from the network. Daily pictures: There are more than 250,000 different people doing various actions in all pictures, and the skeleton coordinates are marked for these different people. The number of skeleton key points is 18, so we can see the data set. The amount of data far exceeds that of MPII.

(3) LSP dataset: The full name of the LSP data set is Leeds Sports Pose Dataset. This data set has a small amount of data. The overall data set contains about two thousand images of skeleton annotation information. Most of these images in the data set are from the data set obtained by Filcker which is mainly for the data set related to the task of moving target characters. Therefore, all images are required to contain not only the annotation information of the human skeleton but also the category labels of various sports, which is convenient to distinguish different sports categories and the number of key points of the human skeleton Is 14.

In the process of training the model, this article performs preprocessing such as enhancement to the data set, removing the traditional rotation of the picture by a certain angle, zooming in a certain proportion, etc.; this article uses a data enhancement processing method different from the above method; here, we call this data standardization. Reference [53] proposed a new standardization method for data, such as formula (10), to find a λ to minimize it.

$$\operatorname{argmin}_{\lambda}\left\|\lambda P_c - P_u\right\|^2. \tag{10}$$

P_c is the x, y coordinate value of the two-dimensional skeleton coordinate point, and P only takes the x, y coordinate value of the corresponding three-dimensional skeleton coordinate. λ is obtained by this method, and the accuracy obtained in the final training is significantly improved. This article studies several methods to solve the vector λ size is $N \times 1 \times 2$, where N generally corresponds to batch size (the value in this experiment is 1024).

4.2.2. 3D Human Pose Estimation Datasets. Because the production of 3D human skeleton data sets is very difficult and complicated, its collection requires a large number of cameras, sensors, and the full support of depth cameras to ensure the accuracy of the marked 3D coordinates. Therefore, many 3D skeleton data sets are indoors. It is difficult to have outdoor data, which leads to the lack of natural data to make the network learn better. However, the corresponding two-dimensional skeleton data is indeed diverse and accurate, containing a large amount of indoor and outdoor various types of data. This is also the reason why the existing two-dimensional pose estimation methods are more stable and accurate. As a result, many advanced three-dimensional pose estimations use two-dimensional data. Based on pose estimation, 3DPoseNet is also based on inputting two-dimensional skeleton key points to predict and output corresponding three-dimensional skeleton key points.

Human3.6M dataset is the main evaluation data set of existing advanced paper methods. This article is mainly based on this data set evaluation. There are 3.6 million images in the data set. There are 11 experimenters (6 males and 5 females) in the data, that is, 11 different characters. Each image contains the corresponding 2D skeleton and 3D skeleton coordinates. The data is captured by 4 digital cameras, 1 time sensor, and 10 sports cameras. It contains 17 action scenes such as discussion, eating, calling, sports, sitting, standing, walking, and greeting, etc. The number of key points of the skeleton is 17. Usually, data S1, S5, S6, S7,

(a) (b)

Figure 4: The 2D pose detection result on the LSP dataset.

S8 are used as the training set, while S9, S11 are used as the corresponding test set.

4.3. DPoseNet Experimental Results. According to the above experiments on many two-dimensional pose estimation methods, an independent and clear target person processed by 2DPoseNet (2DPN) will be obtained. Essentially, it is the process of transforming the two-dimensional human pose coordinates to the three-dimensional human pose coordinates, that is, improving the dimensionality. In this process, it is necessary to make full use of the position information between the key points of each joint of the human body. And 2DPoseNet generates 2D key points funny and accurately, especially for self-occlusion. The method belongs to the top-down method, so it has certain limitations. For example, when there are too many people in the multiperson 2D human pose estimation, our method will have certain defects in speed and accuracy. The position of each joint point of the human body is marked from the image, and each estimated position contains only a two-dimensional coordinate, so the final output of the two-dimensional human body pose estimation method is a two-dimensional human body pose. The position coordinates of each joint of the human body are drawn according to the image. That is, each estimated position corresponds to only one two-dimensional coordinate. Regarding our 2DPoseNet model training, we still use traditional data set training, such as LSP.

As shown in Figure 4 and Tables 1–3 , the single-person 2D human pose estimation is very accurate in the complex environment.

4.4. DPoseNet Experimental Results. This section will show the training process of the 3DPoseNet network. That is, to train the 3DPoseNet in this article, the initial learning rate used is $lr = 0.001$, and each cycle experiences exponential decay. Define dropout $p = 0.25$. For the convolutional layer, set the convolution kernel $W = 1$, and set it to $C = 1024$ output channels. Compared with the previous algorithm, we can see that the accuracy of the final 2D pose skeleton obtained by the method in this paper has been significantly improved. Based on the Human3.6M data set, as shown in Table 4, it is the result of comparing the advanced methods in this paper. Regarding the evaluation methods from 2D human pose estimation to 3D human pose estimation, this paper uses a series of 2D keypoint sequences as input and uses the same evaluation indicators as related papers on 3D pose estimation, namely P_MPJPE. Procrustes analysis MPJPE (P_MPJPE): Denoted as Protocol 2, which is MPJPE based on Procrustes analysis. After rigid transformations such as scaling, translation, and rotation are performed on the network output, the MPJPE is calculated after aligning to the true value.

As shown in Table 4, it can be clearly found that the accuracy of this method is greatly improved.

As shown in Figures 5 and 6 , the visual results of the single-person pose predicted by the method in this paper are all single-person 3D human pose renderings based on 30fps video processing in this paper. As shown in Figure 6, it is the effect of the multiperson 3D pose predicted by the method in this paper. The method in this paper can basically accurately predict the movements of the multiperson movement.

Table 1: Average precision of joint detection on LSP.

LSP	Ankle	Knee	Hip	Wrist	Elbow	Shoulder	Neck	Head	mAP
Wang et al. [26]	0.25	0.34	0.23	0.23	0.43	0.32	0.21	0.34	0.31
Pishchulin et al. [27]	0.37	0.31	0.36	0.31	0.43	0.33	0.32	0.76	0.36
Ours	**0.47**	**0.41**	**0.49**	**0.39**	**0.44**	**0.52**	**0.59**	**0.80**	**0.49**

Table 2: Average precision of joint detection on MSCOCO.

MSCOCO	Ankle	Knee	Hip	Wrist	Elbow	Shoulder	Neck	Head	mAP
Wang et al. [26]	0.20	0.31	0.21	0.19	0.39	0.31	0.19	0.31	0.30
Pishchulin et al. [27]	0.32	0.28	0.32	0.33	0.42	0.36	0.31	0.72	0.33
Ours	**0.42**	**0.39**	**0.45**	**0.34**	**0.43**	**0.49**	**0.47**	**0.77**	**0.41**

Table 3: Average precision of joint detection on MPII.

MPII	Ankle	Knee	Hip	Wrist	Elbow	Shoulder	Neck	Head	mAP
Wang et al. [26]	0.31	0.30	0.25	0.28	0.39	0.33	0.27	0.32	0.34
Pishchulin et al. [27]	0.39	0.33	0.32	0.30	0.43	0.39	0.33	0.74	0.33
Ours	**0.42**	**0.46**	**0.43**	**0.37**	**0.48**	**0.51**	**0.49**	**0.78**	**0.45**

Table 4: The results of the Mean Per Joint Position Error(MPJPE), with less values, mean better results.

MPII	Direct.	Disc.	Eat	Greet	Phone	Photo	Pose	Purch.	Sit
Martinez et al. [28]	0.39	0.43	0.46	047	0.51	0.56	0.41	0.40	0.56
Sun et al. [29]	0.42	0.44	0.45	0.45	0.51	0.53	0.43	0.41	0.59
Fang et al. [30]	0.38	0.41	0.43	0.44	0.48	0.55	0.40	0.38	0.54
Pavlakos et al. [31]	0.34	0.39	0.41	0.38	0.42	0.47	0.38	0.36	0.50
Yang et al. [32]	0.26	0.30	0.39	0.39	0.43	0.47	0.28	0.29	0.36
Hossain et al. [33]	0.35	0.39	0.43	0.43	0.47	0.54	0.38	0.37	0.51
Pavllo et al. [34]	0.34	0.36	0.37	0.37	0.36	0.42	0.34	0.33	0.45
Ours	**0.26**	**0.31**	**0.28**	**0.29**	**0.32**	**0.38**	**0.29**	**0.27**	**0.36**

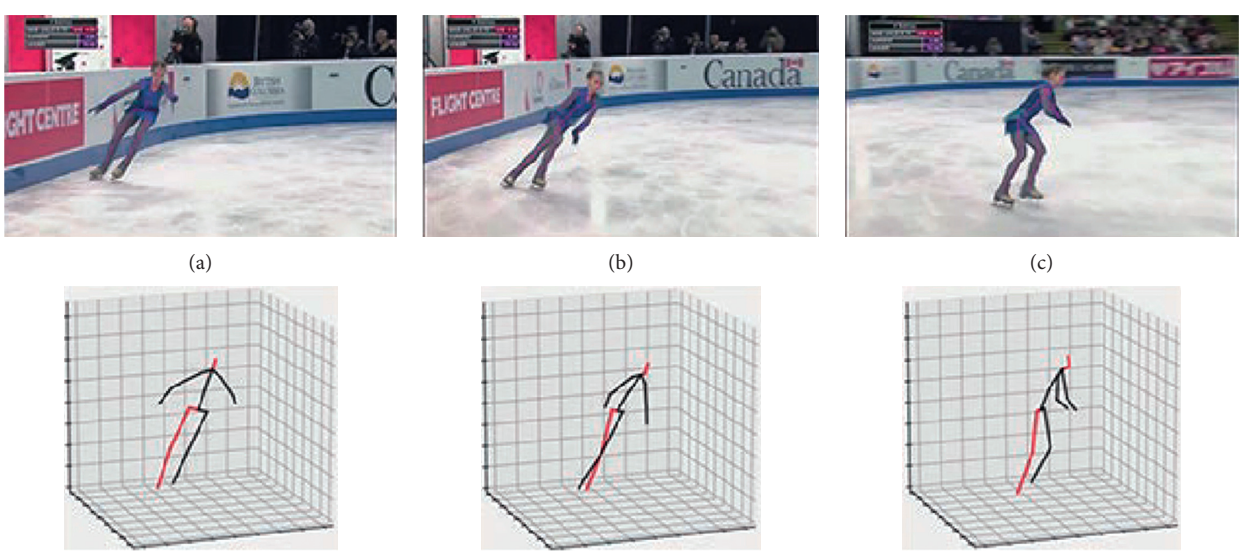


Figure 5: The single-person 3D pose detection results.

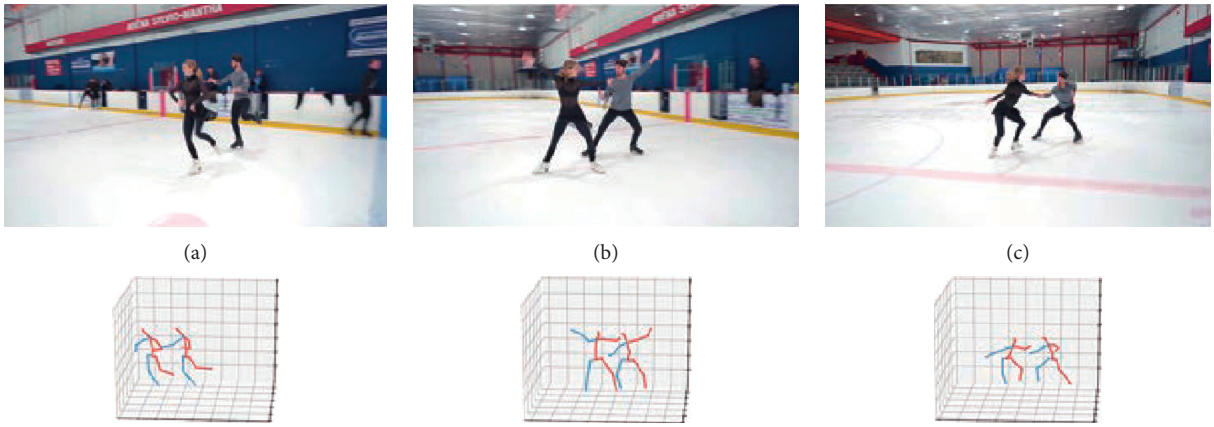


Figure 6: The Multiperson 3D pose detection results.

5. Conclusion

Since football player's posture detection has important application value in sports competition for different purposes of analysis. Therefore, this paper proposes a football motion detection approach combining foreground detection and deep learning for real-time detection of football players' posture. The main task of foreground target detection is to extract the foreground target of interest in the real surveillance scene and use it as the target of interest for subsequent analysis. Then, we propose a triple DetectNet detection framework based on deep learning technology, which can quickly and reliably realize the three-dimensional pose estimation of multiperson motion. For input, the triple DetectNet framework uses three neural networks and is executed in three stages; the first stage is to use the DetectNet network to detect the bounding box of each person separately, the second stage uses the 2DPoseNet network to estimate each person's response, and the third stage uses the 3DPoseNet network to obtain the person's 3D pose. The validity of the proposed approach has been tested by conducting several experiments on four data sets, and the results proved the success and superiority of the algorithm.

Data Availability

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

Conflicts of Interest

The author does not have any possible conflicts of interest.

References

[1] G. Hinton, "Deep belief networks," *Scholarpedia*, vol. 4, no. 5, p. 5947, 2009.

[2] X. Ning, K. Gong, W. Li, L. Zhang, X. Bai, and S. Tian, "Feature refinement and filter network for person re-identification," *IEEE Transactions on Circuits and Systems for Video Technology*, 2020.

[3] W. Cai and Z. Wei, "PiiGAN: generative adversarial networks for pluralistic image inpainting," *IEEE Access*, vol. 8, pp. 48451–48463, 2020.

[4] Y. Tong, L. Yu, S. Li, J. Liu, H. Qin, and W. Li, "Polynomial fitting algorithm based on neural network," *ASP Transactions on Pattern Recognition and Intelligent Systems*, vol. 1, no. 1, pp. 32–39, 2021.

[5] W. Cai and Z. Wei, "Remote sensing image classification based on a cross-attention mechanism and graph convolution," *IEEE Geoscience and Remote Sensing Letters*, 2020.

[6] X. Ning, Y. Wang, W. Tian, L. Liu, and W. Cai, "A biomimetic covering learning method based on principle of homology continuity," *ASP Transactions on Pattern Recognition and Intelligent Systems*, vol. 1, no. 1, pp. 9–16, 2021.

[7] W. Cai, Z. Wei, R. Liu, Y. Zhuang, Y. Wang, and X. Ning, "Remote sensing image recognition based on multi-attention residual fusion networks," *ASP Transactions on Pattern Recognition and Intelligent Systems*, vol. 1, no. 1, pp. 1–8, 2021.

[8] X. Zhang, Y. Yang, Z. Li, X. Ning, Y. Qin, and W. Cai, "An improved encoder-decoder network based on strip pool method applied to segmentation of farmland vacancy field," *Entropy*, vol. 23, no. 4, p. 435, 2021.

[9] Z. Chu, M. Hu, and X. Chen, "Robotic grasp detection using a novel two-stage approach," *ASP Transactions on Internet of Things*, vol. 1, no. 1, pp. 19–29, 2021.

[10] R. Liu, X. Ning, W. Cai, and G. Li, "Multiscale dense cross-attention mechanism with covariance pooling for hyperspectral image scene classification," *Mobile Information Systems*, vol. 2021, 2021.

[11] Z. L. Yang, S. Y. Zhang, Y. T. Hu, Z. W. Hu, and Y. F. Huang, "VAE-Stega: linguistic steganography based on variational auto-encoder," *IEEE Transactions on Information Forensics and Security*, vol. 16, pp. 880–895, 2020.

[12] X. Ning, X. Wang, S Xu et al., "A review of research on co-training," *Concurrency and Computation: Practice and Experience*, 2021.

[13] A. Keivani, J. R. Tapamo, and F. Ghayoor, "Motion-based moving object detection and tracking using automatic K-means," in *Proceedings of the 2017 IEEE AFRICON*, pp. 32–37, IEEE, Cape Town, South Africa, September 2017.

[14] K. E. Van De Sande, J. R. Uijlings, T. Gevers, and A. W. Smeulders, "Segmentation as selective search for object recognition," in *Proceedings of the 2011 International*

Conference on Computer Vision, pp. 1879–1886, IEEE, Barcelona, Spain, November 2011.

[15] K. He, X. Zhang, S. Ren, and J. Sun, "Spatial pyramid pooling in deep convolutional networks for visual recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 37, no. 9, pp. 1904–1916, 2015.

[16] R. Girshick, "Fast R-CNN," in *Proceedings of the IEEE International Conference on Computer Vision*, pp. 1440–1448, Santiago, Chile, December 2015.

[17] X. Zhou, D. Wang, and P. Krähenbühl, "Objects as points," 2019, https://arxiv.org/abs/1904.07850.

[18] J. Redmon and A. Farhadi, "Yolov3: an incremental improvement," 2018, https://arxiv.org/abs/1804.02767.

[19] A. Beaugendre and S. Goto, "Block-propagative background subtraction system for UHDTV videos," *Information and Media Technologies*, vol. 10, no. 2, pp. 259–262, 2015.

[20] A. Beaugendre, S. Goto, and T. Yoshimura, "Real-time UHD background modelling with mixed selection block updates," *IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences*, vol. E100.A, no. 2, pp. 581–591, 2017.

[21] A. Beaugendre and S. Goto, "Adaptive block-propagative background subtraction method for UHDTV foreground detection," *IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences*, vol. E98.A, no. 11, pp. 2307–2314, 2015.

[22] M. F. Savaş, H. Demirel, and B. Erkal, "Moving object detection using an adaptive background subtraction method based on block-based structure in dynamic scene," *Optik*, vol. 168, pp. 605–618, 2018.

[23] D. Sakkos, H. Liu, J. Han, and L. Shao, "End-to-end video background subtraction with 3D convolutional neural networks," *Multimedia Tools and Applications*, vol. 77, no. 17, pp. 23023–23041, 2018.

[24] C. Yan, G. Pang, X. Bai, J. Zhou, and L. Gu, "Beyond triplet loss: person Re-identification with fine-grained difference-aware pairwise loss," *IEEE Transactions on Multimedia*, 2021.

[25] C. Wang, X. Bai, X. Wang et al., "Self-supervised multiscale adversarial regression network for stereo disparity estimation," *IEEE Transactions on Cybernetics*, 2020.

[26] F. Wang and Y. Li, "Beyond physical connections: tree models in human pose estimation," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 596–603, Portland, OR, USA, June 2013.

[27] L. Pishchulin, M. Andriluka, P. Gehler, and B. Schiele, "Poselet conditioned pictorial structures," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 588–595, Portland, OR, USA, June 2013.

[28] J. Martinez, R. Hossain, J. Romero, and J. J. Little, "A simple yet effective baseline for 3D human pose estimation," in *Proceedings of the IEEE International Conference on Computer Vision*, pp. 2640–2649, Venice, Italy, October 2017.

[29] X. Sun, J. Shang, S. Liang, and Y. Wei, "Compositional human pose regression," in *Proceedings of the IEEE International Conference on Computer Vision*, pp. 2602–2611, Venice, Italy, October 2017.

[30] H. S. Fang, Y. Xu, W. Wang, X. Liu, and S. C. Zhu, "Learning pose grammar to encode human body configuration for 3D pose estimation," *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 32, no. 1, 2018.

[31] G. Pavlakos, X. Zhou, and K. Daniilidis, "Ordinal depth supervision for 3D human pose estimation," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 7307–7316, Salt Lake City, UT, USA, June 2018.

[32] W. Yang, W. Ouyang, X. Wang, J. Ren, H. Li, and X. Wang, "3D human pose estimation in the wild by adversarial learning," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 5255–5264, Salt Lake City, UT, USA, June 2018.

[33] M. R. I. Hossain and J. J. Little, "Exploiting temporal information for 3d human pose estimation," in *Proceedings of the European Conference on Computer Vision (ECCV)*, pp. 68–84, Munich, Germany, September 2018.

[34] D. Pavllo, C. Feichtenhofer, D. Grangier, and M. Auli, "3D human pose estimation in video with temporal convolutions and semi-supervised training," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 7753–7762, Long Beach, CA, USA, June 2019.

Hindawi
Scientific Programming
Volume 2021, Article ID 2348494, 8 pages
https://doi.org/10.1155/2021/2348494

Research Article

Recognition and Classification Model of Music Genres and Chinese Traditional Musical Instruments Based on Deep Neural Networks

Ke Xu

School of Art and Design, Qingdao University of Technology, Qingdao 266033, Shandong, China

Correspondence should be addressed to Ke Xu; xuke@qtech.edu.cn

Received 19 April 2021; Revised 31 May 2021; Accepted 7 June 2021; Published 22 June 2021

Academic Editor: Shah Nazir

The teaching of ideological and political theory courses and daily ideological and political education are two important parts of education for college students. With the iterative update of information technology, the individualized development of students, and the reform and innovation of ideological and political education, higher goals and requirements have been put forward for ideological and political education. Some universities have developed new paths in the teaching model, but they have not considered the evaluation module and paid little attention to their own development. They only paid attention to the fact that it injected fresh blood into the reform of education model and ideological education but ignored the improvement of their own quality. Therefore, with these limitations, the learning effect is not satisfactory. Keeping in view these issues, this article defines the concept of deep learning and ideological and political education of college students as the starting point and then analyzes the new precise and personalized concepts, new forms of intelligent teaching and evaluation, and new models of intelligent learning that deep learning brings to college students' ideological and political education. This is a new path of intelligent linkage with the subject, object, and mediator. It can deepen the reform of the education and teaching mode of individualization, accuracy, interactivity, and vividness of college students' ideological and political education and improve the evaluation and management of college students' ideological and political education. The experimental results of the study showed the effectiveness of the proposed study.

1. Introduction

Music is an abstract art that uses sound as a means of expression to reflect human emotions in real life [1]. Music, as an important component of human spiritual life, has occupied an important position in human daily life. Music can improve concentration, relieve people's pressure on work and study, and be good for physical and mental health [2]; music can bring people aural pleasure and spiritual enjoyment [3], help get rid of bad emotions such as sadness, loneliness, and sadness, and make people full of energy and passion. With the rapid development of Internet technology and digital multimedia technology, digital media resources represented by audio and video have obtained good transmission channels and convenient storage media, and digital music resources and Internet music entertainment consumer users have shown explosive growth. The era of digital music has arrived.

Music information extraction [4] has become a popular research direction in the field of computer science. Music genre classification is an important research content in the field of music information extraction. Music genre is a notable label that distinguishes music, and it is also the category that listeners pay most attention to and retrieve the most. In the past, the classification of music genres mostly used manual labeling methods. Manual labeling is to ask professionals with a professional background in music and higher musical literacy to label music works by category. With the continuous emergence of music creation and online uploads, the digital music resource library on the Internet has become increasingly large, and manual labeling methods have gradually failed to meet the needs. To classify a

large digital music resource library, if manual labeling is used, it will consume a lot of manpower and time, and the labeling results are more subjective, and the labeling standards cannot be completely unified, which is limited by different professionals who label music. Therefore, the automatic classification of music [5–7] has gradually become a research hotspot for researchers. The automatic classification of music genres can effectively solve the problem of high cost and time-consuming human labeling. Through the algorithm, a unified classification standard can be formulated, and the algorithm can be continuously optimized, and a highly accurate and objective classification result can be obtained.

Due to the limited application of manual extraction of music features [8], the robustness is poor, and it is difficult to describe the deep features and timing characteristics of music. Moreover, in the current music genre classification tasks, traditional machine learning classifiers are mainly used, including BP neural networks, support vector machines, and nearest neighbor algorithm classifiers. Due to its shallow structure, the classifier limits the learning of music features, and it is difficult to extract more effective features to represent music, which affects the accuracy of classification. In recent years, deep neural networks [9–12] have achieved good results in natural language processing, computer vision [13–16], and other research fields. The deep neural network model can automatically learn deeper features from the shallow features and can reflect the local relevance of the input data. Deep learning provides a new solution for the automatic classification of music.

Therefore, this study first studied the music genre recognition [17, 18] and classification algorithm [19, 20] based on deep neural network and improved the algorithm. Compared with the classic algorithm that directly extracts the acoustic features or music features of music and trains with a classifier to obtain the recognition and classification results, this algorithm improves the accuracy of the recognition and classification of music genres. At the same time, for the recognition and classification of musical instruments, this study proposes a Chinese traditional musical instrument recognition and classification algorithm based on the deep belief network in deep learning. The deep belief network is used in the feature extraction task of traditional Chinese musical instrument music [21], which reduces the work of manual extraction and identification of features. At the same time, the recognition and classification effect has also been improved compared with the classic algorithm. The following are the main innovation points of this study:

(i) Combining Bi-GRU and attention mechanism, a novel music genre classification model is proposed, which can learn more significant music features, thereby improving the accuracy of classification.

(ii) A Chinese traditional musical instrument recognition and classification algorithm was proposed based on a deep belief network. The deep belief network is used in the feature extraction task of Chinese traditional musical instrument music, which limited the recognition and classification effect.

This study is structured as follows. Section 2 shows the background of the study. The methodology section of the study is given in Section 3 with details in the subsections. Section 4 briefly explains the experiments and results of the study performed. The study is ended in Section 5 which is the conclusion.

2. Background

With the rising and advancements in information technology, the individualized development of students, and the reform and innovation of ideological and political education, higher goals and requirements have been put forward for ideological and political education. The following are the details of this section.

2.1. Blues. It originated from the amateur music of poor black slaves in the south of the United States in the past. It had no accompaniment, but a solo singing with emotional content, and later combined with the European chord structure to form music of singing and guitar alternately. The blues are based on the pentatonic scale, which is composed of five scales arranged in pure fifths.

2.2. Classical. It is the traditional musical art of Western music, and it is music created under the background of mainstream European culture. The most prominent feature of classical music is that its works generally use notation to record the score, so that the rhythm and pitch can be recorded in detail, and it is also conducive to the direct coordination of multiple performers. Many types of musical instruments are used in classical music, including woodwind, brass, percussion, keyboard, bowed, and plucked stringed instruments.

2.3. Jazz. It originated from the blues, combining and absorbing classical music, folk music, and other musical styles on the basis of African music traditions, and gradually formed today's diverse jazz music.

2.4. Country. It originated in the southern United States. It is a kind of popular music with ethnic characteristics. The main characteristics of country music are its simple tune, steady rhythm, being mainly narrative, and a strong local flavor, mostly in the form of ballads, with body, two-part, or three-part form. Country music is mostly solo or chorus, with harmonica, guitar, violin, and other accompaniment. The themes of country music are generally love, country life, cowboy humor, family, God, and country.

2.5. Rock. It originated in the mid-1950s and was developed under the influence of blues and country music. It is characterized by prominent vocals and played with guitar, bass, and drum accompaniment, and keyboard instruments such as electronic organs, organs, and pianos are often used. Rock music has a strong beat, centered on various guitar sounds.

2.6. Metal. It is a kind of rock music, which was developed in Britain and the United States in its early days. Metal music has the characteristics of high explosive power, weight, and speed. Its weight is reflected in the low scale of electric guitars and point bass. The speed is reflected in the beat, the beat of metal music can reach more than 200 BPM, and the beat range of general pop music is only 80–130 BPM. The core instruments of metal music are electric guitar, electric bass, and drums, which control the rhythm and melody.

2.7. Disco. It is a kind of electronic music, which originated from African American folk dance and jazz dance. In rhythm, the characteristics of rock music, jazz, and Latin American music are mixed. As ballroom music, disco is characterized by a strong sense of rhythm, arranged by lively string music. Disco is generally 4/4 shots, and every shot is strong, about 120 BPM.

2.8. Pop. It originated in Britain and the United States in the mid-1950s. Popular music is eclectic, often borrowing elements of other styles of music. But pop music also has its core elements: its structure is relatively short, usually about three minutes.

2.9. Hip-Hop. It originated in New York, USA, when it was popular among African Americans and neighborhood gatherings. Hip-hop consists of two main components: rap and DJing. The performer sings in the way of saying words according to the rhythm of the instrument or synthesis.

2.10. Reggae. It is derived from the popular music of Ska and Rock Steady, which evolved in Jamaica. It is the general term for various dance music in Jamaica.

2.11. Electronic. It is a kind of music made using electronic musical instruments and electronic technology. In electronic music, a variety of genres are often combined, and they are modulated into unique timbres through electronic musical instruments and synthesizers to form a unique style. Commonly used electronic musical instruments include electric guitars, electric basses, synthesizers, and electronic organs.

2.12. Punk. It is simple rock music derived from Garage Rock and pre-punk rock, consisting of three chords and a simple main melody.

3. Methodology

3.1. Feature Sequence Extraction of Music Segment. The process of extracting the feature sequence of the music segment is shown in Figure 1. First, analyze the music file to extract the note feature matrix, then perform the main melody extraction and segment division based on the note feature matrix, and then combine the time points of the segment division and the main melody of the music to extract the feature vector based on the main melody for each segment. It composes the feature sequence of the music segment and serves as the input of the later classifier.

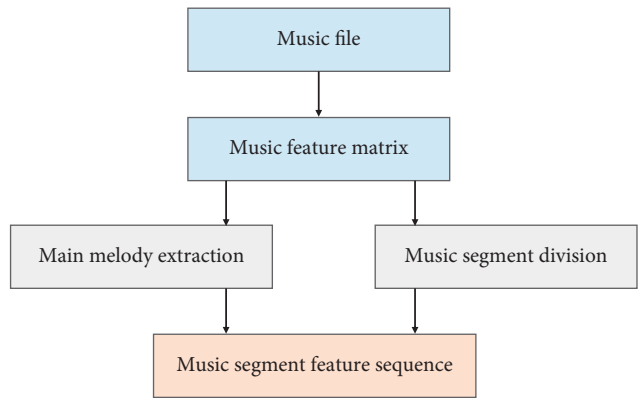


Figure 1: The specific process of extracting the feature sequence of music segment.

3.1.1. Main Melody Extraction. Listening to a piece of music, the perceptual information that people mainly obtain from the sense of hearing is the main melody of the music. The main melody is the soul of music and interprets the theme of music. The main melody of music is the key to music classification and an important basis for distinguishing music genres. This section studies and implements a fast and effective Skyline main melody extraction algorithm for extracting the main melody from music files.

We define the relevant attributes of the notes. Let n_i and n_{i+1} denote two adjacent notes, s_i and s_{i+1} denote the start time of these two notes, respectively, p_i and p_{i+1} denote the pitch of these two notes, respectively, and e_i and e_{i+1} denote the end of these two notes, respectively.

The input of the Skyline algorithm is the note feature matrix. The following describe the specific steps of the Skyline algorithm:

(1) Arrange the note vectors in the note feature matrix in ascending order of their starting time, and remove the note vectors of channel 10 percussion instruments.

(2) Traverse the note feature matrix. For note vectors with the same starting time, keep the note vector with the highest pitch and discard other note vectors.

(3) For two adjacent note vectors n_i and n_{i+1}, if $s_i < s_{i+1}$, $e_i > s_{i+1}$, and $p_i < p_{i+1}$ are satisfied, let $e_i = s_{i+1}$.

3.1.2. Music Segment Division. Firstly, the sound file is sampled, framed, and coded, and the piano roll matrix is used to model the music playing; then the similarity between any two frames is calculated by Euclidean distance to generate a self-similar matrix, and a special Gaussian is constructed. The convolution kernel convolves along the diagonal of the self-similar matrix to generate a novelty curve. The novelty curve is a time-series curve describing the changes in musical performance. Finally, the peak points are

extracted from the novelty curve and segmented. The core idea of the algorithm is to estimate the instantaneous music novelty by analyzing the local self-similarity of music playing; at a significant novel point in time, the music played in the past or the future at that point in time is within a short period of time. It has a high degree of self-similarity, and there is a fairly low cross-similarity between the past and the future at this point in time. Simply put, in a short period of time before this point in time, the musical style of playing is similar. After this point in time, the musical composition is changed to another style of playing. The artistic style of playing music has undergone major changes, and the emotions and themes expressed have also changed, so the music segments can be divided.

3.2. Attention Mechanism. When humans are observing visual images, the human brain quickly scans the images that appear in the field of view and controls the line of sight to fall on the area that you want to focus on. The human brain will allocate different attention to observation according to different areas in the field of view image. For the areas that the field of view focuses on, the human brain will allocate more attention resources to observe carefully to obtain more details of the target area. Information will be ignored for other useless areas of view. The attention mechanism [22, 23] in deep learning is similar to this. It is also a mechanism of attention resource allocation. It can filter out key information that is more conducive to deep learning tasks from a large amount of information, thereby improving the performance of deep learning tasks such as detection [24], prediction [25], and recognition [26, 27].

Figure 2 shows a simplified schematic diagram of the encoding and decoding model that introduces the attention mechanism. The codec model with the attention mechanism can effectively improve its limitations. The encoder no longer converts all the information of the input sequence into a fixed-length context vector. For different outputs, it will focus on finding significant useful information related to the current output from the input data, and calculate different context vectors. Allow the model to better learn the alignment of input and output.

Taken separately, the attention mechanism can be understood as a query calculation process. Figure 3 is a generalized structure diagram of the attention mechanism. x is the input sequence data, and y is the query. First, input y, calculate the attention score of y and each input x_i through the function f, and then map the probability distribution between 0 and 1 through the softmax function. Finally, the probability distribution and each input are correspondingly weighted. Then, calculate the output value of the attention mechanism.

The calculation equation of the attention mechanism is as follows:

$$\text{attention} = \sum_{i=1}^{n} \text{soft}\max\left(f\left(x_i, y\right)\right) * x_i. \tag{1}$$

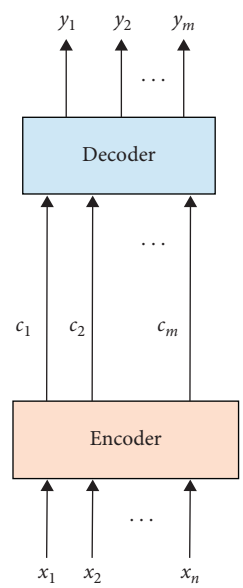


Figure 2: Schematic diagram of the codec model with the introduction of the attention mechanism.

3.3. Classification Model. Compared with 2D convolutional networks, 3D convolutional networks can better model time information through 3D convolution and 3D pooling operations. In a two-dimensional convolutional network, the process of convolution and pooling is completed in space. In a three-dimensional convolutional network, they perform in time and space. In the introduction of 3D convolutional network above, it was proposed that images should be output when 2D convolutional network is processing images, and images should also be output when multiple images (which are regarded as different channels) are operated. Therefore, the time information of input data will be lost after each convolution operation in the two-dimensional convolutional network. Only three-dimensional convolution can preserve the time information of the input signal and produce the output quantity. The same principle can be applied to 2D pooling and 3D pooling.

Figure 4 is the network model structure diagram of the classification of music genres in this article. The classification network model designed in this study can be divided into three parts according to different functions, namely, the input layer, the hidden layer, and the output layer. The input of the input layer is a sequence of musical segment features extracted from music. The main function of the hidden layer is to learn the final feature representation of music. The hidden layer is composed of Bi-GRU, attention mechanism, and fully connected layer.

In the attention mechanism, this article uses the following formula to calculate the attention score corresponding to each feature vector:

$$e_t = \tanh\left(WH_t + b\right), \tag{2}$$

where e_i is the attention score of the feature vector H_t at the time t in H.

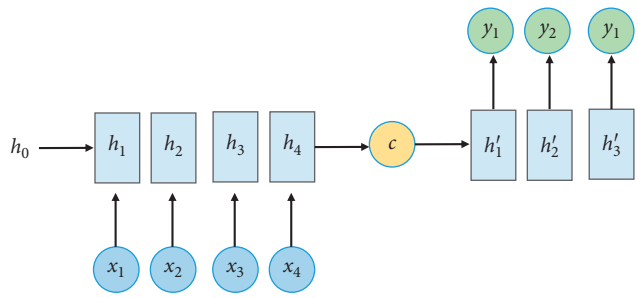


Figure 3: Attention mechanism.

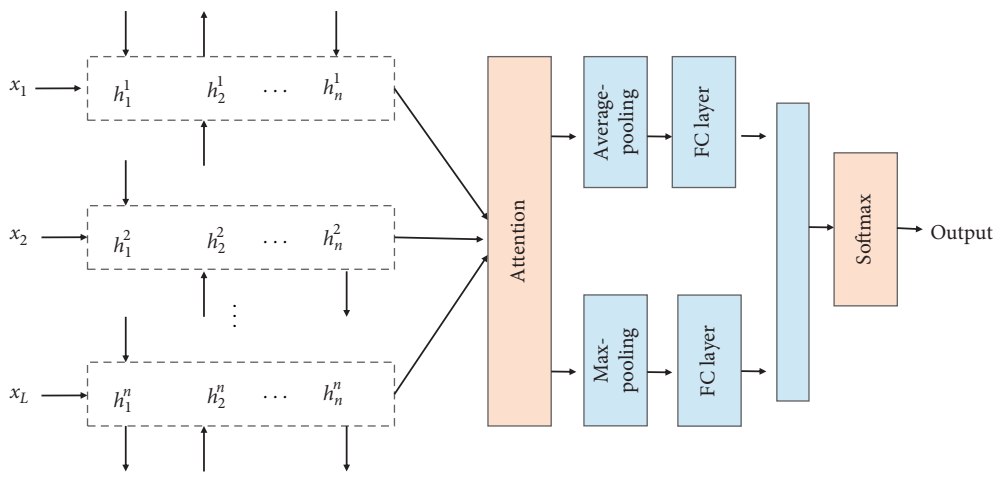


Figure 4: Classification model.

Then, the calculated attention score is mapped to the value range (0, 1) through the softmax function, and the attention probability distribution of each feature vector is obtained:

$$a_i = \text{soft max}(e_i) = \frac{\exp(e_i)}{\sum_{k=1}^{L} \exp(e_k)}. \tag{3}$$

The calculated attention probability distribution and each feature vector of the feature representation H are weighted and summed to obtain the feature vector representation v of the music file:

$$v = \sum_{i=1}^{L} a_i H_i. \tag{4}$$

The combination of Bi-GRL and the attention mechanism network allows the model to effectively learn the value information of the different weights of the genre classification of each piece of music, including forward and backward value information, and more accurately from the input music. The useful information learned from the segment feature sequence is helpful to improve the accuracy of classification.

At the end of the hidden layer, the music feature vector powder extracted by the attention mechanism of the fully connected layer is used to calculate the confidence score of each genre. The output layer uses the softmax function to map the output of the hidden layer to the probability of each genre label to which the music file belongs. Finally, the genre tag with the highest probability is selected as the genre tag of the music file.

3.4. Instrument Recognition. This study uses the music signal characteristics of traditional Chinese musical instruments to identify musical instruments. We regard the 2-second-segment musical instrument music signal as a sample, use the MFCC of the sample as the input feature, and input it into a deep belief network with H hidden layers (as shown in Figure 5), and through the output layer, the softmax layer outputs the predicted label of the musical instrument.

4. Experiments and Results

4.1. Cross-Entropy Cost Function. The essence of neural network training is to continuously iterate to minimize the loss function and the process of model parameter convergence. This study uses the cross-entropy loss function to describe the difference between the predicted value output by the network model and the target expected value. The output layer of the network model calculates the probability of each genre through the softmax function and then calculates the cross-entropy loss function. The definition of the cross-entropy loss function is as follows:

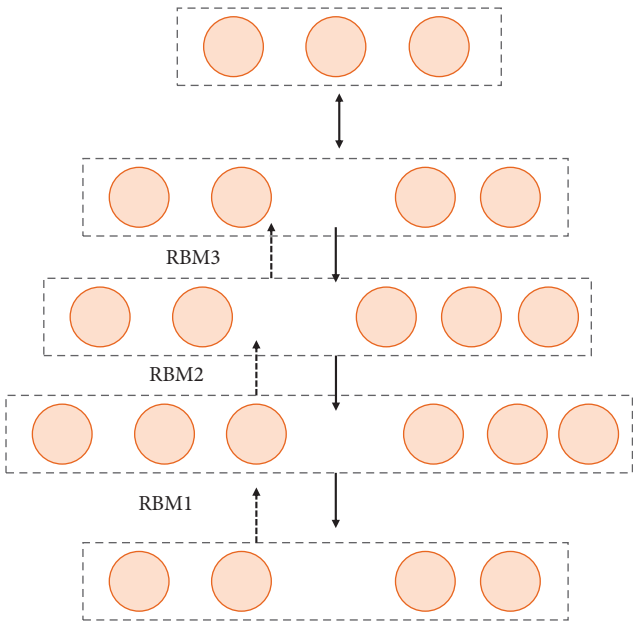


Figure 5: DBM-based recognition and classification network structure diagram of traditional Chinese musical instruments.

$$C = \frac{1}{n}\sum_{x} y \ln(a) + (1-y)\ln(1-a), \tag{5}$$

where C represents the loss, n is the number of samples, x is the input sample, a is the output predicted value of the network model input x, and a is the target expected value of the network model input x.

4.2. Adam. In the process of training the network model, the size of the learning rate has an important impact on the improvement of the model's performance, and the learning rate is one of the hyperparameters that are difficult to set. This article uses Adam optimization algorithm as the optimization method of the network model. Adam algorithm is an adaptive learning rate algorithm, which has excellent performance in practice and is widely used. The Adam algorithm designs independent adaptive learning rates for different parameters by calculating the first-order moment estimation and the second-order moment estimation of the gradient. The calculation formula for adjusting the network parameters is as follows:

$$\begin{aligned} v_t &= k_1 v_{t-1} + (1-k_1) g_t, \\ s_t &= k_2 s_{t-1} + (1-k_2) g_t^2, \\ \hat{v}_t &= \frac{v_t}{1-k_1^t}, \\ \hat{s}_t &= \frac{s_t}{1-k_2^t}, \\ \Delta\theta &= \alpha \frac{\hat{v}_t}{\sqrt{\hat{s}_t + \varepsilon}}. \end{aligned} \tag{6}$$

4.3. Evaluation Environment. In order to carry out the experiment smoothly, we prepare the experimental data in advance. This article downloads genre-labeled MIDI music files from the Internet dedicated to sharing music, constructs a real data set, and collects a total of 2000 music files. There are 5 genres in the data set, including classical, country, dance, folk, and metal. The number of 1VIIDI music files of each genre is shown in Table 1.

4.4. Experimental Results. A special Gaussian convolution kernel is used to convolve along the diagonal of the self-similar matrix to obtain the novelty curve. After smoothing the novelty curve, the peak point is extracted from it and used as the time point for segment division. The smoothed novelty curve and the extracted peak points are shown in Figure 6.

According to the experimental settings, 6 groups of comparative experiments were carried out. The brief description of the experimental settings is shown in Table 2. By comparing Experiment 1 and Experiment 2, it can be concluded that the classification effect of the extracted feature set input to BP neural network for classification experiment is far lower than the classification effect of the 11 features explored and selected in Experiment 2 according to the genre classification task input to BP neural network. It can be seen that the extracted music features are not suitable for the classification task of music genres in this study, which indicates that feature extraction is not easy to be universal, and feature sets usually need to be constructed according to the actual classification task. Meanwhile, the validity of the feature sets selected in this study in the classification task of music genres is verified.

Comparing Experiment 2 and Experiment 3, we can obtain that, in Experiment 3, we divide the MIDI file into sections, use the section as the analysis unit, extract the features of the section with the same feature combination, form the section feature sequence, and input it into the classification network. Bi-GRL can learn the deeper expression of music about time sequence and semantic information from the input music segment feature sequence, which can effectively improve the accuracy of music classification, and the classification effect is better than the traditional 1VIIDI music classification method based on BP neural network.

Comparing Experiment 4, Experiment 5, and Experiment 6, the music segment is divided into different methods, and the extracted music segment feature sequence will affect the final classification performance. In Experiment 5 and Experiment 6, the music was divided into segments with equal time intervals of 5 seconds and 10 seconds, and the final classification accuracy of the experiment was lower than that obtained in Experiment 4 using the segment division method introduced in this article. The possible reason is that the development of music melody is a process of repetition and change, and there is a certain transition boundary. The division of music with equal duration does not take into account this music characteristic, and the extracted segment feature sequence cannot describe the

Table 1: Number of music files in five genres.

Classical	Country	Pop	Rock	Metal	Total
400	386	375	440	399	2000

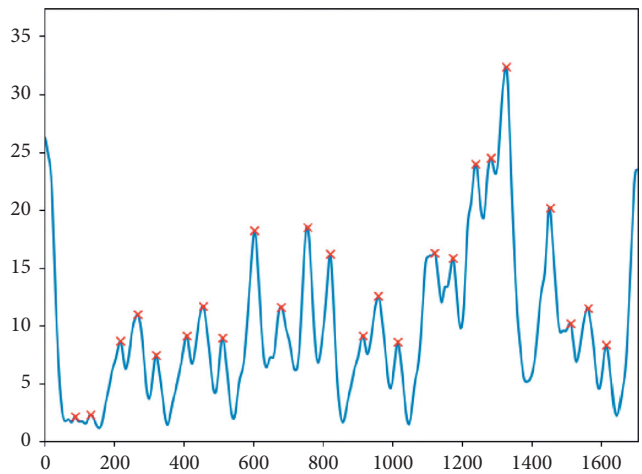


Figure 6: Novelty curve and peak point.

Table 2: Comparison of experimental results of music classification.

Experiment number	Category	Acc
1	BP neural network + local features	0.75
2	BP neural network + global features	0.86
3	Bi-GRU + dense	0.88
4	**Bi-GRU + attention + dense**	**0.91**
5	Bi-GRU + attention + dense (5 seconds)	0.87
6	Bi-GRU + attention + dense (10 seconds)	0.88

music well, so it affected classification performance. In Experiment 4, this study finds the mutation points of music playing to divide the music segment, which can achieve a higher classification effect. The experimental results verify the effectiveness of the music segmentation method used in this study.

5. Conclusion

In this study, we propose a method of music genre classification based on deep learning. According to the characteristic sequence of the input music segment, the cyclic neural network and attention mechanism are studied, and the Bi-GRU and attention mechanism are used to design the classification network model. Bi-GRU is good at processing sequence data. It can learn the contextual semantics and deep features of music from the sequence feature sequence. The attention mechanism is added to automatically assign different attention weights to the features learned by Bi-GRU from different segments and learn more significant music features, thereby improving the accuracy of classification. In addition, this study also proposes a recognition and classification algorithm for traditional Chinese musical instruments based on deep belief networks. The experimental results of the study have achieved credible results.

Data Availability

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

Conflicts of Interest

The author has no conflicts of interest regarding the publication of this study.

References

[1] J. A. Sloboda and P. N. Juslin, "Psychological perspectives on music and emotion," *Music and Emotion: Theory and Research*, pp. 71–104, Oxford University Press, Oxford, UK, 2001.

[2] A. Kresovich, M. K. Reffner Collins, D. Riffe, and F. R. D. Carpentier, "A content analysis of mental health discourse in popular rap music," *JAMA Pediatrics*, vol. 175, no. 3, pp. 286–292, 2021.

[3] T. Eerola, J. K. Vuoskoski, H.-R. Peltola, V. Putkinen, and K. Schäfer, "An integrative review of the enjoyment of sadness associated with music," *Physics of Life Reviews*, vol. 25, pp. 100–121, 2018.

[4] Y. Li, W. Hu, and Y. Wang, "Music rhythm customized mobile application based on information extraction," in *Proceedings of the 4th International Conference on Smart Computing and Communication*, pp. 304–309, Birmingham, UK, October 2019.

[5] F. Medhat, D. Chesmore, and J. Robinson, "Automatic classification of music genre using masked conditional neural networks," in *Proceedings of the 2017 IEEE International Conference on Data Mining (ICDM)*, pp. 979–984, IEEE, New Orleans, LA, USA, November 2017, In press.

[6] S. Vishnupriya and K. Meenakshi, "Automatic music genre classification using convolution neural network," in *Proceedings of the 2018 International Conference on Computer Communication and Informatics (ICCCI)*, pp. 1–4, IEEE, Coimbatore, India, January 2018.

[7] S. Shetty and S. Hegde, "Automatic classification of carnatic music instruments using MFCC and LPC," *Data Management, Analytics and Innovation*, Springer, Singapore, pp. 463–474, 2020.

[8] Y. T. Chen, C. H. Chen, S. Wu, and C. C. Lo, "A two-step approach for classifying music genre on the strength of AHP weighted musical features," *Mathematics*, vol. 7, no. 1, 19 pages, 2019, In press.

[9] R. Liu, X. Ning, W. Cai, and G. Li, "Multiscale dense cross-attention mechanism with covariance pooling for hyperspectral image scene classification," *Mobile Information Systems*, vol. 2021, Article ID 9962057, 15 pages, 2021.

[10] C. Yan, G. Pang, X. Bai, Z. Zhou, and L. Gu, "Beyond triplet loss: person re-identification with fine-grained difference-aware pairwise loss," *IEEE Transactions on Multimedia*, 2021.

[11] Y. Ding, X. Zhao, Z. Zhang, W. Cai, and N. Yang, "Multiscale graph sample and aggregate network with context-aware learning for hyperspectral image classification," *IEEE Journal*

of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 14, pp. 4561–4572, 2021, In Press.

[12] Y. Tong, L. Yu, S. Li, J. Liu, H. Qin, and W. Li, "Polynomial fitting algorithm based on neural network," *ASP Transactions on Pattern Recognition and Intelligent Systems*, vol. 1, no. 1, pp. 32–39, 2021.

[13] X. Ning, K. Gong, W. Li, L. Zhang, X. Bai, and S. Tian, "Feature refinement and filter network for person re-identification," *IEEE Transactions on Circuits and Systems for Video Technology*, 2020.

[14] W. Cai, Z. Wei, R. Liu, Y. Zhuang, Y. Wang, and X. Ning, "Remote sensing image recognition based on multi-attention residual fusion networks," *ASP Transactions on Pattern Recognition and Intelligent Systems*, vol. 1, no. 1, pp. 1–8, 2021.

[15] X. Zhang, Y. Yang, Z. Li, X. Ning, Y. Qin, and W. Cai, "An improved encoder-decoder network based on strip pool method applied to segmentation of farmland vacancy field," *Entropy*, vol. 23, no. 4, p. 435, 2021.

[16] X. Ning, X. Wang, S. Xu et al., "A review of research on co-training," *Concurrency and Computation: Practice and Experience*, 2021.

[17] D. Bisharad and R. H. Laskar, "Music genre recognition using convolutional recurrent neural network architecture," *Expert Systems*, vol. 36, no. 4, Article ID e12429, 2019.

[18] S. Iloga, O. Romain, and M. Tchuenté, "A sequential pattern mining approach to design taxonomies for hierarchical music genre recognition," *Pattern Analysis and Applications*, vol. 21, no. 2, pp. 363–380, 2018.

[19] S. Oramas, F. Barbieri, O. Nieto, and X. Serra, "Multimodal deep learning for music genre classification," *Transactions of the International Society for Music Information Retrieval*, vol. 1, no. 1, pp. 4–21, 2018.

[20] H. Bahuleyan, "Music genre classification using machine learning techniques," 2018, http://arxiv.org/abs/1804.01149.

[21] R. Yang, L. Feng, H. Wang, J. Yao, and S. Luo, "Parallel recurrent convolutional neural networks-based music genre classification method for mobile devices," *IEEE Access*, vol. 8, pp. 19629–19637, 2020, In press.

[22] W. Cai and Z. Wei, "Remote sensing image classification based on a cross-attention mechanism and graph convolution," *IEEE Geoscience and Remote Sensing Letters*, pp. 1–5, 2020, In Press.

[23] W. Cai, B. Liu, Z. Wei, M. Li, and J. Kan, "TARDB-Net: triple-attention guided residual dense and BiLSTM networks for hyperspectral image classification," *Multimedia Tools and Applications*, vol. 80, no. 7, pp. 11291–11312, 2021.

[24] Z. Chu, M. Hu, and X. Chen, "Robotic grasp detection using a novel two-stage approach," *ASP Transactions on Internet of Things*, vol. 1, no. 1, pp. 19–29, 2021.

[25] W. Sun, P. Zhang, Z. Wang, and D. Li, "Prediction of cardiovascular diseases based on machine learning," *ASP Transactions on Internet of Things*, vol. 1, no. 1, pp. 30–35, 2021.

[26] L. Sun, W. Li, X. Ning, L. Zhang, X. Dong, and W. He, "Gradient-enhanced softmax for face recognition," *IEICE Transactions on Information and Systems*, vol. E103.D, no. 5, pp. 1185–1189, 2020.

[27] Y. Zhang, W. Li, L. Zhang, X. Ning, L. Sun, and Y. Lu, "AGCNN: adaptive gabor convolutional neural networks with receptive fields for vein biometric recognition," *Concurrency and Computation: Practice and Experience*, Article ID e5697, 2020, In press.

Hindawi
Scientific Programming
Volume 2021, Article ID 3045418, 8 pages
https://doi.org/10.1155/2021/3045418

Research Article

Recognition of Basketball Player's Shooting Action Based on the Convolutional Neural Network

Rui Liu,[1] Ziqi Liu,[2] and Shuyong Liu[3]

[1]*Lingnan Normal University, Zhanjiang 524048, Guangdong, China*
[2]*Institute Genetics and Developmental Biology, Chinese Academy of Sciences, Beijing 100101, China*
[3]*P. E. Scientific College, Harbin Normal University, Harbin 150025, China*

Correspondence should be addressed to Shuyong Liu; liushuyong2251@126.com

Received 30 April 2021; Revised 21 May 2021; Accepted 4 June 2021; Published 22 June 2021

Academic Editor: Shah Nazir

In the field of basketball, the formulation of the existing training plan mainly relies on the coaches' artificial observation and personal experience, which is inevitably subjective. The application of body domain network technology in athletes' training and recognition of athletes' postures can help coaches to assist decision-making and greatly improve athletes' competitive ability. The human movements reflected in basketball are more complex which need deep understanding. The accuracy of basketball players' shooting movements recognition plays a positive and important role in basketball games and training practice. Based on the prior knowledge of the convolutional neural network study, environment light conditions change the dynamic characteristics of basketball image analysis, capture images of the basketball goal algorithm of minimum circumscribed rectangle of the object, and based on the convolutional neural network, introduce two types of prior knowledge, one kind is based on the feature matching method that defined a priori knowledge, while another kind is based on training the convolution neural network model. The test results of the network model are taken as the prior knowledge, and then, a convolutional neural network dynamic target recognition model is constructed based on the prior knowledge. The construction process of the model is organized as the basketball target image is collected under any illumination conditions, the convolutional neural network model is trained with the convolutional neural network as the input data, and the standard illumination conditions are determined according to the test results of the network model. Then, put it into the trained network model to test and get the recognition results of basketball players' shooting movements. The research is validated with performing experiments and the results revealed the success of the study.

1. Introduction

Basketball is a collective sport that put the ball into the opponent's basket to score and prevents the opponent from getting the ball and scoring under certain rules. Compared with other ball games, basketball has a variety of techniques [1–3], diverse tactics [4], and strong skills of players [5–7]. It also reflects the characteristics of individual combat and coordination. In a basketball game [8], the player's basketball skill level has a very obvious impact on the entire team. If the player's basketball level is insufficient, the team's weaknesses will be exposed, and the defense and offense level will be greatly reduced, which is not conducive to the team's performance in the basketball game. It is very necessary to carry out scientific and reasonable basketball training [9–11] for athletes. In traditional basketball training, coaches make training plans based on athletes' training and competition. This method relies on the training theory mastered by the coach and his own experience and has a certain degree of subjectivity. In addition, it is difficult to avoid wrong actions and possible damage to athletes' muscles [12–14], soft tissues [15], and bones through scientific observations during training, which will affect the normal training and even shorten the athlete's lifespan [16, 17]. From the perspective of training quality evaluation, the evaluation work is performed manually. Coaches need to calculate the training performance of each athlete with reference to different test standards. This method also has some drawbacks. First of all,

the testing of athletes is performed manually, which requires a lot of time for the coaches, the process is complicated, and the accuracy is poor; second, the testing methods have limitations, and it is difficult to directly measure some important sports parameters such as acceleration and angular velocity. Information such as muscle tension, sprinting ability, and body balance cannot be measured during exercise. Third, coaches lack scientific evaluation methods, and it is difficult to make corresponding decision-making plans based on test data. Therefore, if the athlete's sports parameters can be accurately collected in real-time, the athlete's motion posture can be analyzed and recognized [18], and the training effect evaluation model can be constructed; the coach can make reasonable adjustments to the training program and scientifically evaluate the training quality. The improvement of athletes' competitive ability and coaches' decision-making ability is of great significance.

There are two main recognition methods for human body gesture recognition [19–22], namely, the recognition technology based on image analysis and the recognition technology based on inertial sensors [23]. Image analysis-based recognition technology mainly uses video, image, and other information to recognize human posture. Therefore, it is necessary to place a camera and other monitoring equipment in the detection environment in advance to collect data. Image analysis technology is applied to human posture recognition earlier. The technology is relatively mature; the early ones are based on monocular video research and multiview video research. In addition, Iosifidis et al. [24] used multiple cameras to perform multiangle detection of human action poses and used neural network algorithms to image and video data that are trained and classified. Although this method can identify people's daily actions more accurately, it is difficult to realize real-time monitoring due to the large amount of data contained. There are still many shortcomings in the recognition technology based on image analysis. The equipment requires high accuracy, and the equipment is relatively heavy and not portable. Video capture is prone to blind spots, and some places are not easy to be observed. The monitoring range is obviously limited, and the image captures the large amount of data that can easily lead to insufficient storage and fail to achieve the purpose of real-time monitoring.

Therefore, a basketball player's shooting action recognition algorithm [25] based on a priori knowledge of the convolutional neural network [26–30] analyzes the dynamic characteristics of the basketball image under the conditions of light changes and proposes the smallest enclosing rectangle algorithm for intercepting the circular target object in the image [31]. Then, a fast recognition algorithm of dynamic target based on the convolutional neural network based on prior knowledge is proposed, and a blue-gray basketball is taken as an example to carry out related experiments. The main contributions of this study are as follows:

(i) In this study, neural network technology is introduced innovatively in the prediction task of graphic design multimedia communication, and it has achieved more accurate results than traditional prediction algorithms.

(ii) Based on the analysis of the traditional recursive model, this study uses a special form of the recursive neural network LSTM to predict the law of multimedia communication in graphic design, carries out a comparative experiment, and the experimental results prove the superiority of the proposed algorithm.

2. Related Work

2.1. Digital Image Processing. Gamma correction [32, 33] performs a nonlinear operation on the gray value of the input image, so that the gray value of the input image and the gray value of the output image have an exponential relationship, thereby improving the contrast effect of the image. The calculation equation is as follows:

$$V_{\text{out}} = V_{\text{in}}^{\gamma}, \tag{1}$$

where γ is the gamma coefficient, V_{in} is the gray value of the input image, V_{out} is the gray value of the output image, and $V_{\text{in}}, V_{\text{out}} \in [0, 1]$.

It can be seen from Figure 1 that

(1) When $\gamma < 1$, the gray value of the darker area of the image increases, and the image becomes brighter as a whole.

(2) When $\gamma > 1$, the gray value of the brighter area of the image becomes smaller, and the image becomes darker as a whole. The gray value of the brighter area decreases, the gray value of the overall darker area decreases, the overall gray value becomes smaller, and the overall image becomes darker.

2.2. CNNs. As shown in Figure 2, a typical convolutional neural network model generally composed of an input layer, a convolutional layer, an activation layer, a pooling layer, a fully connected layer, and an output layer. The convolutional neural network proposes image features by convolution operation on a large amount of input data, continuously reduces the dimensionality of the image through the pooling operation, summarizes the proposed image features through the fully connected layer, and then realizes the image characteristics through the softmax function of the output layer. They conduct classification and finally output the classification results.

In mathematics, convolution is an important analytical operation. It is a mathematical operator that generates a third function through the sum of two functions and represents the area of the overlapping part between function two and the shifted function W. Its calculation equation is as follows:

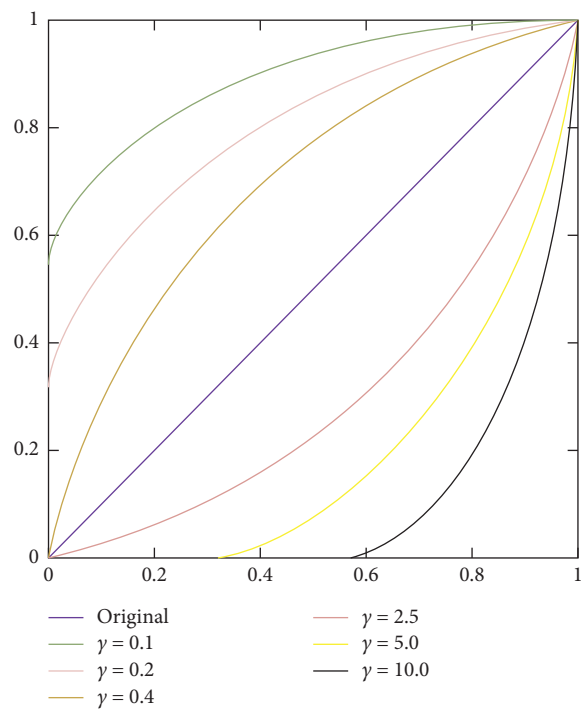


Figure 1: Different γ value gamma conversion curves.

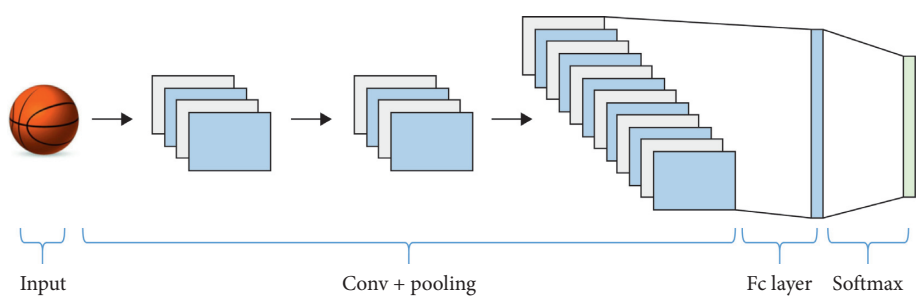


Figure 2: The typical convolutional neural network model.

$$s(t) = f(t) * g(t) = \sum_{\tau=-\infty}^{\infty} f(\tau) g(t-\tau), \tag{2}$$

$$s(t) = x(t) * w(t) = \sum_{\tau=-\infty}^{\infty} x(t-\tau) w(\tau). \tag{3}$$

The integral form is as follows:

$$s(t) = x(t) * w(t) = \int_{-\infty}^{\infty} x(t-\tau) w(\tau) \mathrm{d}\tau = \int_{-\infty}^{\infty} x(\tau) w(t-\tau) \mathrm{d}\tau. \tag{4}$$

The matrix expression is as follows:

$$s(t) = (X * W)(t), \tag{5}$$

where $*$ represents the convolution, X represents the input, and W represents the convolution kernel. In addition, the two-dimensional convolution calculation equation is as follows:

$$s(i, j) = (X * W)(i, j) = \sum_{m} \sum_{n} x(i-m, j-n) w(m, n), \tag{6}$$

where m and n are the size of the convolution kernel, respectively.

3. Methodology

So far, various approaches have been presented in the field of research. Figure 1 represents different γ value gamma conversion curves.

Figure 2 is the representation of the typical convolutional neural network model.

3.1. Analysis of Dynamic Characteristics of Shooting Action Images under the Condition of Changing Illumination. This section takes two actual pictures taken as examples for analysis. The two images in Figure 3 are pictures of basketballs in the same position and changing lighting conditions. According to observations, the light distribution of the two pictures is uneven, and both show that the upper part of the picture has a higher brightness than the lower part of the picture; the overall brightness of the first picture is dark, and the overall brightness of the second picture is moderate.

The color description of the two pictures in Figure 3 in the HIS color model is realized by Matlab simulation, and the simulation results are shown in Figures 4 and 5. Obviously, the brightness of Figure 5 is higher than that of Figure 4, so under the condition of changing illumination, the color characteristic of the image of the target object is dynamically changed.

3.2. Minimum Bounding Rectangle Algorithm. In this study, the interception of basketball goal object in the image of the biggest external rectangular algorithm is the first choice of target image in different light conditions that appropriate gamma coefficient, reoccupy gamma correction to school optical processing of images, and select the appropriate Gaussian template, using a Gaussian filter to filter the image, and roundness based on Hough transform detection basketball round objects in the image. The maximum outer rectangle of the circle is determined, and the vertex coordinates of the rectangle are calculated.

As shown in Figure 6, the image coordinates of vertices are A, B, C, and D, and the calculation equation is as follows:

$$\begin{cases} A_u = C_u = X_{\text{center}} - r, \\ B_u = D_u = X_{\text{center}} + r, \\ A_v = B_v = Y_{\text{center}} - r, \\ C_v = D_v = Y_{\text{center}} + r. \end{cases} \tag{7}$$

According to the vertex coordinates, intercept the smallest bounding rectangle of the circular target object in the original image.

3.3. Fast Recognition Algorithm for Shooting Action. Under the condition of changing ambient light, the characteristics of basketball players' shooting images are changing dynamically, which makes it very difficult to recognize the basketball object. The target recognition method based on convolution typical neural network models is often a target of acquisition under the condition of different illumination object image sample sets and into the convolutional neural network training, which requires the collected images of the sample set to cover all light shooting images under the condition of basketball player, so the convolution neural network can learn the image characteristics under different light conditions. The trained convolutional neural network model can contain the shooting image features of basketball players under all illumination conditions, but this requires a large number of image sample sets, and the training of the convolutional neural network model needs a lot of time, and the target object recognition accuracy is not high under illumination changes.

In order to improve the rapidness and accuracy of object recognition under the condition of illumination change, this study proposes a convolutional neural network based on prior knowledge to recognize basketball player's shooting image. Based on the convolutional neural network, the algorithm introduces two kinds of prior knowledge and combines prior knowledge with the convolutional neural network to reduce the randomness of the network model in searching for image features, reduce useless exploration, shorten the training time of the convolutional neural network model, and improve the accuracy of target recognition.

3.3.1. Priori Knowledge. A standard illumination condition is assumed. Under this standard illumination condition, prior knowledge is introduced to collect athletes' shooting image set S, and the collected pictures have certain characteristics. Under any certain illumination conditions, the image P of the target object is randomly collected without introducing any prior knowledge, and the image P is rotated for a certain number of times, so that the image features of the image P collected under any certain illumination conditions are included in the image features of the image set S collected under standard illumination conditions.

To satisfy the above conditions, the operation on image P is as follows: rotate the image $x\,(x = 2k)$ times and denote the maximum number of rotations as $x\_\max$. When the number of rotations x is not equal to 0, each rotation is a radian, and the radian value a is the ratio of the whole circle radian a to the maximum number of rotations $x\_\max$. When the number of rotations x is equal to 0, the gray value a is 0. The new test picture M_k is obtained by rotating the image P by $x\alpha$ radian through the rotation function, where $R_{xa}(\cdot)$, and its calculation equation is as follows:

$$x\_\max = \max(1, 2, \cdots, x), \tag{8}$$

$$\begin{cases} x\_\max \times \alpha = 2\pi, & x \neq 0, \\ x\_\max \times \alpha = 0, & x = 0, \end{cases} \tag{9}$$

$$M_k = R_{x\alpha}(P), \tag{10}$$

that is, the image features of the newly obtained test picture $M_k\,(k = 1, 2, \cdots, x\_\max)$ and the original picture P are included in the image features of image set S. A new picture

Figure 3: The original picture of the basketball under different lighting conditions.

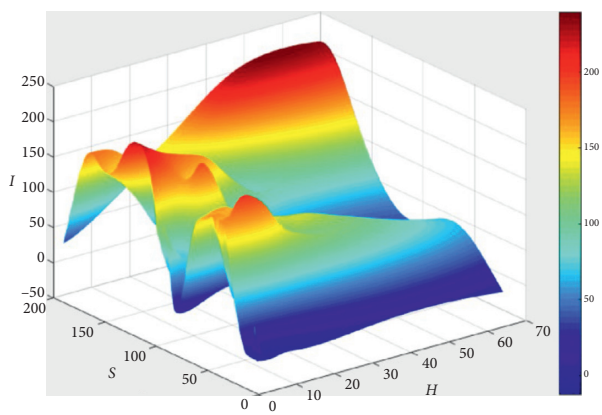


Figure 4: The original picture of the basketball under different lighting conditions.

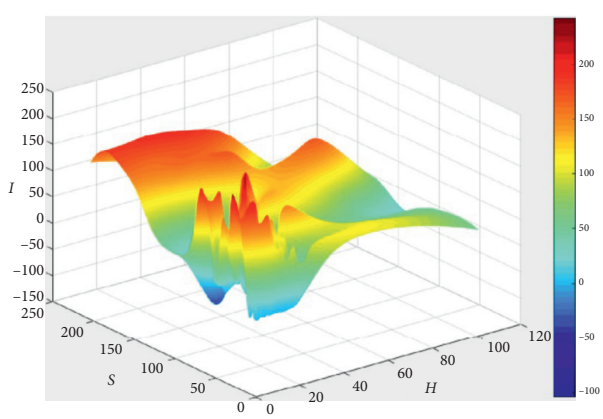


Figure 5: The original picture of the basketball under different lighting conditions.

set is composed of k test set pictures $M_k\,(k = 1, 2, \cdots, x\_\max)$ and the original picture P.

3.3.2. Model and Training. This is shown in Figure 7. The first layer is the input layer, the second, fourth, sixth, and eighth layers are the convolution layer, the third, fifth, seventh, and ninth layers are the pooling layer, the tenth and

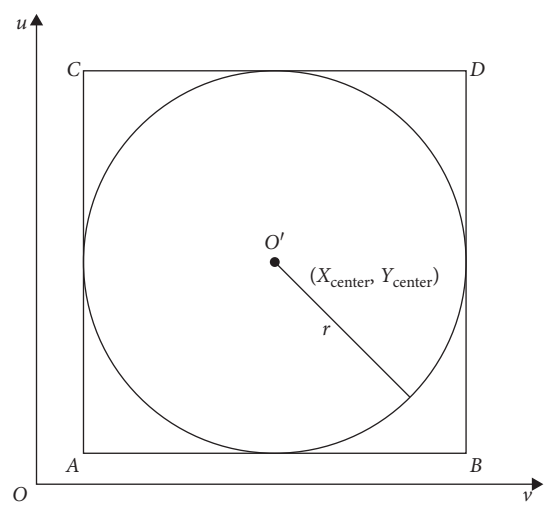


Figure 6: The image coordinate system.

eleventh layers are the full connection layer, and the twelfth layer is the output layer.

We collect sports basketball shooting images as the training set and validation set of the convolutional neural network and use the training set as the input data for training the convolutional neural network. The image features are extracted through the convolutional layer and the pooling layer, and the fully connected layer will be extracted. The output image features are combined, and the output value is calculated through the output layer. Introduce the dropout method, select the appropriate loss function, add a regularization term to the loss function, and iterate repeatedly through the gradient descent algorithm to update the network parameters until the end of the iteration. Finally, the validation set is used to verify the effect of the trained convolutional neural network model.

4. Experiments and Results

4.1. Experimental Setup. The experimental software platform uses a Win-based desktop system, and the hardware platform uses an Intel Core i5-7200 processor with 4 GB of running memory. The algorithm is written in Python

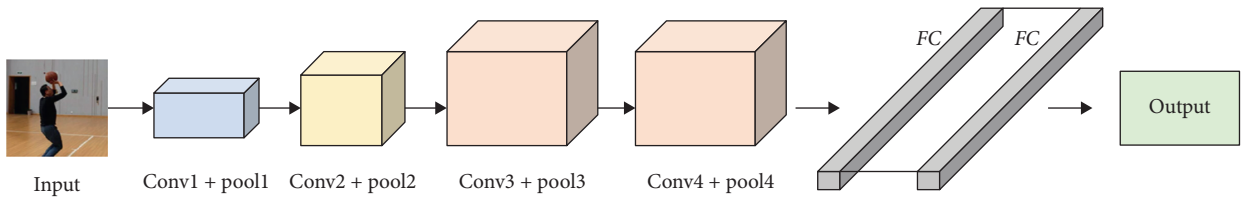


Figure 7: The structure diagram of the convolutional neural network model.

Table 1: The parameter table corresponding to the experimental environment.

Environments	Parameter
OS	Win7
CPU	Inter(R) Core(TM) i7
RAM	8.00 GB
Programming language	Python
Deep-learning tools	PyTorch

Table 2: The parameter table corresponding to the experimental environment.

Behavior	SVM	BN	Ours
Catch the ball	0.9652	0.9290	0.9782
Pass	0.9352	0.9352	0.9825
Shot	0.9125	0.9145	0.9745

language and TensorFlow deep-learning tool, and the experimental environment configuration is given in Table 1.

4.2. Experimental Dataset Collection. During the data collection process, 9 types of walking, running, jumping without the ball and standing dribble, walking dribble, running dribble, shooting, passing, and receiving were completed for 8 male testers. The actions are collected separately. Each action is repeated 50 times. There are a total of 5,000 samples. Among them, the upper limb movements when holding the ball include standing dribbling, walking dribbling, running dribbling, shooting, passing, and receiving, a total of 2,400, and a total of 2,600 lower limb movements, including those without the ball walking dribbling, running dribbling, and shooting when walking, running, jumping, and holding the ball. During the sampling process, each tester completed the prescribed actions as required, and the monitoring personnel recorded the number of actions.

4.3. Experimental Results. The completion of basketball shooting movement is mainly through the coordination of the upper and lower limbs of players to complete the overall movement; so in the recognition of basketball movements, we need to discuss the upper and lower limbs movements, respectively. In the process of data collection, according to the different placement positions of sensor nodes in the body, the collected upper limb movement data and lower limb movement data are discussed and identified, respectively. So, in view of the onset of action, the structure classifier is to identify, respectively, through the combination of the onset of action to determine the athletes do action, and this study analyzes the characteristics of the classification of different classifiers, compares the different classification of the gesture recognition classifier for basketball performance, according to different body movement data, builds the corresponding classification algorithm for training, and the recognition effect is analyzed from two aspects of accuracy and recall rate, as given in Table 2. The whole experiment process is carried out in the same environment, and the ten-fold cross-validation method is adopted.

5. Conclusion

In this study, the convolutional neural network is based on prior knowledge studies, the dynamic characteristics analysis of basketball images when the environmental lighting conditions change, and then intercepts the minimum bounding rectangle algorithm of the basketball target object in the image. The research introduces two types of priors based on the convolutional neural network knowledge, one is to define prior knowledge based on the feature matching method, and the other is to pretrain the convolutional neural network model. The study uses the test results of the network model as prior knowledge and then constructs a convolution based on the prior knowledge neural network dynamic target recognition model. The model construction process is to collect basketball target images under arbitrary lighting conditions, use them as input data of the convolutional neural network to train the convolutional neural network model, and determine the standard lighting conditions according to the test results of the network model. Then, put it into the trained network model to test and get the recognition result of the basketball player's shooting action. Use standard lighting conditions as prior knowledge, and under standard lighting conditions, introduce prior knowledge defined based on feature matching methods, collect target object images, and use them as input data for convolutional neural networks to train convolutional neural network models. At the same time, introduce the prior knowledge based on the feature matching method to process the test image data, put it into the trained network model for testing, and get the target recognition result.

Data Availability

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

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

[1] H. Faal Moghanlo, F. S. Hosseini, and F. Mikaili Manee, "Comparison the impact of spark motor program and basketball techniques on improving gross motor skills in educable intellectually disabled boys," *Journal of Ardabil University of Medical Sciences*, vol. 14, no. 3, pp. 274–284, 2014.

[2] W. Xinhua, "The theoretical thinking and breakthrough in basketball techniques and tactics," *Journal of Guangzhou Physical Education Institue*, vol. 3, 1996.

[3] A. Mardiana, M. Doewes, and S. K. Purnama, "Development of learning media based on video tutorial on basketball based shooting techniques," *Journal of Education, Health and Sport*, vol. 9, no. 5, pp. 298–303, 2019.

[4] C. Minghua, Z. Tuxuan, and S. Shuijun, "The development tendency of positional attack tactics in the basketball match of rio olympic games," *Bulletin of Sport Science & Technology*, vol. 1, 2017.

[5] N. Apostolidis and Z. Emmanouil, "The influence of the anthropometric characteristics and handgrip strength on the technical skills of young basketball players," *Journal of Physical Education and Sport*, vol. 15, no. 2, pp. 330–337, 2015.

[6] G. Fiorilli, E. Iuliano, G. Aquino et al., "Mental health and social participation skills of wheelchair basketball players: a controlled study," *Research in Developmental Disabilities*, vol. 34, no. 11, pp. 3679–3685, 2013.

[7] C. E. Silva, H. M. Carvalho, C. E. Gonçalves et al., "Growth, maturation, functional capacities and sport-specific skills in 12-13 year-old-basketball players," *Journal of Sports Medicine and Physical Fitness*, vol. 50, no. 2, pp. 174–181, 2010.

[8] M. Perše, M. Kristan, S. Kovačič, G. Vučkovič, and J. Perš, "A trajectory-based analysis of coordinated team activity in a basketball game," *Computer Vision and Image Understanding*, vol. 113, no. 5, pp. 612–621, 2009.

[9] P. G. Montgomery, D. B. Pyne, and C. L. Minahan, "The physical and physiological demands of basketball training and competition," *International Journal of Sports Physiology and Performance*, vol. 5, no. 1, pp. 75–86, 2010.

[10] O. Stoll, A. Lau, and J. Stoeber, "Perfectionism and performance in a new basketball training task: does striving for perfection enhance or undermine performance?" *Psychology of Sport and Exercise*, vol. 9, no. 5, pp. 620–629, 2008.

[11] N. A. Maffiuletti, C. Gometti, I. G. Amiridis, A. Martin, M. Pousson, and J. C. Chatard, "The effects of electromyostimulation training and basketball practice on muscle strength and jumping ability," *International Journal of Sports Medicine*, vol. 21, no. 6, pp. 437–443, 2000.

[12] A. Moreira, K. Nosaka, J. A. Nunes, L. Viveiros, A. Z. Jamurtas, and M. S. Aoki, "Changes in muscle damage markers in female basketball players," *Biology of Sport*, vol. 31, no. 1, pp. 3–7, 2014.

[13] C. Alfredo, F. Diego, M. Juan, S. Calvo, and C. G. A. Jesús, "Effect of magnesium supplementation on muscular damage markers in basketball players during a full season," *Journal of Magnesium Research*, vol. 30, pp. 61–70, 2017.

[14] C. G. Marques, V. C. Santos, A. C. Levada-Pires et al., "Effects of DHA-rich fish oil supplementation on the lipid profile, markers of muscle damage, and neutrophil function in wheelchair basketball athletes before and after acute exercise," *Applied Physiology, Nutrition, and Metabolism*, vol. 40, no. 6, pp. 596–604, 2015.

[15] N. Tamai, T. Minematsu, T. Maeda, K. Yabunaka, and H. Sanada, "The relationship between skin ultrasound images and muscle damage using skin blotting in wheelchair basketball athletes," *Spinal Cord*, vol. 8, 2020.

[16] B. Calvo, E. Peña, M. A. Martinez, and M. Doblaré, "An uncoupled directional damage model for fibred biological soft tissues. formulation and computational aspects," *International Journal for Numerical Methods in Engineering*, vol. 69, no. 10, pp. 2036–2057, 2007.

[17] S. Lemez, N. Wattie, and J. Baker, "Do "big guys" really die younger? an examination of height and lifespan in former professional basketball players," *PloS One*, vol. 12, no. 10, Article ID e0185617, 2017.

[18] T. Hachaj, M. Piekarczyk, and M. Ogiela, "Human actions analysis: templates generation, matching and visualization applied to motion capture of highly-skilled karate athletes," *Sensors*, vol. 17, no. 11, p. 2590, 2017.

[19] B. W. Hwang, S. Kim, and S. W. Lee, "A full-body gesture database for automatic gesture recognition," in *Proceedings of the 7th International Conference on Automatic Face and Gesture Recognition (FGR06)*, pp. 243–248, IEEE, Washington, DC, USA, April 2006.

[20] T. Gonzalez-Sanchez and D. Puig, "Real-time body gesture recognition using depth camera," *Electronics Letters*, vol. 47, no. 12, pp. 697-698, 2011.

[21] F. Noroozi, D. Kaminska, C. Corneanu, T. Sapinski, S. Escalera, and G. Anbarjafari, "Survey on emotional body gesture recognition," *IEEE Transactions on Affective Computing*, vol. 12, no. 2, 2018.

[22] O. Patsadu, C. Nukoolkit, and B. Watanapa, "Human gesture recognition using Kinect camera," in *Proceedings of the 2012 Ninth International Conference on Computer Science and Software Engineering (JCSSE)*, pp. 28–32, IEEE, Bangkok, Thailand, May 2012.

[23] J. Chen, C. Du, Y. Zhang, P. Han, and W. Wei, "A clustering-based coverage path planning method for autonomous heterogeneous UAVs," *IEEE Transactions on Intelligent Transportation Systems*, vol. 11, pp. 1–11, 2021.

[24] A. Iosifidis, A. Tefas, and I. Pitas, "Multi-view action recognition based on action volumes, fuzzy distances and cluster discriminant analysis," *Signal Processing*, vol. 93, no. 6, pp. 1445–1457, 2013.

[25] A. Schmidt, "Movement pattern recognition in basketball free-throw shooting," *Human Movement Science*, vol. 31, no. 2, pp. 360–382, 2012.

[26] L. Zhang, X. Wang, X. Dong, L. Sun, W. Cai, and X. Ning, "Finger vein image enhancement based on guided tri-Gaussian filters," *ASP Transactions on Pattern Recognition and Intelligent Systems*, vol. 1, no. 1, pp. 17–23, 2021.

[27] X. Zhang, Y. Yang, Z. Li, X. Ning, Y. Qin, and W. Cai, "An improved encoder-decoder network based on strip pool method applied to segmentation of farmland vacancy field," *Entropy*, vol. 23, no. 4, p. 435, 2021.

[28] Y. Tong, L. Yu, S. Li, J. Liu, H. Qin, and W. Li, "Polynomial fitting algorithm based on neural network," *ASP Transactions on Pattern Recognition and Intelligent Systems*, vol. 1, no. 1, pp. 32–39, 2021.

[29] R. Liu, "Multiscale dense cross-attention mechanism with covariance pooling for hyperspectral image scene classification," *Mobile Information Systems*, vol. 2021, 2021.

[30] Y. Zhang, W. Li, L. Zhang, X. Ning, L. Sun, and Y. Lu, "AGCNN: adaptive gabor convolutional neural networks with

receptive fields for vein biometric recognition," *Concurrency and Computation: Practice and Experience*, vol. 21, p. e5697, 2020.

[31] J. Zhang, J. Sun, J. Wang, and X.-G. Yue, "Visual object tracking based on residual network and cascaded correlation filters," *Journal of Ambient Intelligence and Humanized Computing*, vol. 13, 2020.

[32] Y. Cheng, S. Yue, J. Li, L. Deng, and Q. Quan, "Errors of machine translation of terminology in the patent text from English into Chinese," *ASP Transactions on Computers*, vol. 1, no. 1, pp. 12–17, 2021.

[33] L. Liang, Q. Yin, and C. Shi, "Exploring proper names online and its application in English teaching in university," *ASP Transactions on Computers*, vol. 1, no. 1, pp. 24–29, 2021.

Hindawi
Scientific Programming
Volume 2021, Article ID 2287596, 11 pages
https://doi.org/10.1155/2021/2287596

Research Article

Optimization of Post-Based Merit Salary Based on CPI and University Characteristic Data

Yuan Cao [1] and Qing Cao [2]

[1] *Donghua University, Shanghai 201620, China*
[2] *Shandong University of Science and Technology, Qingdao 266590, China*

Correspondence should be addressed to Yuan Cao; caoyuan@dhu.edu.cn

Received 6 April 2021; Accepted 11 June 2021; Published 21 June 2021

Academic Editor: Shah Nazir

In order to improve the rationality and efficiency of post-based merit salary setting for teaching and research staff in colleges and universities, this paper analyses and discusses the influencing factors and determining principles of post-based merit salary for teaching and research staff in colleges and universities based on the consumer price index (CPI) and the characteristics of colleges and universities. A dynamic and optimal post-based merit salary generation model and its determination method are put forward. The model comprehensively takes changes in price levels into account and the key factors such as the teacher's professional technical level, length of service, and performance appraisal results and can dynamically and intelligently calculate the post-based merit salary of teachers. The test results of the examples prove that the post-based merit salary calculated by the model is objective and reasonable, and it is easy to dynamically adjust according to the changes in social environment and the characteristics of colleges and universities. It can provide a scientific method and useful reference for determining post-based merit salary in colleges and universities.

1. Introduction

In order to achieve good pay for good work [1] and to encourage the enthusiasm of the faculty of colleges and universities [2], according to the national salary and income distribution system, all colleges and universities in our country have established a merit salary distribution system that suits their own characteristics [3]. The current merit salary system implemented by colleges and universities divides the salary into four parts: post salary, pay grade salary, merit salary, and allowance [4], while merit salary can be divided into post-based merit salary and rewarding merit salary, and so on. Post-based merit salary is an important component of merit salary. This part mainly reflects post responsibilities and workload matched with the professional technical level of teachers, which not only have an important impact on the income of faculty, but also are a clear response to fairness [5], incentives, and social adaptability [6]. This article researches the optimization methods of post-based merit salary.

The faculty of colleges and universities is divided into four groups: teaching and research staff [7], other specialized technical staff, management staff, and technical workers. Teaching and research staff are the largest group and also contribute the most to the development of colleges and universities. Therefore, this article focuses on researching the post-based merit salary of teaching and research staff (hereinafter, referred to as "teachers") in colleges and universities and establishes their job performance optimization methods and dynamic adjustment mechanisms based on social development and the university's own characteristic data.

There are several kinds of institutions of higher learning in China, such as teaching and research-oriented universities [8], application-oriented universities, and higher vocational colleges. The construction of teaching and research-oriented universities is the focus of China's higher education construction, and the author has long been engaged in salary management and research in teaching and research-oriented

university. For the above-mentioned reasons, this article mainly discussed the post-based merit salary of teachers in teaching and research-oriented universities. Unless otherwise specified, "university" means teaching and research-oriented university in this article.

The main contributions of this article are summarized as follows:

(1) This article studies and proposes a model for dynamic optimization that generates the post-based merit salary of university teachers. This model analyses and discusses the influencing factors and determining principles of post-based merit salary for teachers in universities based on CPI and the characteristics of universities and comprehensively takes changes in price levels into account and the key factors such as the teacher's professional technical level, length of service, and performance appraisal results and can dynamically and intelligently calculate the post-based merit salary of teachers.

(2) In this article, CPI is introduced as a parameter for the generation model of post-based merit salary. A linkage mechanism for post-based merit salary to rise with the price level is established, which makes post-based merit salary change dynamically with the price level, in order to ensure the income level of the university teachers, guarantee the basic living needs of faculty of college and university, and maintain the salary competitiveness of university.

(3) In this article, the effectiveness of the generation model of post-based merit salary proposed is tested with the actual information of 8 teachers, and the results are compared with salary data obtained by traditional methods. The test results prove that the dynamic and optimized generation model of post-based merit salary developed in this article is scientifically valid. This model can calculate reasonable post-based merit salary, which provides a useful reference for salary managers of universities.

2. Analysis on the Status and Problems of Post-Based Merit Salary of University Teachers by Data

The post-based merit salary of university teachers is a kind of compensation based on the post setting result of the university, which mainly reflects the post responsibilities and basic workload of the teachers. Taking a university in Shanghai as an example, the university adjusted the standard of post-based merit salary once in 2013, 2017, and 2020, respectively. The data of the three adjustments are shown in Table 1. Another example is the post-based merit salary data of a university in Nanjing, as shown in Table 2.

Table 1: Post-based merit salary of a university in Shanghai in 2013, 2017, and 2020.

Professional and technical level	2013	2017	2020
2	90,000	140,400	151,200
3	82,200	128,400	138,240
4	74,400	110,400	118,920
5	68,400	88,800	96,000
6	64,200	82,800	89,520
7	60,000	76,800	83,040
8	55,200	65,400	70,800
9	51,600	61,200	66,240
10	48,000	57,000	61,800
11	44,400	52,800	57,120
12	43,200	51,000	55,200
13	40,200	47,400	51,240

Table 2: Post-based merit salary of a university in Nanjing.

Professional and technical level	Post-based merit salary
2	120,000
3	100,000
4	80,000
5	60,000
6	55,000
7	50,000
8	45,000
9	40,000
10	35,000
11	30,000
12	27,000

Analysing the current situation of the setting of post-based merit salary of teachers in different universities, it is found that there are several problems as follows:

(1) Post-based merit salary cannot be adjusted dynamically with the price level. The average increment speed of post-based merit salary should match with the increment speed of social production efficiency and price increases; that is, salary increment speed needs to be adjusted with the increment speed of per capita GDP and the increment speed of CPI. However, the changes in price levels have not been considered in the current setting of post-based merit salary of many colleges and universities.

(2) Post-based merit salary structure is relatively simple. China's public institutions (university is a kind of public institutions in China) began implementing the employee salary system in 1993, implementing the post allowance system in 1999, and established the merit salary system in 2006 [9]. Although many colleges and universities have completed the setting of post-based merit salary, they still use the same way as the original post allowance system. They fail to follow the principle of "focus on efficiency with due consideration given to equity," "more pay for more work," and "good pay for good work" [10] in the setting of post-based merit salary.

(3) The setting of difference between post-based merit salary levels lacks an objective basis, so it is not persuasive. The current difference between the post-based merit salary levels of university teachers is mostly based on the original job allowance salary grade differential multiplied by the corresponding

coefficient, which lacks an objective basis and scientific calculation method.

(4) The determination method of post-based merit salary is not smart enough. The current determination method of post-based merit salary is mainly through the static table, and this method lacks the ability to change intelligently and dynamically, which makes it difficult for post-based merit salary to keep pace with the times. It also forces universities to manually adjust post-based merit salary rates every few years to keep up with the development of the times, thereby increasing the difficulty of the work.

3. The Determining Principle and Function Design of Post-Based Merit Salary of University Teachers

3.1. The Influencing Factors and Determining Principles of Post-Based Merit Salary of University Teachers. The salary distribution system for university teachers must not only comply with national policies and relevant regulations but also be affected by the social environment and the specific conditions of the university. Among the four parts of university teachers' salary, the three parts, namely, post salary, pay grade salary, and allowance, must be implemented in strict accordance with national regulations. Only the merit salary part is flexibly determined by the university according to its own situation and social influence. Therefore, in order to promote its own development and progress, every university should make scientific decisions in order to optimize its merit salary distribution plan. The influencing factors and determining principles of university teachers' post-based merit salary are as follows:

(1) Focus on motivation effect and guarantee internal fairness.

The main distribution principles of post-based merit salary are as follows: "focus on efficiency with due consideration given to equity," "highlight the role of performance," and "good pay for good work." According to these principles, colleges and universities should appropriately widen the income difference between teachers of different positions and levels in order to enhance the incentive of compensation and at the same time should focus on improving internal fairness. Therefore, according to the different job levels of teachers, the salary increase range should be allocated reasonably to balance efficiency and fairness.

(2) Adjust dynamically according to social development and price changes.

The salary increase mechanism of university teachers must be synchronized with the economic development of the country and the region. Price level increase is an accompanying phenomenon of social and economic development, and the consumer price index (CPI) [11] is constantly changing. Therefore, it is necessary to establish a linkage mechanism in which merit salary increases with the price level increase to ensure the fundamental living demands of university teachers and help them share social development achievement.

(3) Reasonably determine the post-based merit salary of teachers according to the situation of the university.

When specifically determining the post-based merit salary, each university should make reasonable arrangements according to its own situation, and the following four key factors need to be considered.

Firstly, reasonable income differential: colleges and universities should set up a reasonable income differential between teachers of different post levels. As mentioned above, in the process of determining post-based merit salary, both incentive and balance should be carefully studied and considered, reasonable differences of post-based merit salary should be formed, and efficiency and equity [12] should be balanced as much as possible.

Secondly, being able to link up with job performance assessment and evaluation system [13]: when determining the post-based merit salary, we should fully consider the post and performance appraisal results so that the salary can better reflect the work performance of teachers, better play the role of salary leverage, and reflect the distribution principle of "more pay for more work," "reward the good and fine the bad," and "reward the diligent and punish the lazy."

Thirdly, reasonable growth and flexible and dynamic adjustment: in order to fully stimulate the enthusiasm of teachers, the post-based merit salary of university teachers should have a reasonable growth year by year and can be flexibly and dynamically adjusted according to the specific situation of the university.

Fourthly, the limitation of the total amount of school merit salary: the verification of total post-based merit salary should be related to the overall development level of colleges and universities, and a linkage mechanism should be established to link the total post-based merit salary with the development of colleges and universities and the financial level. The amount of post-based merit salary for teachers (teaching and scientific research staff) is allocated within the total post-based merit salary of the school, and the post-based merit salary for each teacher is reasonably determined within this amount.

3.2. The Function Design of the Post-Based Merit Salary of the Teachers in Colleges and Universities. In order to optimize the post-based merit salary of teachers according to the principles of determining the post-based merit salary of teachers in colleges and universities and establish a scientific, dynamic, and intelligent method for determining the post-based merit salary of teachers, the functions of the method are planned and designed as follows:

(1) Under the same circumstances of other variables, the higher the level of professional and technical posts, the higher the post-based merit salary. The highest level and structure proportion of professional and technical posts shall be determined in accordance with the functions, specifications, affiliations, and professional and technical level of public institutions and in accordance with the current relevant regulations on professional and technical post management and the guiding opinions on post setting management of the industry. In colleges and universities, professional and technical posts are divided into 13 levels, including senior titles, vice-senior titles, intermediate titles, junior titles, and staff titles. Among them, technical levels 1–4 are senior titles, levels 5–7 are vice-senior titles, levels 8–10 are intermediate titles, levels 11-12 are junior titles, and level 13 is staff titles. In particular, the professional technical level 1 is generally academician, and the salary system of academician and other high-level talents is generally planned separately. Therefore, this model does not consider the professional technical level 1 and other high-level talents.

(2) Under the same circumstances of other variables, the longer the length of service, the higher the post-based merit salary. The length of service reflects the working time of the employees. The post-based merit salary increases with the length of service, which is conducive to maintaining the stability of the employees and meets the needs of the sustainable development of colleges and universities.

(3) Under the same circumstances of other variables, the better the performance appraisal result is, the higher the post-based merit salary is. Post-based merit salary can reflect the job content and workload. The better the performance appraisal result is, the higher the salary amount is, motivating employees to work hard and improve work efficiency.

(4) Post-based merit salary should change dynamically with the price level. At present, with the rapid development of society and the continuous improvement of the price level, the adjustment of post-based merit salary is obviously delayed. In order to make up for this deficiency, job performance pay should change dynamically and intelligently with the price level so as to maintain the salary competitiveness of university and attract [14] and retain talents.

(5) The total amount of post-based merit salary is controllable. Considering the relevant policies of performance pay, the employment cost of colleges and universities, and the actual situation of colleges and universities, the total amount of post-based merit salary should be designed as a controllable value.

4. Research on the Optimization Model of Post-Based Merit Salary for University Teachers

4.1. Research on the Method of Determining the Post-Based Merit Salary of University Teachers. According to the above analysis, we design the post-based merit salary of university teachers according to the dynamic, controllable, and hierarchical principles.

(1) Determining the grade differential of post-based merit salary hierarchically.

Reasonable salary grade differential design is the key to determine post-based merit salary. To distinguish the difference of post-based merit salary among scientific research personnel with different post levels and service years, we should consider not only fairness but also efficiency and not only incentive effect but also maintaining stability.

According to the work complexity of different post levels and considering the contribution rate of each post to the development of colleges and universities, the difference of post-based merit salary between different post grades should be reasonably drawn. The design adopts three layers: senior titles, vice-senior titles, intermediate and lower titles, and salary grade differentials are set in each level. The specific treatment uses the mathematical model of the salary grade differential in sections and adjusts the gradient flexibly of salary grade differential by adjusting the parameters.

In addition, in order to reflect the factors of "fairness" and "stability" of the sign of employment years, a coefficient related to the employment years should be set when determining the salary grade differential so that the post-based merit salary increases with the service period reasonably.

(2) Dynamic adjustment of post-based merit salary with the price level.

With the continuous change of national and local price levels, the income of staff should be adjusted accordingly. Although the state advocates the coordination between the salary of staff in public institutions and the development of the national economy, many institutions have not yet done so. On the one hand, Japan's practice can be used for reference: the Japanese personnel academy investigates the wage level, national standard living expenses, and living price index of employees in private enterprises every year and writes the civil service wage adjustment plan after analysis. The cabinet formulates the wage adjustment bill accordingly and reports it to the parliament for deliberation and implementation.

In order to ensure the income level of the staff in public institutions and maintain the competitiveness of colleges and universities to attract talents, CPI is introduced as a parameter to determine the post-based merit salary of teachers. For specific operations, we can check the CPI index of local urban residents in the previous year published by the national bureau of statistics. If it is accurate to month, we can check the CPI index of local urban residents in the previous month published by the national bureau of statistics to adjust the results of the salary model.

(3) Determine the total post-based merit salary according to the specific situation of the university and dynamically control it.

The setting of post-based merit salary is important to the performance and development needs of specific universities teaching and scientific research and should be able to make reasonable adjustment according to the post performance. The total amount of post-based merit salary in colleges and universities should be designed as controllable parameters, and colleges and universities can make regular (generally one year) adjustments according to their own specific conditions.

4.2. Dynamic Generation Model of Post-Based Merit Salary for University Teachers

4.2.1. Basic Generation Model of Post-Based Merit Salary. According to the determination method and work demand of post-based merit salary of university teachers, this article designs the generation model of post-based merit salary of university teachers, and its basic model is as follows:

$$S_i = \delta \times F \times (1 + \mathrm{CPI}) \times \rho, \quad \text{subject to} \quad \sum_{i=1}^{N} S_i < T, \tag{1}$$

where S_i is the annual post-based merit salary of teaching staff and its unit is Yuan; δ is the coefficient of salary grade differential and $\delta \geq 1$; F is post-based merit salary baseline defined as the value of S_i last year when $\delta = 1$ and $\rho = 1$. CPI is the consumer price index of local urban residents, ρ is the performance appraisal coefficient of the previous year, N is the total number of teachers, and T is the total amount of post-based merit salary with the unit of Yuan.

4.2.2. Discussion and Parameter Determination of Post-Based Merit Salary Generation Model. We analyse and discuss the basic model (1) and study the method of determining each parameter as follows (Figure 1).

(1) δ coefficient of salary grade differential is a key parameter in the post-based merit salary generation model, which reflects the impact of the two dimensions of professional technical level and length of service on post-based merit salary.

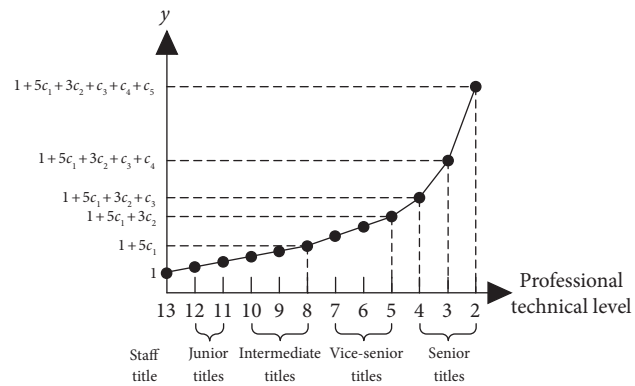


Figure 1: Curve of y, professional technical level parameter.

$$\delta = y + p. \tag{2}$$

Among them, y is the professional technical level parameter and p is the length of service parameter. These two parameters, respectively, reflect the influence of the professional technical level and length of service on δ, the coefficient of salary grade differential.

We first discuss the method of determining y the professional technical level parameter. The coefficient of salary grade differential between professional technical levels can be increased in four ways. Firstly, proportional salary grade differential: that is to say, the wages of each professional technical level are increased step by step with the same salary grade differential percentage. It has two characteristics. One is the amount of wage increased by the same percentage, and the salary grade differential is gradually expanded with the absolute amount, but the gap is not very large, and the incentive effect is obvious. The other is that it is convenient for labor cost budgeting and enterprise salary plan formulation. Secondly, the progressive salary grade differential: progressive salary grade differential means that the wages of each grade increase step by step by a progressive percentage. The wage salary grade differential determined in a progressive way has obvious absolute disparity between levels and large income gap. Compared with the proportional salary grade differential, it has a stronger incentive effect on employees and is more suitable for some jobs that need to highlight individual abilities. Thirdly, the regressive salary grade differential: that is, the wage levels are gradually increased in a regressive proportion. The regressive salary grade differential is suitable for jobs that require high labor intensity, small technical differences, and regular upgrades to employees. Fourthly, the irregular salary grade differential: that is to say, the changes in the percentage of the salary grade differential and the absolute amount of the salary grade differential are determined in accordance with the "segmental type" between the wages of each grade. In accordance with

the above-mentioned method of "determining the grade differential of post-based merit salary hierarchically," y, the teacher's professional technical level parameter, should be determined by this method.

Irregular salary grade differential is more flexible than other salary grade differentials in determining methods and is more in line with the law of wage distribution. At the same time, considering that the promotion standards for senior titles are quite different from those of intermediate, junior, and staff titles, it is not advisable to choose a unified salary grade differential. Therefore, the irregular gradation method is used to set the salary differential coefficient in the post-based merit salary generation model, and the coefficients of salary grade differential are set as three sectional salary grade differentials, which are staff to intermediate titles, vice-senior titles, and senior titles. From staff titles to junior titles to intermediate titles, a linear salary grade differential distribution is used, as shown in the following equation:

$$y = -c_1 \times \text{Post} + d, \tag{3}$$

where Post is the corresponding professional technical level, c_1 is the slope of the straight line, which reflects the growth rate of professional technical level parameter with the promotion of professional technical level, and d is the intercept of the straight line. Set the professional technical level parameter y of the new employee v of professional technical level 13 as 1, so there is $-c_1 \times 13 + d = 1$, i.e., $d = 1 + 13 \times c_1$. Then, equation (3) is obtained as

$$y = 1 + (13 - \text{Post}) \times c_1. \tag{4}$$

When the professional and technical level rises to the vice-senior titles, the grade difference should increase correspondingly due to the improvement of talent level, and the slope of professional technical level parameter is increased to c_2. Similarly, in the senior titles, the higher the professional technical level is, the more difficult it is to promote, and the talent level rises greatly, and the slope of professional technical level parameter is increased to c_3. There are three main reasons to consider when setting the three parameters c_1, c_2, and c_3. Firstly, doctors and masters can get intermediate titles through the orientation of professional and technical level, which is less difficult, so the slopes of intermediate and junior titles are the same and the minimum. Vice-senior is more difficult to promote than intermediate, while senior is more difficult than vice-senior, and the return should be more abundant, so the slope gradually increases. Secondly, from the perspective of performance appraisal, the higher the level of professional and technical posts, the higher the performance appraisal requirements, and the salary level should also be improved accordingly, which is in line with the principle of more pay for more work and distribution according to one's performance. Finally, considering the number of new promotion places each year, the higher the professional and technical level, the fewer the promotion places, which is in line with the pyramid talent structure of colleges and universities, and the salary should be tilted to more senior technical personnel, which has a greater incentive effect on the staff of colleges and universities. So, we finally set $c_3 > c_2 > c_1$. From this, we can get y the professional technical level parameters from staff title 13 to senior title 2 as follows:

$$y = \begin{cases} 1 + (13 - \text{Post})c_1 + (8 - \text{Post})c_2 + (5 - \text{Post})c_3, & \text{Post} \le 5, \\ 1 + (13 - \text{Post})c_1 + (8 - \text{Post})c_2, & 5 \le \text{Post} \le 8, \\ 1 + (13 - \text{Post})c_1, & \text{Post} \ge 8, \end{cases} \tag{5}$$

where $c_3 > c_2 > c_1$. According to the reasons described in Section 3.2, the professional technical level 1 is not considered in the model. A schematic diagram of the professional technical level parameter curve is shown in Figure 1.

In addition to the professional technical level, length of service is also an important factor to be considered in the post-based merit salary generation model. The contribution and return of senior teachers are reflected by the job performance bonus brought by the length of service. The length of the service parameter is determined by the following formula:

$$p = \text{Year} \times \eta \times c_1, \tag{6}$$

where Year is the length of service of the teacher (less than one year is calculated as one year); new employee Year = 1. η is a coefficient that can be set on demand and $\eta \times c_1$ reflects the growth rate of grade difference coefficient with length of service. Generally, this speed should be far less than the growth rate of professional technical level parameter with professional technical level promotion, so η is usually far less than 1.

(2) F is post-based merit salary baseline. The post-based merit salary baseline F is constant, defined as the value of S_i with $\delta = 1$ and $\rho = 1$ in the previous year. Therefore, F changes with CPI the year-on-year, as shown in the following equation:

$$F_{\text{the } N\text{th year}} = F_{\text{the } (N-1)\text{th year}} \times (1 + \text{CPI}). \quad (7)$$

When using this model, the value of F in the first year can be determined in the following ways. Firstly, it can be reasonably estimated and determined according to the salary data of the school in the previous year. Secondly, for special cases, such as the new school, it can be reasonably estimated and determined by referring to the per capita disposable monthly income of local urban residents or the per capita consumption expenditure of local urban residents.

(3) Consumer price index (CPI) is a relative number reflecting the change trend and range of the price level of goods and services purchased and consumed by residents. It is closely related to people's lives and plays a very important role in the price system of the whole national economy. This index includes not only the prices of all kinds of consumer goods needed by urban and rural residents in daily life but also the prices of various services closely related to people's life, such as water, electricity, transportation, education, and medical care. It can comprehensively reflect various factors of market price changes and their impact on the actual life of residents.

Adding the CPI parameter item into the post-based merit salary generation model is the concrete embodiment of the method of "dynamically adjusting the post-based merit salary with the price level." The purpose is to make the post-based merit salary change with the price level dynamically and keep pace with the development of the national economy so as to ensure the basic living needs of university teachers.

In some exceptional circumstances, such as the COVID-19 epidemic in 2020, there may be sharp changes in CPI. Therefore, we should give CPI a range of values in case of this kind of sharp change happens. Considering recent years' situation, we think when CPI > 3% or CPI ≤ 3%, the CPI in the model should be substituted with 3% or −3%, not the true value of CPI.

(4) ρ is the performance appraisal coefficient of the previous year. Because the post-based merit salary should reflect the work responsibilities, basic workload, and performance appraisal results, it is closely combined with the performance and demand of teaching and research work in specific universities, and the performance appraisal coefficient ρ of the previous year can be set to make reasonable adjustment to this situation. ρ is determined according to the comprehensive evaluation of the workload and performance appraisal results in the previous year and positively related to the comprehensive evaluation results. If the evaluation results are in a regular way, ρ is taken as 1. Correspondingly, if the evaluation results are higher or lower than the average level, they are taken as greater than 1 or less than 1, respectively. The value can be determined according to the specific performance appraisal results, one value for one person, in principle $0.7 \le \rho \le 1.3$.

4.2.3. Dynamic and Optimal Generation Model of Post-Based Merit Salary. The final generation model of post-based merit salary is established by introducing all the above parameters into the basic model (1). The final model is a dynamic and optimized generation model of the post-based merit salary of university teachers, as shown in the following:

$$S_i = \begin{cases} \left(1 + (13 - \text{Post})c_1 + (8 - \text{Post})c_2 + (5 - \text{Post})c_3 + \text{Year} \times \eta \times c_1\right) \times F \times (1 + \text{CPI}) \times \rho, & \text{Post} \le 5, \\ \left(1 + (13 - \text{Post})c_1 + (8 - \text{Post})c_2 + \text{Year} \times \eta \times c_1\right) \times F \times (1 + \text{CPI}) \times \rho, & 5 \le \text{Post} \le 8, \\ \left(1 + (13 - \text{Post})c_1 + \text{Year} \times \eta \times c_1\right) \times F \times (1 + \text{CPI}) \times \rho, & \text{Post} \ge 8, \end{cases} \quad (8)$$

$$\text{subject to} \quad \sum_{i=1}^{N} S_i < T, \quad c_3 > c_2 > c_1.$$

The meaning of each parameter in (8) is the same as those mentioned before, and it will not be repeated. Some parameters are explained as follows.

For the performance appraisal coefficient ρ in the previous year, the benchmark value is 1. If the performance of the current year is poor, it can be reduced as appropriate, such as 0.95. If the performance is good, it can be increased as appropriate, such as 1.05. η is a parameter used to adjust the influence percentage of the length of service on post-based merit salary. $1/\eta$ can be understood as the number of years of service required to increase the allowance of a professional technical level. $c_1 > 0$ refers to the difference of professional technical level parameters between adjacent professional technical levels in level 13 to level 8. The larger c_1 is, the greater the salary grade differential is. $c_2 > c_1$ refers to the difference of professional technical level parameters between adjacent professional technical levels in level 8 to level 5. $c_3 > c_2$ refers to the difference of professional technical level parameters between adjacent position levels in level 5 to level 2.

5. Dynamic Optimization Determination Method for Post-Based Merit Salary of University Teachers

5.1. Determination Method and Optimization Adjustment of Post-Based Merit Salary for University Teachers.

According to this study, the post-based merit salary of university teachers is calculated and determined according to equation (8) and adjusted dynamically. We should pay attention to the following problems in the specific work.

(1) Solve the coordination problem between the total amount of post-based merit salary and related parameters in colleges and universities:

In the dynamic optimization model shown in (8), the constraint is the total amount of post-based merit salary of a specific university. However, according to the above research and discussion, each parameter in the equation has its reasonable value range. Therefore, in the specific determination of a year's post-based merit salary, although it can be calculated according to the model, it often needs to be adjusted according to the actual situation of the university, and finally, determine the total post-based merit salary of the university and the value of the relevant parameters.

The specific treatment method of this problem is as follows:

Due to the changes in the social environment and specific situations of colleges and universities each year, the total performance salary of colleges and universities will change every year. For a specific year, the total performance salary is a fixed value. The decision-making staff of colleges and universities cut the total performance salary according to the components of performance salary and determine the total post-based merit salary according to the appropriate proportion; that is, the total post-based merit salary T discussed in this article is a component of the total performance salary.

In practical work, if the other parts of the total performance salary can be compressed, it can ensure that the value of T is not less than the total demand of the post-based merit salary calculated by the current model, and all the relevant parameter values do not need to be adjusted adaptively. In this case, T = the total amount of post-based merit salary calculated by the model.

If, for some reason, the value of T cannot meet the total demand of post-based merit salary calculated by the current model, then through calculation, the performance baseline F of this year will be reduced adaptively. Because S and F are related in each position grade model, without changing other parameters, each employee's post-based merit salary S will decrease accordingly. Thus, the post-based merit salary of all teachers is reduced by equal proportion based on total amount control.

(2) Reasonably determine the salary grade differential of post-based merit salary according to the current situation of colleges and universities.

The salary differences between different post levels are adjusted by parameters c_1, c_2, c_3, in which $c_3 > c_2 > c_1$. Although it can reasonably widen the difference in post-based merit salary between different post levels, there are still two problems: the values of c_1, c_2, c_3 and the gap between three parameters. These two problems still need to be solved through experience and analysis in practical work.

The recommended approach to this problem is that the values of c_1, c_2, c_3 are determined comprehensively according to the proportion of the teachers of each post level among the full-time teachers in the university in the current year and the performance appraisal requirements for teachers of each post level in the current year. Meanwhile, there is secondary management in colleges and universities. The management institutions of secondary colleges can adjust c_1, c_2, c_3 to meet the requirements of the staff structure and performance appraisal of the college according to the personnel structure of their respective colleges.

5.2. Case Analysis on the Determination of Post-Based Merit Salary of University Teachers. Taking the actual situation of a university in Shanghai as an example, this article analyses the feasibility and effectiveness of the proposed optimization method of post-based merit salary.

5.2.1. Parameter Setting. Through investigation, the CPI of cities and towns in 2019 is 2.9%. Combined with the actual situation of a university in Shanghai, set $F_{2019} = 40,000$, $\rho = [0.8,\ 1,\ 1.2]$, $\eta = 0.25$, $c_1 = 0.13$, $c_2 = 0.22$, $c_3 = 0.31$. These parameters are brought into equation (8), and the model is transformed into the following:

$$S_i = \begin{cases} (6 - 0.66 \times \text{Post} + 0.0325 \times \text{Year}) \times 41160 \times \rho, & 2 \le \text{Post} \le 5, \\ (4.45 - 0.35 \times \text{Post} + 0.0325 \times \text{Year}) \times 41160 \times \rho, & 5 \le \text{Post} \le 8, \\ (2.69 - 0.13 \times \text{Post} + 0.0325 \times \text{Year}) \times 41160 \times \rho, & \text{Post} \ge 8, \end{cases} \tag{9}$$

$$\text{subject to} \quad \sum_{i=1}^{N} S_i < T.$$

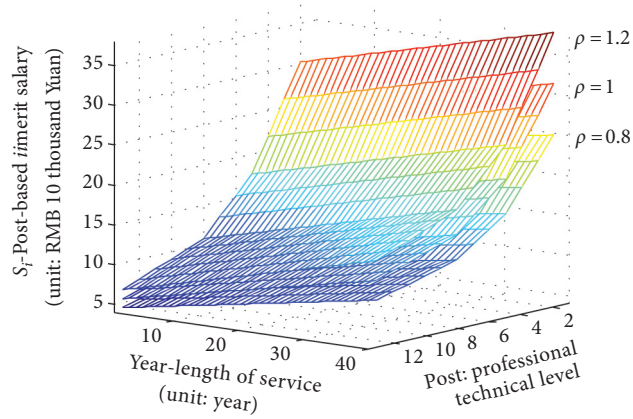


Figure 2: Example simulation results of post-based merit salary generation model.

Table 3: Test sample group.

No.	Year	Professional technical level	Length of service	Performance appraisal result	Post-based merit salary by our model	Current post-based merit salary
A	*B*	*C*	*D*	*E*	*F*	*G*
1	2020	10 (intermediate)	8 year	1	67,914	73,800
2	2020	7 (vice-senior)	8 year	1	93,022	95,040
3	2020	7 (vice-senior)	15 year	1	102,386	95,040
4	2020	7 (vice-senior)	15 year	0.9	92,147	85,536
5	2020	7 (vice-senior)	15 year	1.1	112,624	95,040
6	2020	5 (vice-senior)	15 year	1	131,198	108,000
7	2020	4 (senior)	15 year	1	158,363	130,920
8	2019	4 (senior)	15 year	1	153,900	130,920

5.2.2. Calculation Results. Simulate the equation (9) model with MATLAB, and draw the mesh image of post-based merit salary, as shown in Figure 2.

It can be seen from Figure 2 that the post- and performance-based salaries generated by this model can set reasonable differences with the increase of post level, the differences of level 13 to level 8 are balanced, and the differences of level 8 to level 5 and level 5 to level 2 are appropriately improved. At the same post level, the coefficient of gradation also rises to a certain extent as the length of service increases. Meanwhile, the results of job appraisal also have a certain range of influence on post-based merit salary.

In order to more specifically observe the running results of the model, we take 8 teachers' samples for testing, and the results are shown in Table 3. Among them, column *F* is the post-based merit salary value of each teacher obtained by this model, and column *G* is the current post-based merit salary of corresponding teachers in a university in Shanghai.

By analysing the results of column *F*, we can see the rules of post-based merit salary of this model: from teachers 1 and 2 and teachers 3, 6, and 7, the higher the post level, the higher the post-based merit salary of teachers with the same length of service and the same performance appraisal results. From teachers 2 and 3, it can be seen that the longer the length of service, the higher the post-based merit salary of teachers with the same post level and performance appraisal results. From teachers 3, 4, and 5, we can see that the better the performance appraisal results, the higher the post-based merit salary. From teachers 7 and 8, it can be seen that the post-based merit salary of teachers with the same post level, length of service, and performance appraisal results dynamically change with the CPI year by year.

Analysing the results of column *G*, we can see the current law of post-based merit salary. Teachers 3, 6, and 7 show that the higher the post level, the higher the post-based merit salary. Teachers 2 and 3 show that the length of service has no effect on post-based merit salary. From teachers 3, 4, and 5, it can be seen that, compared with teachers with average performance appraisal results, teachers with poor performance appraisal results have post-based merit salary, but teachers with good performance appraisal results have no increase in post-based merit salary. From teachers 7 and 8, it can be seen that post-based merit salary does not change dynamically with the CPI year by year.

In summary, the test results prove that compared with the current post-based merit salary system of a university in Shanghai, the dynamic and optimized generation model of post-based merit salary developed in this article is reasonable and effective, and it can reflect the teachers' own post level, length of service, and performance appraisal results. Furthermore, the post-based merit salary generated by this model can also change dynamically with the price level.

5.3. Model Improvement and Problem Discussion. Although the method and optimization model of university teachers' post-based merit salary proposed in this article satisfactorily solve the problems in the current salary system

such as single structure and cannot be adjusted dynamically according to price changes, it still needs to be gradually improved in the application. Related issues and their solutions are discussed as follows.

The difficulty of the determination method and dynamic generation model of university teachers' post-based merit salary is the unity of its realistic rationality, social influence, and intelligent optimization. On the premise of studying and putting forward the dynamic generation model of post-based merit salary, the difficulty is transformed into the problem of reasonable determination of each parameter in the model. As mentioned above, although this article puts forward the principles and methods of parameter determination, it still needs discussion, research, and specific determination by relevant leaders. In the practical work, due to the need for different departments in colleges and universities to complete the work, there is a gap in mutual coordination, resulting in the actual effect being often difficult to achieve the satisfaction of teachers and the real sense of optimization. This problem still needs to be solved by the effective cooperation of different departments in colleges and universities.

6. Conclusion

Based on the consumer price index and the characteristics of colleges and universities, this articl analyses and discusses the influencing factors and determination principles of post-based merit salary for teaching and scientific research personnel in colleges and universities, and studies its dynamic and optimal generation model and determination method.

(1) The grade differential of post-based merit salary for university teaching and scientific research personnel should be determined at different levels and should be dynamically adjusted according to the price level. At the same time, the total post-based merit salary should be determined according to the specific conditions of the university, and it should be dynamically controllable.

(2) The dynamic and optimized generation model of post-based merit salary for teaching and scientific research personnel established in this article can truly reflect the various elements of post-based merit salary through simulation experiments and case analysis. The amount of post-based merit salary is reasonably adjusted according to the post level of the teacher, the length of service, and the result of performance appraisal and can change dynamically according to the level of social consumption. This model not only conforms to the principle of giving priority to efficiency and giving consideration to fairness but also increases the competitiveness of universities in attracting talents. Therefore, the post-based merit salary optimization model and method proposed in this article is an innovation in this field, which can be flexibly applied according to the actual situation of different universities.

(3) The model and method proposed in this article leave room for further research in its parameter determination methods, which should be solved in practice and further research.

Data Availability

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

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work was supported in part by the Project of the Humanities and Social Sciences on Young Fund of the Ministry of Education in China (Research on Swarm Intelligence Collaborative Robust Optimization Scheduling for High-Dimensional Dynamic Decision-Making System, no. 20YJCZH052) and 2020 Donghua University Higher Education Planning Special Topic Project (Research on Performance Wage Distribution Method Based on the Sense of Pay Fairness, no. 206-99-0243030).

References

[1] W. Jones, R. Haslam, and C. Haslam, "What is a 'good' job? modelling job quality for blue collar workers," *Ergonomics*, vol. 60, no. 1, pp. 1–22, 2016.

[2] S. Rahayuningsih, "The influence of compensation and work environment towards teacher's job satisfaction," *SHAHIH Journal of Islamicate Multidisciplinary*, vol. 2, no. 2, pp. 157–169, 2017.

[3] X. Hu, "Research on performance pay policy for college teachers in China: logical evolution, theoretical framework and future development," *Journal of National Academy of Education Administration*, no. 1, pp. 89–95, 2020.

[4] X. Liu, "Difficulties and countermeasures of performance pay reform in public institutions," *Party School of the Central Committee of the Chinese Communist Party*, vol. 23, no. 1, pp. 30–34, 2019.

[5] M. Mihai, P. Constanța, T. Ana-Maria, and T. Valerica, "Modelling discrete choice variables in assessment of teaching staff work satisfaction," *PloS One*, vol. 10, no. 4, Article ID e0115735, 2015.

[6] Y.-M. He, Yu-L. Pei, B. Ran, J. Kang, and Y.-T. Song, "Analysis on the higher education sustainability in China based on the comparison between universities in China and America," *Sustainability*, vol. 12, 2020.

[7] G. Rodriguez, C. Zhou, and M. Carrio, "Creativity in biomedical education: senior teaching and research staff's conceptualization and implications for pedagogy development," *International Journal of Engineering Education*, vol. 33, no. 1, pp. 30–43, 2017.

[8] J. Arsenault, J. Boustani, J. Talbot, R. Gonzales, and K. Manaugh, "The environmental footprint of academic and student mobility in a large research-oriented university," *Environmental Research Letters*, vol. 14, no. 9, Article ID 095001, 2019.

[9] W. Zhen, "Exploration of the reform of the performance-related pay system in institutions in the new era," *Value Engineering*, vol. 37, no. 22, pp. 111-112, 2018.
[10] G. Yan, "The trajectory, problems and future of academic appointment system reform," *China Higher Education Research*, no. 10, pp. 1–9+19, 2019.
[11] G. A. Sarantitis, T. Papadimitriou, and P. Gogas, "A network analysis of the United Kingdom's consumer price index," *Computational Economics*, vol. 51, no. 2, pp. 173–193, 2018.
[12] A. Cohn, E. Fehr, and L. Goette, "Fair wages and effort provision: combining evidence from a choice experiment and a field experiment," *Management Science*, vol. 61, no. 8, pp. 1777–1794, 2015.
[13] C. R. Sánchez, D. Díaz-Cabrera, E. H. Fernaud, Does effectiveness in performance appraisal improve with rater training," *PLoS One*, vol. 14, no. 9, pp. 1–20, 2019.
[14] C. Björn, M. Lindberg, and D. Rissén, "Significant factors for work attractiveness and how these differ from the current work situation among operating department nurses," *Journal of Clinical Nursing*, vol. 25, no. 1-2, pp. 109–116, 2016.

Hindawi
Scientific Programming
Volume 2021, Article ID 9608147, 8 pages
https://doi.org/10.1155/2021/9608147

Research Article

Health Analysis of Footballer Using Big Data and Deep Learning

Tao Yang,[1] Guoliang Yuan ,[2] and Jing Yan[3]

[1]*Institute of Physical Education, North Minzu University, Yinchuan, Ningxia 750021, China*
[2]*College of Physical Education, HengShui University, Hengshui, Hebei 053000, China*
[3]*Chinese Academy of Customs Administration, Qinhuangdao, Hebei 066004, China*

Correspondence should be addressed to Guoliang Yuan; 601254@hsnc.edu.cn

Received 7 April 2021; Revised 30 April 2021; Accepted 24 May 2021; Published 16 June 2021

Academic Editor: Shah Nazir

With the development of information technology, health management and big data have risen and developed in recent years. Big data need proper analysis and shape in order to extract meaningful information from it. This paper analyzes the application status and prospects of big data in the field of health management. The results show that the most widely used big data in health management are intelligent wearable devices. Big data applications in football players' mental health monitoring systems and chronic disease health management systems also have a good prospect. The intelligent wearable device is applied to several aspects of sports work: teaching and sports training, real-time monitoring of football players' physical exercise process, collecting football players' heart rate, calorie consumption, exercise steps, and track, blood pressure, blood oxygen, and other physical exercise data; through monitoring the heart rate, we can get the intensity and duration of football players' physical exercise in school; through the calculation, we can also get the football players' time energy consumption and understand the overall situation of football players' physical exercise in school; through step counting and track monitoring, we can master the number of steps and track of football players; by monitoring the changes of blood oxygen and blood pressure of football players, we need to build a third-party residents' health information storage and analysis system and further realize the marketization of residents' health big data. The experimental results of the proposed study show the effectiveness of the proposed work.

1. Introduction

As people pay more attention to health problems, people pay more and more attention to health-related products, especially football players. With the social development and economic progress, more and more football players wear intelligent wearable devices. More and more football players use intelligent wearable devices to monitor their health problems. More and more people begin to pay attention to sports health management. With the development of the electronic communication industry, the smartphone industry has developed to an unprecedented peak. During this period, big smartphone manufacturers such as apple and Samsung encountered bottlenecks in mobile phone innovation. People put the development of smart devices into smart wearable devices, and various smart wearable devices have sprung up like mushrooms. In 2013, about 6.75 million smart wearable devices were sold in China, and it is expected to exceed 73.5 million in 2016 and develop into a huge market with a scale of 16.94 billion yuan [1] as intelligent wearable devices need to contact with the body and they have the advantage of detecting the body quality. Therefore, many manufacturers have launched intelligent wearable devices with fitness functions, and the development opportunities of intelligent wearable devices in the sports field will be unprecedented. As the intelligent wearable device is a new product of scientific and technological development, there is no unified conclusion on the definition of the intelligent wearable device [2].

In this paper, through the analysis of football players' personalized teaching, Sports Prescription Theory and attribution, cognitive dissonance, and behavior reinforcement theory, this paper studies and analyzes the football players in various countries. By using the research on the status quo of football players' physique and other aspects, this paper analyzes the optimization [3] effect of intelligent wearable

equipment on the sports health management of football players. Also, through the support of society, government, enterprises, and the specific measurement of football players' physical condition, we can understand the various functions of wearable devices and analyze the different functions of different intelligent wearable devices. At present, the field of intelligent wearable devices is developing rapidly, and there are various kinds and quantities of intelligent wearable devices. Now for our football players, it is the priority to understand the real benefits that intelligent wearable devices can bring us. Secondly, we should make a targeted selection of devices and, finally, formulate a reasonable sports health plan to improve football games mobilize their physical quality to achieve the goal of sports health management. The following are the key contributions of the paper:

- To analyze the application status and prospects of big data in the field of health management
- To study big data applications in football players' mental health monitoring systems and chronic disease health management systems also have a good prospect
- To apply intelligent wearable device for several aspects of sports work
- To show the effectiveness of the proposed study through experimental results

This paper is organized as follows: Section 2 presents the research overview of the intelligent wearable devices. Section 3 is the representation of the design and implementation of a wearable human motion capture system. The details of design of the motion capture hardware platform are given in Section 4. Section 5 briefly represents the report on the total amount of sexual activity and calorie consumption of football players. Section 6 represents the details of classification and function of intelligent wearable devices. The value of intelligent wearable devices in sports activities is given in Section 7. This paper is concluded in Section 8.

2. A Research Overview of Intelligent Wearable Devices

The Baidu Encyclopedia defines intelligent wearable devices as "wearable intelligent devices" as the general term of wearable devices, such as glasses, gloves, watches, clothing, and shoes, by applying wearable technology to intelligently design and develop wearable devices. Li Yang believes that "wearable devices refer to intelligent devices with the independent operating system, system application, upgrading and scalability, worn by the human body, and realizing continuous interaction." This paper agrees with Li Yang's statement that wearable devices refer to intelligent devices that adopt the independent operating system, have system applications, upgrade and expansion, are worn by the human body, and realize continuous interaction [4]. The research on intelligent wearable devices is less, more just as science and technology news reports, domestic and foreign research on intelligent wearable devices in the field of sports is less, so this study has a certain degree of innovation for the research of intelligent devices in the field of sports to provide a reference and reference.

2.1. Development and Evolution of Intelligent Wearable Devices. Early wearable devices can be traced back to the 13th century, when the first pair of glasses with crystal lenses appeared. In China, in the 17th century, the abacus ring, the rudiment of intelligent wearable devices, appeared. In 1762, John Harrison invented the pocket watch. In 1907, the first watch was born. People began the watch era. In 1977, HP developed the HP 01 algebra calculator watch, which sold for $600. Computer shoes were designed in 1978 to predict roulette results. In 1993, BBN completed the Pathfinder system. The system is a wearable computer equipped with GPS and a radiation detection system. In 1994, with the advent of wearable cameras, Steve Mann developed a "wearable wireless camera." This device can transmit point-to-point through amateur TV frequency, and images can be transmitted from head-on analog cameras to SGI base stations. In 1998, Seiko launched a smartwatch rupture compatible with Windows 95. In 2010, Nike launched the Nike + SportBand device, one of the founders of smart wearable devices, which can "communicate" with sensors hidden in sneakers and let users know more about their running. In April 2012, Google released Google glasses, bringing intelligent wearable devices into public view. In 2013, Samsung launched the Samsung gear smartwatch. In 2015, Apple launched the Apple Watch smartwatch, and then, 2015 was called the first year of smart wearable devices.

3. Design and Implementation of a Wearable Human Motion Capture System

Motion capture technology is to obtain 3D motion information of the human body in real-time by computer vision [5–7] technology and then analyze and use the information. Fleischer invented dynamic mask technology in 1915, which is regarded as the ancestor of motion capture technology. In the 1970s, the Computer Graphics Laboratory of New York designed an optical motion capture device, which can display actors' actions on the screen in real time, which has become the beginning of modern motion capture technology. Modern motion capture systems are based on different capture principles, including acoustic, electromagnetic, optical, mechanical, video capture, and inertial. The accuracy of acoustic and electromagnetic capture systems is relatively poor, and it is affected by environmental noise and magnetic fields. An optical capture system has high precision, but it is complex to install, costly, and has high environmental requirements. The mechanical motion system is difficult to wear and will affect the free activity of the human body. Although the video capture system does not need to wear sensor nodes on the human body, its algorithm is complex, and it is difficult to realize. The motion capture system based on an inertial sensor has the characteristics of portable wear, simple operation, and low cost. It can continuously collect real-time motion information of all parts of the human body without the limitation of time and place. It has been widely

used in fall monitoring, sports training, military training, somatosensory games, virtual reality, and augmented reality. For example, in the aspect of fall monitoring, through wearing inertial sensors on the wrist or other parts, real-time acquisition of the acceleration, angular velocity, and other information of the wearing part, when the human body suddenly falls, the inertial data of the wearing part will change suddenly, to judge that the human body falls and timely ask the family members and medical staff for help.

To capture human motion in real time, a wearable human motion capture system is designed. Each IMU is composed of a micro-MEMS 3-axis gyroscope, MEMS 3-axis accelerometer, MEMS 3-axis magnetometer, and MCU. An MCU obtains the data of each sensor and calculates the attitude angle of corresponding parts by using the extended Kalman filter based on quaternion. The data are uploaded to the computer in real-time through the CAN bus and Bluetooth module. The computer drives the virtual human body movement through VC++ and OpenGL program to realize real-time human motion reproduction.

3.1. The Overall Design of the System Software Platform. The human motion capture system is composed of a lower computer and an upper computer. The lower computer is distributed in the whole body of sensor nodes, real-time sensor node data, through the wireless module, and sent to the PC. The upper computer vision [8–10] platform runs on the PC. The PC receives the data of sensor nodes in real-time through the wireless antenna. The software platform analyzes and processes these data and reconstructs the calculated human posture into computer graphics in real time. To make the software platform achieve certain functions and usability, it is combined with the requirements of the motion capture system itself.

3.1.1. Serial Data Receiving. It is mainly responsible for real-time monitoring of the wireless data received in the serial port of the computer [11], receiving the data distributed in each key sensor node of the human body in real time, distinguishing the nodes of each sensor data, and sending these data to the data preprocessing module.

3.1.2. Data Preprocessing. It mainly completes the preprocessing process of the original signal discussed in Section 3. The signal with error and noise is compensated and filtered. The data are divided into two parts: one part is output to the attitude calculation module. The other part saves the data in real time to facilitate other related simulation experiments.

3.1.3. Loading and Driving of the Human Model. This is in charge of importing the human model made in Section 4 into the 3D scene of the software and driving the human body model in real time by using the data processing module to solve the human posture.

3.1.4. Software Platform Appearance Beautification. This part of the function is mainly responsible for the visual control, title bar, title, and menu bar interface beautification work. In the VC++ development environment, there is no corresponding beautification interface for visual control, menu bar, and toolbar, which needs to be customized and processed.

3.2. Detailed Design of the Software Platform. The system uses wireless to the serial port to complete the communication between the upper computer and the lower computer. In VC++, the serial port needs to be monitored in real time to complete the data receiving. In common VC++, the methods of receiving serial data include calling system API directly, third-party serial port class, and COM-based serial port control. Serial port control has good performance, shielding some complex bottom operations, stable performance, and shortening the software development cycle. This software platform uses the ActiveX control MSComm developed by Microsoft. McComb control supports both the active query mode and event-driven model. In the process of serial port initialization, firstly, the serial port control is initialized, including setting serial port number, baud rate, and input and output buffer space size. Finally, the buffer data are read in advance once, and the data remaining in the buffer are cleared. To prevent data deposition and loss, the data receiving part is encapsulated into a separate module. At the same time, the software improves the concurrency and opens up new threads in the module. In VC++, each thread is inherited from a twin thread class. There are two types of threads, one is a worker thread without a message queue, and the other is a user interface thread with a message queue and the maintenance message loop mechanism. User interface threads are mainly used to interact with users, respond to peripheral input in real-time and process events triggered by the operating system. The worker thread does not need to interact with the user and is usually used to deal with tasks with a large amount of computation, such as printing service and transmission. The data receiving module does not need to interact with the user, it only needs to ensure the safe and stable data receiving, so it is more appropriate to use worker thread for receiving thread.

The format of the data frame is sent by the sensor node. When reading the serial buffer data, the original data of the sensor node in the frame are analyzed. On the one hand, the data are transferred to the attitude calculation module; on the other hand, the original data are saved as a file in a certain format, which is convenient for further analysis and research of sensor node data offline. Generally, the offline data collected are mainly used in algorithm research, which must be stored according to certain rules, and the file format is the general format. At the same time, the software platform saves the original data collected by the gyroscope, velocimeter, and magnetometer to the subfolder under the software path. To read the data conveniently, the data are saved as an XLS file.

VC++ is selected as the software development environment, and the MFC framework and DirectX graphical

interface library are used to complete the loading and driving of the human body model. A multifunctional application platform integrating data receiving and storage, real-time attitude calculation, and real-time human motion reconstruction is completed. The software of human motion and posture capture system is designed, and the functions of real-time motion capture and three-dimensional structure are completed.

4. Design of the Motion Capture Hardware Platform

According to the principle of inertial motion capture equipment and current technology, generally, for single-person motion capture, the model shown in Figure 1 can be used to describe. For the model, users wear 11 inertial motion sensors, which are distributed in the left upper arm, left forearm, right big arm, right forearm, left thigh, left leg, right thigh, right leg, head, chest, and waist. Each motion sensor outputs acceleration, magnetic field strength, and angular velocity inertial parameters. The sensors are connected by a CAN bus, which can supply power to the sensors with high data rate and reliability. The concentrator is the bridge between the sensor and the computer. The sensor node first sends the data to the concentrator and then forwards it to the receiver. The receiver transmits the data to the computer through a USB port. The concentrator and sensor are connected by a wire, and wireless communication is adopted between concentrators and computers.

4.1. The Part of the Motion Sensor. The functions of the motion sensor include data acquisition, attitude angle settlement, and wired communication for three different types of inertial devices. The block diagram of the hardware system is shown in Figure 2. As can be seen from Figure 2, the motion sensor mainly includes five parts: an MCU, acceleration sensor, gyroscope sensor, magnetometer, and CAN interface. The MCU consists of an asynchronous serial communication bus (I2C), asynchronous receiver transmitter (UART), keys, and other modules. It controls a series of operations of nodes. The nodes are powered by wire, and the MCU controls the data acquisition, attitude angle calculation, and CAN transceiver of the three sensors.

The main chips used in motion sensors include STM32F103 (MCU) of St, mpu6050 (3-axis acceleration juice and 3-axis gyroscope) of InvenSense, hmc5883 l (3-axis magnetometer) of honeywe11, and sn65hv D230 (can transceiver) of Ti. The parameters of each chip are shown in Table 1.

5. Report on the Total Amount of Sexual Activity and Calorie Consumption of Football Players

The specific calculation of football players' sports energy consumption is also an important way to understand the number of football players' activity calories are an important reference index to calculate human energy consumption. There are many ways to calculate the energy consumption of football players in stage activities. At present, there are mainly the direct heat measurement methods, indirect side heat method, double standard water detection method, heart rate detection method, pedometer detection method, and triaxial acceleration sensor detection method [12]. Each detection method has its advantages and disadvantages. At present, the commonly used energy monitoring methods in intelligent wearable devices are step counting monitoring, heart rate monitoring, and triaxial acceleration sensor monitoring. Step counting monitoring mainly records the number of human movement steps and then estimates the energy consumption of the human body. Because in some unconventional running and walking movements, the energy consumption of the human body is not only manifested as the distance between the number of steps and the moving position, there is a big error in calculating the energy consumption of football players by using the step counting principle [13];heart rate detection mainly measures the energy consumption of human body according to the change of human heart rate. Heart rate can effectively judge the

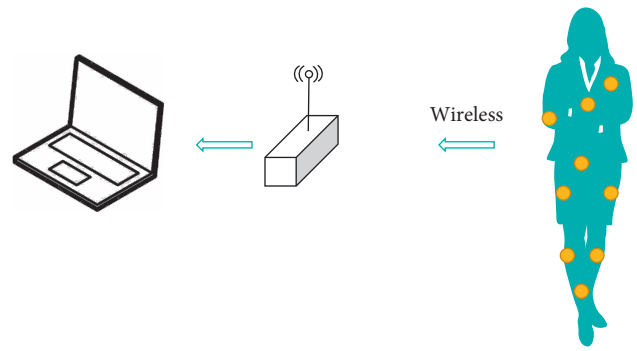


Figure 1: Motion system model.

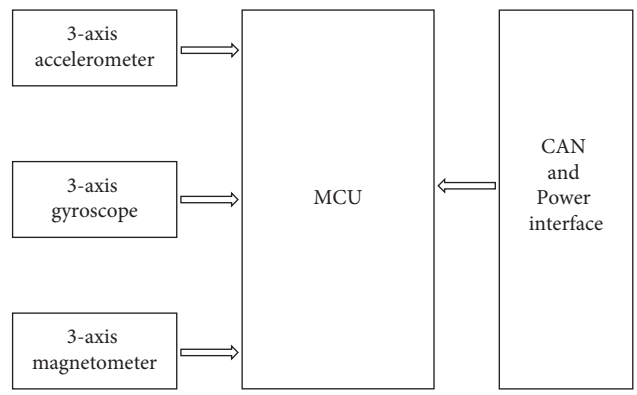


Figure 2: Block diagram of the motion sensor system.

Table 1: Chip parameters.

Type	Name	Parameter
Accelerometer	MPU6050	±250/500/1000/2000°sec
Gyroscope	MPU6050	±2/4/8/16g
Magnetometer	HMC5883 L	±8gauess
MCU	STM32F103	90DMIPS
CAN	SN65HVD230	1 Mb/s

intensity of human exercise and oxygen consumption. According to the corresponding intensity to measure the energy consumption of football players, the equipment requirements are simple, and the accuracy rate of heart rate measurement is also high. However, heart rate is easily affected by the external environment and internal differences of testers, which leads to detection. There is an error. Meijer and other related researchers have also found that the measurement of human energy consumption through heart rate, step counting, three-dimensional acceleration, and other related principles is the ideal way to measure the energy consumption of the human body [14]. The calculation and construction of energy consumption based on the characteristic value GM is one of them. GM value is the geometric average value of three-axis acceleration data as the characteristic quantity to calculate the energy consumption of human body movement:

$$GM = \sqrt{\frac{\left(A_x^2 + A_y^2 + A_z^2\right)}{3}}. \quad (1)$$

A_x, A_y, and A_z are the output values of the acceleration sensor in the X-, Y-, and Z-axis, respectively, and the unit is g. After studying the correlation between the geometric mean of the three-axis acceleration data and the energy consumption of the human body, Yang Huiliang proposed a calculation equation of human energy consumption based on the eigenvalue GM.

$$EF = \frac{\left((1/2)u^{\bullet} mg^{\bullet} \int_{a1}^{a2} \mathrm{d}a \int_{t1}^{t2} t\mathrm{d}t\right)}{4.18}. \quad (2)$$

6. Classification and Function of Intelligent Wearable Devices

There are various types of intelligent wearable devices. "From the perspective of the products used in the wearable intelligent devices, 30% of the products are worn on the wrist and 26% and 22% of the products are worn on the upper body and head, respectively. 15% of the products do not have clear requirements for the wearing parts. However, the feet seem to be the least favored part of the intelligent wearable devices, and only 7% of the products choose the feet." [13].

A fan distribution map of intelligent wearable devices is classified by function. From this, we can see that the current wearable devices are mainly used in the field of medical health and sports fitness. Using wearable devices that can monitor ECG signals to monitor the heart condition of users in walking, running, and other activities can prevent heart disease. A wearable sensor is used to detect the change of heart rate and the number of walking steps during exercise and calculate the heat and calories consumed by the user, evaluate the health level of the user, and formulate a reasonable fitness plan for the user.

According to the wearing position and the temperature monitoring circuit function module, the wearable device is subdivided according to the position, and the wearable device is classified into the following categories according to the functional classification. Wearable devices can be electronic products worn on the head, neck, arm, wrist, or finger or embedded in clothes or shoes. As can be seen from the figure, wearable devices are mainly worn at the wrist. It is because people are used to wearing bracelets or watches, and the wearable devices on the wrist are portable, consume low power, and small in size.

7. The Value of Intelligent Wearable Devices in Sports Activities

According to Fjord's research and analysis of 27 well-known wearable intelligent devices, "about 70% of the main functions of these products are used in sports and health, 23% is used in social communication, and the remaining 7% is used for monitoring and improving sleep." In all the listed intelligent wearable devices, almost all intelligent wearable devices have the same function, that is, the collection and collation of human motion data, which shows that intelligent wearable [15] devices have great development value in the field of sports.

At present, there are various types of intelligent products in the market, including medical and healthcare products, intelligent rings, intelligent headbands, intelligent bracelets, intelligent glasses, smartwatches, and intelligent antiloss devices. Easy operation is the basic requirement of intelligent equipment, and intelligence is its development trend. According to the application functions of the equipment, it is divided into the following five categories.

The first category is life sports health equipment. This kind of equipment is mainly used to collect human movement data, physical fitness health data, and life work and rest data. The main form of the product is the bracelet. The main functions of the products are heart rate monitoring, exercise measurement, blood oxygen measurement, sedentary reminder, sleep quality monitoring, exercise mode, body surface temperature monitoring, and calorie consumption. For example, obese people can use the device's exercise record and calorie consumption calculation function to make a practical exercise plan for themselves [16]. Under the supervision of the equipment, they can control the exercise heart rate, achieve their own exercise goals, and improve the efficiency of physical exercise. This kind of equipment can help people better understand their physical conditions, find their body abnormalities in time, and track the change process of various indicators of the body. In the current environment where people pay more and more attention to their health problems, life sports health products are an important part of the intelligent wear market.

The second category is leisure and entertainment equipment, mainly to meet people's daily leisure and entertainment needs; the main forms of products are watches and glasses. For example, Google glasses can get information notification and remind in time to avoid missing important information; at the same time, it supports voice input, so that you can process the important information received in time, without affecting the use of some mobile phones when it is inconvenient. Also, Google glasses have a convenient photo-taking function, which can help people capture beautiful

moments in time. Another example is VR intelligent virtual glasses, with a high-definition display function, which can make game pictures and movie pictures more realistic, enhance people's experience, and 3D vector sound field effect, which makes people feel like they are in the scene [17]. There are also smartwatch products and smartwatches with extended functions of mobile phones, voice navigation, message push, call reminder, and intelligent alarm clock functions so that people's lives are more convenient.

The third category is safety protection equipment, which mainly meets the needs of the elderly and children. Children's products are mainly reflected in children's watch products. Guardians can remotely monitor children's phone calls. Through the GPS positioning system, the specific location of children can be known, and the family number can be set. When necessary, children can be contacted in time, and communication can be joined through control record friends and refuse strangers to add, to further ensure the safety of children's external environment. Compared with the mobile phone, the telephone watch also plays a communication function, but it is more convenient for children to use. At the same time, the smartphone watch also has the function of an alarm when the watch falls off, which makes the phone watch a real tool for parents to find their children in time. The safety products for the elderly mainly use the GPS positioning function to let the family know the whereabouts of the elderly in time to prevent loss and accidents.

The fourth category of products is medical and healthcare equipment, mainly to meet the needs of specific groups of people to monitor their health indicators [18]. For example, the household finger clip measuring instrument is equivalent to a mobile "physical examination doctor." It can detect the basic health indicators of the human body at any time, including heart rate, blood oxygen, blood pressure, respiratory rate, PI, and other indicators, which can be automatically detected in 60 seconds, upload data records, generate physical examination reports, and provide analysis suggestions and is a health housekeeper. The data recording function of the device can generate weekly reports, monthly reports, and annual reports, which is convenient for users to know their past health status at any time.

The fifth category is smart home products, which mainly meet the control needs of household appliances and furniture in daily life [19–23]. Through WiFi or other wireless ways, the control of household appliances such as the water heater, TV, and air conditioner can be realized to save electricity and avoid fire accidents. At the same time, HD equipment can be used to intelligently monitor the situation of the home; for example, there are nannies taking care of children in the family, worrying about the health of the elderly, worrying about the pets in the family, and worrying about theft of property in the family by thieves. Effective protective measures can be taken for all of these through the intelligent device. Through the use of smart home products, people can enjoy home life more intelligently.

Supervising the movement and adjusting the movement state: intelligent wearable devices can transmit real-time motion parameters to users, such as heartbeat, blood pressure, and energy consumption. The current intelligent wearable devices can monitor the user's activity parameters and timely transmit the real-time motion parameters of users to the intelligent wearable devices. Therefore, during aerobic exercise such as running, swimming, and cycling, users can monitor their heart rate, blood pressure, and other information at any time. In the process of sports, according to these parameters, we can adjust our own sports situation so that the sport's participants can adjust their state at any time and control their sports.

7.1. Correct Movement and Efficient Training. Recently, some companies have made deeper development in the field of wearable technology. The wearable devices of LUMO LFT and Wa Hoo can capture various movements of the arms and legs. By using artificial intelligence algorithms and sensors that can measure the motion conditions in 3D space, we can compare the body posture of users in swimming, running, or archery with the motion data collected from professional athletes by ideal motion models. Therefore, for the beginners of sports, comparing these data parameters can show how to regulate their actions, and the learning process of sports can also be more scientific and efficient [24–27].

7.2. Prevention of Sports Injury of Athletes. The intelligent wearable device can transmit the athlete's body data to the medical and healthcare personnel in real time. According to the athlete's physical condition, the sports training is specially customized, and the athlete's fatigue condition is analyzed and modeled to understand the rise and fall of the athlete's competitive state in real time. Based on this, the athletes can maintain their physical condition in the best competitive state. Through the collection of athletes' sports data, big data can be used for processing. Athletes can also collect sports data through wearable devices in the usual competition. The data modeling can be carried out for the physical condition data of athletes before the injury so that medical and healthcare personnel can track the physical condition of athletes in real time, and early warning can be timely carried out before athletes are injured, to reduce sports injury.

7.3. Improving the Competitive Level of Athletes. The future is an "Internet of Things era"; that is, people can control the automobile, electrical appliances, and other equipment through intelligent devices. In the field of sports, intelligent wearable devices can be designed intelligently with the equipment used by athletes and the equipment of athletes. When athletes use competition equipment, data collection equipment is added to the equipment. At the same time, it is combined with the intelligent wearable equipment of athletes, so that the sports mobilization can understand the advantages and disadvantages of their actions more correctly and the athletes and competition equipment can be realized. Now, "character interconnection," for example, ballplayers can correct their actions according to the feedback of sensors

on the football and the hitting point information, is used to continuously improve their competitive level. At the same time, the intelligent glasses in the intelligent wearable device can reproduce the virtual scene and help the athletes better integrate into the atmosphere of the competition. "The NFL (National Football League) has planned to introduce intelligent wearable technology to the NFL arena, and NFL players are using wearable devices to prepare for their professional sports careers." With the development of science and technology, intelligent wearable devices can be more and more refined, recording sports data without the athletes' awareness, and constantly improve the competitive level of athletes.

8. Conclusions

With Apple's release of the Apple watch, 2015 is known as the first year of intelligent wearable devices. The era of intelligent wearable devices has come, and the impact on sports will be far reaching. However, the development time of intelligent wearable devices is relatively short, and there are still many problems to be solved, such as endurance, comfort, data acquisition, and other issues, which are still an important factor restricting the development of intelligent wearable devices. In all aspects of sports, intelligent wearable devices can be perfectly combined with it, and the development of intelligent wearable devices can better promote the development of sports, making sports more scientific, efficient, and easy to learn. Intelligent wearable devices have begun to penetrate all aspects of sports. Although intelligent wearable devices have just entered the initial stage of development, every scientific and technological progress requires constant attempts. The proposed study analyzes the application status and prospects of big data in the field of health management. To study big data applications in football players' mental health, monitoring systems and chronic disease health management systems that also have a good prospect are presented in this study. The application of an intelligent wearable device for several aspects of sports work is part of this study. Therefore, with the continuous development of science and technology, in the future, intelligent wearable devices will occupy an important position in the field of sports. The effectiveness of the proposed study is shown through experimental results.

Data Availability

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

Conflicts of Interest

The authors declare that they have not conflicts of interest regarding the publication of this paper.

Acknowledgments

This work was supported by "the Fundamental Research Funds for the Central Universities," North Minzu University (2019XYSTY01).

References

[1] M. Jiang, *Thoughts and Suggestions on the Development Prospect of Intelligent Wearable Devices Mobile Communication*, Springer, Berlin, Germany, 2014.

[2] L. I. U. Feng, H. A. N. Jing-Long, Q. I. Ji, Y. U. Jia-Luo, L. I. Wen-Peng, and L. I. Bo-Wei, "Research and application progress of intelligent wearable devices," *Chinese Journal of Analytical Chemistry*, vol. 49, no. 2, pp. 159–171, 2021.

[3] Y. Li and J. Cao, "WSN node optimal deployment algorithm based on adaptive binary particle swarm optimization," *ASP Transactions on Internet of Things*, vol. 1, no. 1, pp. 1–8, 2021.

[4] M. Chen, Y. Ma, Y. Li, D. Wu, Y. Zhang, and C.-H. Youn, "Wearable 2.0: enabling human-cloud integration in next generation healthcare systems," *IEEE Communications Magazine*, vol. 55, no. 1, pp. 54–61, 2017.

[5] J. Zhang, J. Sun, J. Wang, and X.-G. Yue, "Visual object tracking based on residual network and cascaded correlation filters," *Journal of Ambient Intelligence and Humanized Computing*, vol. 20, 2020.

[6] R. Liu, X. Ning, W. Cai, and G. Li, "Multiscale dense cross-attention mechanism with covariance pooling for hyperspectral image scene classification," *Mobile Information Systems*, vol. 14, 2021.

[7] X. Ning, F. Nan, S. Xu, L. Yu, and L. Zhang, "Multi-view frontal face image generation: a survey. concurrency and computation," *Practice and Experience*, vol. 24, Article ID e6147, 2020.

[8] W. Cai, Z. Wei, R. Liu, Y. Zhuang, Y. Wang, and X. Ning, "Remote sensing image recognition based on multi-attention residual fusion networks," *ASP Transactions on Pattern Recognition and Intelligent Systems*, vol. 1, no. 1, pp. 1–8, 2021.

[9] X. Ning, Y. Wang, W. Tian, L. Liu, and W. Cai, "A biomimetic covering learning method based on principle of homology continuity," *ASP Transactions on Pattern Recognition and Intelligent Systems*, vol. 11, no. 1, pp. 9–16, 2021.

[10] Z. Huang, Y. Zhang, Q. Li et al., "Joint analysis and weighted synthesis sparsity priors for simultaneous denoising and destriping optical remote sensing images," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 58, no. 10, pp. 6958–6982, 2020.

[11] Q. Liu, L. Cheng, A. L. Jia, and C. Liu, "Deep reinforcement learning for communication flow control in wireless mesh networks," *IEEE Network*, vol. 35, no. 2, pp. 112–119, 2021.

[12] B. Su, *Design and Implementation of Human Posture Monitoring System Based on Acceleration Sensor*, Northeast University, Boston, MA, USA, 2011.

[13] F. Shuntian, "Development status and trend of wearable devices," *Research and development*, vol. 53, no. 3, 2014.

[14] H. Li, "Discussion on the application prospect of wearable devices in the field of leisure sports," 2015.

[15] J. Deng, H. Liu, X. Yang, and Q. Cheng, "Current situation and future development trend of wearable intelligent devices," 2016.

[16] Q. Han, "Some thoughts on promoting the application of big data under the new situation," *Scientific and Technological Innovation and Application*, vol. 9, 2017.

[17] S. Zhao and X. Xu, "The application of wearable devices in Physical Education Teaching -- an analysis based on the SWOT model," *Contemporary Sports Science and Technology*, vol. 6, no. 34, pp. 111-112, 2016.

[18] G. Tang and H. Qin, "Min. Extension of Internet plus to physical education -- Application of smart wearable

equipment in school physical education," *Sports Technology*, vol. 37, no. 3, pp. 143-144+150, 2016.

[19] Z. He, *Design and Implementation of the Human Energy Consumption Detection System Based on the Acceleration Sensor*, Harbin Institute of Technology, Harbin, China, 2016.

[20] B. Yang, *Analysis of the Current Situation and Future Development Trend of Wearable Devices Wuhan Decision Information Research and Development Center, Decision, and Information magazine*, School of economics and management, Peking University, Beijing, China, 2015.

[21] J. Deng, H. Liu, X. Yang, and Q. Cheng, "Current situation and future development trend of wearable intelligent devices," *Heilongjiang Science and Technology Information*, vol. 28, p. 135, 2015.

[22] Y. Zhao and T. Zhang, "Current situation analysis of intelligent wear in the health field [small," *Chinese Journal of Health Information Management*, vol. 12, no. 4, pp. 354–358, 2015.

[23] B. Li, X. Yu, X. Yang, and Y. Zheng, "Wearable exercise intensity monitoring system," *Computer System Application*, vol. 24, no. 5, pp. 32–39, 2015.

[24] C. Wang, *Design and Implementation of Wearable Multi-Sensor Human Daily Activity Monitoring System*, Nanjing University of Technology, Nanjing, China, 2015.

[25] P. Wang, "Fitness monitoring is an important market for wearable devices," *Import and Export Manager*, vol. 10, pp. 46-47, 2014.

[26] Y. Zhuang, "Sports and health products lead to the development of the wearable device market," *Electronics World*, vol. 21, no. 6, pp. 3–6, 2014.

[27] Y. Geng, Hui, L. Yang, and H. Jiang, "Development status and prospect of wearable devices," *Electronic Science and Technology*, vol. 10, no. 2, pp. 238–245, 2014.

Hindawi
Scientific Programming
Volume 2021, Article ID 7273981, 9 pages
https://doi.org/10.1155/2021/7273981

Research Article

Research on IT English Flipped Classroom Teaching Model Based on SPOC

Lili Liu

School of Foreign Languages, Xuchang University, Xuchang 461000, China

Correspondence should be addressed to Lili Liu; lll2010@xcu.edu.cn

Received 23 April 2021; Revised 21 May 2021; Accepted 31 May 2021; Published 7 June 2021

Academic Editor: Shah Nazir

In order to improve the existing problems in the teaching process of IT English courses and improve the quality of IT English, this paper conducts a research on the flipped classroom teaching mode of IT English based on SPOC. The present situation of IT English teaching is analyzed, and the problems existing in the teaching process are also analyzed. On the basis of the above thought, with the support of SPOC platform build IT turn English classroom teaching mode, namely, by setting the course target, learning for class in advance, choose high-quality class, learning information collection, upload the related resources, and do a good job in teaching design complete teacher preparation, and design the specific teaching unit of teaching process, teaching quality evaluation model was constructed. In this way, the teaching quality evaluation results of IT English flipped classroom are obtained, in order to further improve the teaching quality. The experimental results show that the flipped English teaching mode based on SPOC can effectively improve students' performance and increase students' average daily learning time and course satisfaction, and the practical application effect is good.

1. Introduction

Relevant policies in China propose strengthening the cultivation of international talents. Meanwhile, the outline clearly points out that a group of talents with international vision, familiar with international rules and able to participate in international affairs and international competition, should be cultivated [1, 2]. With the development of global economic integration, more and more IT companies and II enterprises in China are actively exploring the international market. Therefore, the importance of cultivating IT talents suitable for the international market for China is increasingly prominent [3].

The IT talents required by the international market should not only have solid professional knowledge but also have the ability to communicate professional knowledge in cross-cultural context using English fluently. IT English courses in colleges and universities aim to cultivate students' ability to use foreign language to learn professional knowledge and solve professional problems. Meanwhile, they also focus on improving students' ability to use foreign language to communicate and communicate professional knowledge [4, 5]. IT English is an important course to train students to read and understand the software engineering industry, which can improve students' ability to write English software development documents and oral English communication in the environment of project development, so as to make students more competitive in international software companies after graduation [6].

At present, as far as the status quo of higher education in China is concerned, the IT English teaching in most colleges and universities still adopts the traditional "teacher-oriented" cramming teaching mode, which leads to the general deficiency of students' comprehensive professional English application ability, teamwork ability, and oral expression and communication ability. The main reason for this is that China's current professional English teaching materials have many problems, such as the disconnection between the content of the teaching materials, the teaching mode of the courses, and the actual project development [7]. In this situation, in order to promote the training of international IT talents, the innovative research on the teaching mode and

method of IT English classroom is a new topic that needs to be studied urgently [8].

In today's international context, how to promote the IT English course teaching method reform and optimization is very urgent and significant, and IT English teaching reform is a new challenge and opportunity. Therefore, this paper studies the flipped classroom teaching mode of IT English based on SPOC in order to improve the teaching quality of IT English and cultivate more IT talents for the society. The following are the key contributions of the proposed study:

(i) To conduct a research on the flipped classroom teaching mode of IT English based on SPOC.

(ii) The present situation of IT English teaching is analyzed, and the problems existing in the teaching process are also analyzed.

(iii) With the support of SPOC platform build IT turn English classroom teaching mode, namely, by setting the course target, learning for class in advance, choose high-quality class, learning information collection, upload the related resources, and do a good job in teaching design complete teacher preparation, and design the specific teaching unit of teaching process, teaching quality evaluation model was constructed.

(iv) The teaching quality evaluation results of IT English flipped classroom are obtained, in order to further improve the teaching quality.

(v) Experimental results of the proposed study show that the flipped English teaching mode based on SPOC can effectively improve students' performance and increase students' average daily learning time and course satisfaction, and the practical application effect is good.

2. IT English Flipped Classroom Teaching Mode

2.1. Status Quo of IT English Teaching. The contradiction between the IT market's demand for high-quality talents and the unsatisfactory status quo of IT English teaching requires that the unsatisfactory status quo of IT English teaching be changed and the causes of the unsatisfactory status quo of IT English teaching be analyzed in order to realize the change [9].

(1) There is a big difference in the quality of students and their learning quality is not high.

Students' learning foundation and learning attitude affect the teaching effect of IT English to a great extent. When students come to school, their English level is uneven. In order to facilitate teaching, teachers often set uniform goals in teaching and adopt the same teaching methods for all students. This led to some interesting learning for students who could keep up. However, students who cannot keep up with IT English will hold a fear of difficulties, and the more afraid they become, the more difficult they become, and the result is that they simply give up [10, 11]. In addition, students' learning attitude is not correct. They think that as long as they learn professional courses well, IT English is not important. Of course, some students want to learn IT English well, but they fail to balance the time between IT major and IT English, so that they fail to learn IT English persistently [12].

(2) Teachers do not have advantages in quality, and knowledge and methods are updated slowly.

The IT industry is originally a new industry; there have been no normal colleges to train IT English major teachers, generally computer or English major teachers through advanced study or training to lead the course. Therefore, the basic quality of teachers does not have a great advantage. In addition, with the rapid development of IT industry and the rapid updating of knowledge, teachers are required to update knowledge regularly and in a planned way, but schools and society fail to create a good platform for teachers to update their knowledge [13].

(3) The content of the textbook lags behind the actual development of IT industry English.

A good teaching effect must rely on good teaching materials, and a good set of teaching materials often need a long period of time to be produced, but the IT industry develops rapidly and the industry language vocabulary is updated quickly, so the IT English teaching materials often cannot keep up with the development of the IT industry [14].

2.2. IT English Flipped Classroom Teaching Mode Based on SPOC. In 2013, Professor Armando Fox of the University of California at Berkeley proposed SPOC (small private online course). SPOC can not only supplement classroom teaching but also completely replace classroom teaching. It can improve the utilization of classroom and improve students' learning ability and students' participation. By definition, SPOC's small corresponds to MOOC's massive and private corresponds to open. It can be seen that MOOC is mainly aimed at learners from all over the world, while SPOC pays more attention to learners on campus. Flipped classroom refers to readjusting the time inside and outside the classroom and transferring the decision-making power of learning from teachers to students. In this kind of teaching mode, with precious time in the classroom, students can focus more on active project-based learning and jointly study and solve the challenges of localization or globalization and other real-world problems, so as to obtain a deeper understanding. The basic idea of SPOC based IT English flipped classroom teaching mode is to make use of excellent MOOC teaching resources, change the traditional teaching mode, promote blended teaching and student participation, and further improve the teaching quality [15, 16].

The main purpose of this paper is to make high-quality MOOC resources effectively used in IT English teaching through SPOC and flipped classroom, so as to improve the teaching efficiency. IT English flipped classroom teaching

mode includes three parts. The first part is the teacher's preparation before taking this course, including setting course objectives, learning MOOC in advance, selecting high-quality MOOC, collecting learning materials, uploading relevant resources, and doing a good job in teaching design. The second part is the teaching process of a specific teaching unit [17], including before class, in class, and after class, cycle, until the end of the course. The third part is through the construction of teaching quality evaluation model, in order to obtain the evaluation results of IT English flipped classroom teaching mode, in order to improve the existing teaching mode and improve the teaching quality [18].

2.2.1. Teacher Preparation. Before the beginning of the semester course, although the teaching of knowledge points can be replaced by MOOC teaching video, instead of preparing lessons like traditional teaching, teachers need to prepare carefully in advance in the face of new teaching methods. The specific process is shown in Figure 1.

(1) According to the syllabus and students' prior knowledge, teachers set the curriculum objectives and unit teaching objectives.

(2) *Learn MOOC in Advance*. At present, colleges and universities are not the main body of the construction of MOOC, and teachers are not the creators of the course. Therefore, in order to choose the appropriate MOOC, we need to learn the relevant MOOC courses in advance.

(3) *Choose High-Quality MOOC*. The existing MOOC resources are very rich, so teachers need to consider many aspects in the selection, such as the teaching objectives of the courses taught, the content of the textbooks used, and the students' knowledge level to select the most suitable MOOC [19].

(4) *Collect Learning Materials*. Before the class, teachers should record supplementary videos in advance, because the original purpose of MOOC is not just to let the students in the class learn nor is it based on the teaching materials stipulated by our school, so teachers need to be prepared for the parts of teaching videos that do not involve the contents of teaching materials. In addition, we should prepare learning materials for students to read, such as books, websites, and electronic materials, so as to give students more choices of learning resources.

(5) *Do a Good Job in Teaching Design*. The in class stage is carried out in the classroom environment. But students are the main body of the classroom, and teaching activities and students' learning activities are very different. Teachers need to change the previous teaching design, from the teaching objectives, teaching process, and other aspects of detailed teaching design; the effect of teaching design directly affects the students' complete construction of the course content. The preparation of teachers is the key to flipped teaching in SPOC environment. The five links mentioned in this section require teachers to work hard. The more well-prepared the teachers are in the five links, the smoother the IT English course will be and the better the effect will be [20].

2.2.2. Implementation of Unit Teaching. IT English course consists of many teaching units. The so-called unit teaching is a relatively complete unit of knowledge teaching; unit teaching can be a chapter or a section of the course or a knowledge point or a lesson. We divide a unit teaching process into preclass, in class, and after class.

(1) Preclass Stage. The preclass stage is the knowledge imparting stage of this teaching mode. Teachers prepare IT English teaching videos and resources on SPOC platform in advance, and students carry out autonomous learning. In this stage, both teachers and students need to complete many learning links. The specific process of preclass stage is shown in Figure 2.

For IT English teachers, they should first set the teaching objectives of the unit according to the curriculum objectives and syllabus and then compare the selected MOOCS with the teaching materials in detail. If the teacher only publishes similar contents when arranging MOOCS resources but does not match the teaching materials, some students will feel that they do not want to learn or some of the knowledge points in the original MOOCS video cannot be easily understood, which will inevitably reduce students' interest in learning. Therefore, teachers should carefully review the teaching content every week. For the knowledge points not involved in the original MOOCS, teachers should record teaching videos or use other authorized video resources. At the same time, in order to make students better understand the knowledge points, teachers can add some other forms of IT English learning resources (e.g., learning website links, e-books, case materials, etc.) on SPOC platform for students to learn. At the same time, according to the content of the teaching materials, teachers should check and adjust the classroom tests on the original MOOCS and carefully design them. In addition, teachers should participate in the interaction with students before class, try to reply to students as soon as possible, and answer questions online for students. In this process, teachers can also learn the students' mastery of knowledge points, which knowledge is difficult or has been mastered, which students master better, and which students have difficulties in learning. Teachers should make full use of the benefits of discussion. For students, they should first understand the teaching objectives of the unit and then register and log on to SPOC platform to watch the IT English teaching videos of the unit in the selected MOOC, browse various resources, and complete the classroom test according to the requirements of MOOC. It is worth noting that MOOC platform provides online discussion, which can solve the cognitive construction of students' basic knowledge. When students encounter problems that cannot be solved in their study, they can publish topics in the discussion area. Students can help each other to answer them. Generally, some basic knowledge problems can be effectively

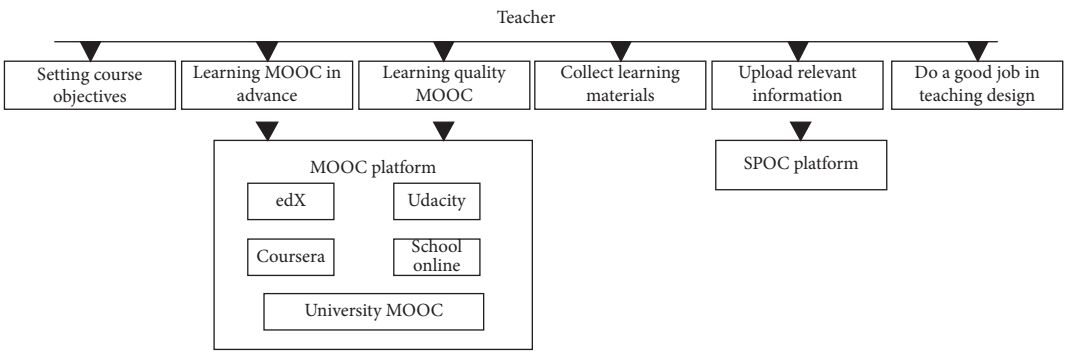


Figure 1: Preparation process of teachers.

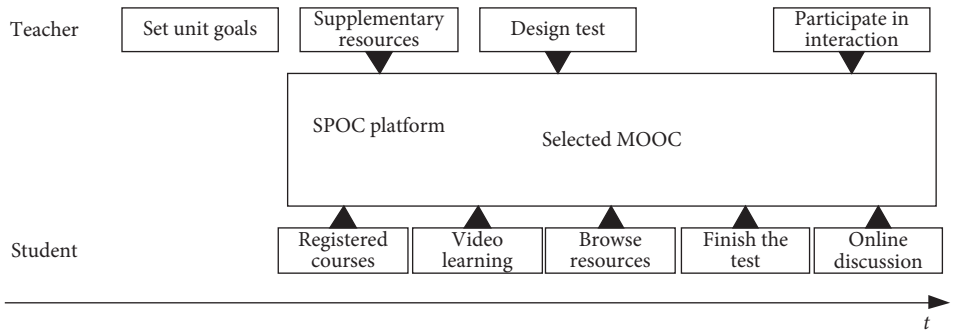


Figure 2: Specific process of preclass stage.

solved. In the discussion area, students can also publish their feelings about the learning of this unit. For example, after learning on the platform, they cannot deeply grasp some knowledge points, and they hope that teachers can further explain and discuss them in class. This is also for teachers to better grasp the learning situation of students on the MOOC platform and id conducive to the development of teaching activities.

(2) In Class Stage. The in class stage is the complete knowledge construction stage of IT English curriculum model. In the classroom environment, students should not only master the knowledge completely but also cultivate their abilities in all aspects, such as the ability to solve comprehensive problems independently and creative thinking ability. Therefore, the in class stage is divided into two parts: one part is to discuss the key and difficult points and clarify the knowledge points, so that students can systematically perceive the knowledge points of this class, and construct the knowledge completely, so as to better consolidate the knowledge; the other part includes the activities of raising questions, thinking alone, determining problems, small group discussion, reporting and communication, teachers' puzzles, and summary and evaluation. In this process, we should firmly grasp the knowledge points and improve the ability of comprehensive application. The specific process of the in class stage is shown in Figure 3.

For teachers, first of all, at the beginning of the IT English class, they spend less time with students to discuss the important and difficult points of the content learned in the class, understand the questions that students have not solved and the knowledge points that are difficult to understand, and sort out the content and knowledge points as a whole to help students complete the construction of knowledge. Then most of the rest of the class time is used to cultivate students' comprehensive ability. In order to cultivate students' ability to use knowledge to analyze and solve problems, teachers need to put forward multilevel comprehensive problems. One is slightly more difficult than the general problems, so as to improve students' understanding of unit knowledge. The other is the problem of the comprehensive use of knowledge points, which is used to cultivate students' comprehensive problem-solving ability and improve their high-level thinking. Let the students think independently first. If they cannot solve the problem by themselves, organize the students to discuss these problems in groups. After listening to the students' reports, if the problems have not been effectively solved, the teacher needs to give detailed answers. Finally, according to the classroom situation, the teacher made a summary. For students, first of all, they need to actively participate in the process of discussing important and difficult points and knowledge points with teachers and consult the teachers about the problems they fail to understand in MOOC learning, so as to actively complete the knowledge construction. Then, I will think independently of the questions raised by the teacher and try to solve the problems with the knowledge I have learned in the MOOC. If I cannot solve the problems independently, I will discuss them in groups and actively communicate with my classmates. In the reporting section, students should make a

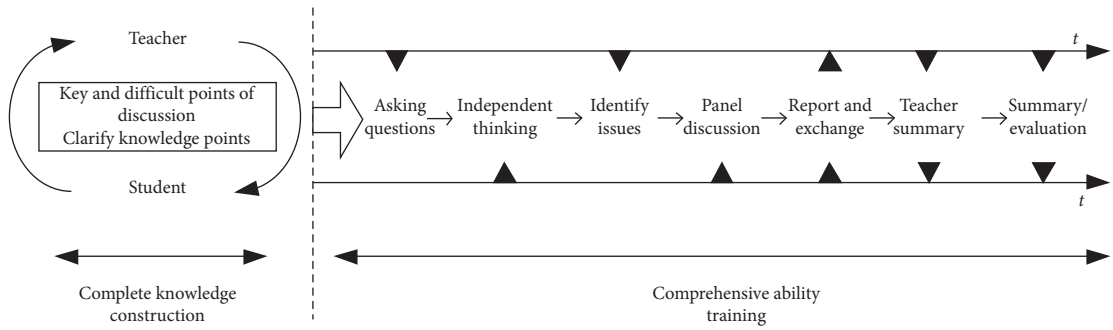


Figure 3: The specific process of the middle class.

report boldly, raise the problems encountered by the group in the discussion process, and listen carefully when the teacher explains the problems and summarizes them, so as to improve their knowledge and ability points.

(3) The After Class Stage. IT is stage after class English teaching mode of study effect test stage, in order to get more accurate results, in addition to the original reference for design of a lesson homework exercises, the teacher on the homework and unit test according to the difficult point problem sets, and knowledge to design to measure whether the overall level of students reached the standard of learning into the next unit. The specific tasks and activities of teachers and students in the after class stage are shown in Figure 4.

For teachers, they should first complete relevant exercises and tests. On this basis, according to the IT English teaching objectives of this unit, homework and unit test questions should be designed from two aspects of knowledge and ability. It is worth noting that, in order to grasp the actual level of students, according to the teaching practice, teachers should appropriately select some topics beyond the teaching objectives when designing homework and unit tests. It is important to point out that students in classroom testing cannot only rely on the class assignments and unit tests they originally desired. This is because the original students who desire classroom teaching are a large number of learners with very different levels of knowledge in all aspects of the world. If you blindly rely on these, you will not be able to accurately understand the real learning situation and effect. Teachers analyze the results of unit tests to give personalized guidance to students who fail to reach the teaching objectives. For students, first of all, they need to complete the homework and unit tests on the SPOC platform, give feedback to teachers to solve the wrong exercises in the homework and tests, and review them accordingly. At the same time, they actively discuss with classmates, help each other, and constantly master and consolidate the knowledge.

2.2.3. Teaching Quality Evaluation Method. The third part is to obtain the teaching quality evaluation results of IT English flipped classroom through the construction of teaching quality evaluation model, in order to improve the existing teaching model and improve the teaching quality. The flipped classroom teaching of IT English is affected by a variety of factors, each of which has different influence on the teaching effect, so IT is difficult to establish an accurate mathematical model. Let the teaching quality evaluation index of IT English flipped classroom be $\{x_1, x_2, \ldots, x_m\}$. The first step to establish the teaching quality evaluation model of IT English flipped classroom to establish an evaluation index system. A reasonable and scientific evaluation index system can improve the accuracy of teaching quality evaluation, but the teaching quality evaluation is affected by a variety of factors, for example, teaching method, teaching attitude, teaching content, and management and teaching effect, and the evaluation index system of teaching quality was established by referring to relevant literature and research, and then the hierarchical method was adopted to analyze and establish the teaching quality evaluation index system, as shown in Table 1.

When using big data technology to calculate the weight value of teaching quality evaluation index of IT English flipped classroom, Hadoop computing mode is used to calculate the evaluation weight value according to the analytic hierarchy process. First, the judgment matrix is constructed, and the judgment matrix W is expressed as

$$W = \left(w_{ij}\right) n * n, \tag{1}$$

where n is the number of horizontal and vertical lines in the matrix, and w_{ij} is the degree coefficient, which is generally as in [1, 10]. Therefore, for formula (1) above, the degree of influence can be expressed as

$$w_{ji} = \frac{1}{w_{ji}}, \quad (i, j = 1, 2, \ldots, n). \tag{2}$$

Normalize the data of each column in judgment matrix W to get matrix W', add each row of W' according to the row to get vector w, use the sum product method to find the maximum eigenroot of the vector, and get

$$\lambda_{\max} = \frac{1}{n} \sum_{i=1}^{n} \frac{W \cdot w^T}{w_{ij}}, \tag{3}$$

where $\lambda_{\max}$ is the maximum characteristic root and T is the coefficient of the summation product method. Using the random consistency ratio CR to judge the consistency of the largest characteristic root, the following calculation can be obtained:

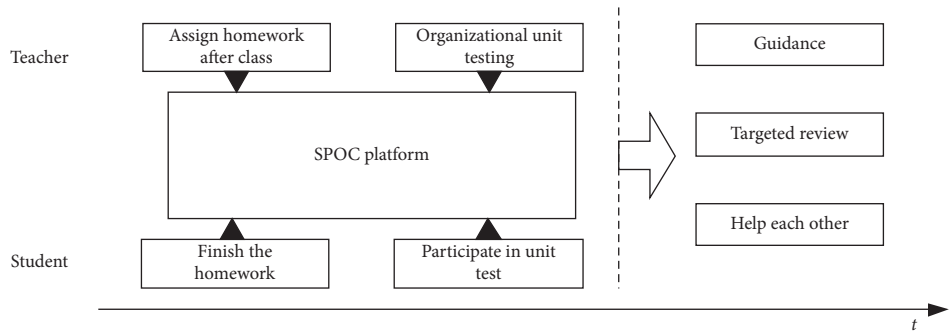


Figure 4: Specific tasks and activities of teachers and students in the after class stage.

Table 1: Teaching quality evaluation index system.

Level indicators	Secondary indicators
Teaching content	Teaching objectives (x_1) Scientific nature of teaching content (x_2) Teaching key and difficult points processing (x_3)
Teaching method	Good at heuristic thinking (x_4) Taking care of personality differences (x_5) Learning method guidance (x_6) Teaching tools and means (x_7)
Teaching attitude	Course preparation (x_8) Degree of engagement (x_9) Personal quality of teachers (x_{10})
Teaching effect	Teaching plan completion (x_{11}) Students' mastery of the teaching content (x_{12})

$$\mathrm{CR} = \frac{\mathrm{CI}}{\mathrm{RI}}, \tag{4}$$

where CI is the consistency index and RI is the average random consistency index. Therefore, the weight value of the final teaching quality evaluation index can be calculated according to the consistency index:

$$\mathrm{CI} = \frac{\lambda_{\max} - n}{n - 1}. \tag{5}$$

In the above formula, n is the order of the judgment matrix. Formulas (3) and (5) are put together, and the comprehensive random consistency table finally obtains the following: When matrix W meets $\mathrm{CR} < 0.1$, the characteristics of the normalized vector can be used as a weighted vector; the final calculation result of formula (2) is the weight of the teaching evaluation index, and the calculation process is repeated to calculate the application-oriented undergraduate teaching evaluation index and its weight value $K = \{k_1, k_2, \ldots, k_m\}$; then the mathematical expression of teaching quality evaluation model is

$$y = f\left(k_1 x_1, k_2 x_2, \ldots, k_m x_m\right). \tag{6}$$

In the formula, f represents the evaluation function. The teaching quality evaluation model mentioned above is used to evaluate the teaching quality evaluation of IT English flipped classroom, and the corresponding evaluation results are output, in order to improve the existing problems of IT English flipped classroom teaching model and further improve the teaching quality.

3. Experimental Design and Result Analysis

3.1. Experimental Design. In order to verify the effectiveness of the SPOC-based IT English flipped classroom teaching mode, IT needs to be designed. The specific experimental scheme is shown below.

In order to ensure the scientific nature and reliability of the experimental results, the experiment needs to be carried out in a unified experimental environment. The specific experimental environment is shown in Table 2.

1000 students in a certain university were selected as the research objects, and the research objects were divided into two groups on average. One group used the traditional IT English teaching mode and was designated as experimental group A. The other group used the IT English flipped classroom teaching mode proposed in this paper, which was recorded as experimental group B, and the changes of various data of students in all groups were recorded and taken as experimental data. The basic data of students are shown in Table 3.

The educational background of students is shown in Table 4.

3.2. Analysis of Experimental Results. Twelve months after the experiment was compared, the average scores of the students in experiment A and experiment B were measured. The results are shown in Table 5.

By analyzing the data in Table 5, it can be seen that the average score of experimental group A presents wavy changes, while the average score of experimental group B presents a steady upward trend. When the experiment lasted for 9 months, the average score of experimental group A reached the maximum of 69.6, an increase of 3.3 compared with the initial average score of students. However, when the experiment lasted for 12 months, the average score of experimental group B reached the maximum of 81.3, an increase of 14.5 compared with the initial average score of students, indicating that, compared with experimental group A, the average score of experimental group B improved more significantly. It shows that the flipped English teaching mode based on SPOC can effectively improve students' performance and verifies the effectiveness of this teaching mode.

On the basis of the above experiments, the average daily learning duration of students during the experiment period is compared, and the specific results are shown in Table 6.

Table 2: Experimental environment.

Name	Describe
Operating system	Windows7
CPU	Intel Core i5-7300HQ
Memory	32 G
Hard disk	500 GB
Run a memory	8 G
Frequency	2.1 GHz

Table 3: Basic data of students.

The basic data	Experimental group A	Experimental group B
Boys	236	243
The number of girls	264	247
Initial average score of students (full mark is 100)	66.3	66.8
Average learning length of initial day (hours)	1.59	1.52

Table 4: Education background of students.

Record of formal schooling	Experimental group A	Experimental group B
A freshman	125	116
A sophomore	121	135
Junior year	116	98
Senior year	85	111
A graduate student	35	26
The master	12	7
A PhD student	6	7

Table 5: Comparison of average scores of students.

The experimental time	Experimental group A	Experimental group B
1 month	67.2	67.4
2 months	66.3	68.3
3 months	65.9	69.2
4 months	68.4	70.4
5 months	69.2	72.6
6 months	66.3	73.4
7 months	67.5	74.5
8 months	68.6	76.9
9 months	69.6	78.1
10 months	66.8	79.9
11 months	67.2	80.2
12 months	67.6	81.3

By analyzing the data in Table 6, we can see that the average daily learning time of the students in experimental group A presents a wave like change, while the average daily learning time of the students in experimental group B presents a steady upward trend. When the experiment lasted for 11 months, the average daily learning time of the students in group a reached the maximum of 1.77 hours, which only increased by 0.18 hours compared with the initial average daily learning time. In the 12 months of the experiment, the average daily learning time of the students in the experimental group B reached the maximum of 2.37 hours, which only increased by 0.85 hours compared with the initial average daily learning time, indicating that the average daily learning time of the students in experimental group B increased significantly, indicating that the IT English flipped classroom teaching mode based on SPOC constructed in this paper can effectively stimulate students' interest in learning and improve students' average daily learning time.

Finally, after the results of the experiment, the satisfaction degree of the students in experimental group A and experimental group B with the existing IT English teaching mode was compared. The full score was 100. Each group of students was divided into 10 groups on average, and the satisfaction score of each group was averaged. Student satisfaction scores are shown in Table 7.

By analyzing the data in Table 7, it can be seen that the average degree of satisfaction of experimental group A is

Table 6: Comparison of average daily learning length of students (hours).

Experimental time	Experimental group *A*	Experimental group *B*
1 month	1.61	1.59
2 months	1.72	1.63
3 months	1.69	1.68
4 months	1.73	1.69
5 months	1.68	1.72
6 months	1.75	1.75
7 months	1.74	1.81
8 months	1.71	1.85
9 months	1.69	1.96
10 months	1.68	2.17
11 months	1.77	2.25
12 months	1.72	2.37

Table 7: Student satisfaction scores.

Experimental group	Experimental group A	Experimental group B
Team 1	67.8	92.1
Team 2	66.3	90.3
Team 3	69.5	91.5
Team 4	67.8	92.7
Team 5	69.5	93.3
Team 6	67.7	94.1
Team 7	69.3	93.2
Team 8	70.3	90.6
Team 9	72.4	91.3
Team 10	70.1	95.6
Average	69.1	92.5

69.1, and that of experimental group B is 92.5. The average degree of satisfaction of experimental group B is 23.4 higher than that of experimental group A, which indicates that students in SPOC based flipped classroom teaching mode can be accepted more, which can effectively stimulate students' interest in learning and improve students' interest in IT English teaching satisfaction of language flipped classroom teaching mode.

4. Conclusion

Due to the rapid development of computer network technology, mankind has been brought into the "information age," and the exchanges of science and technology and culture between China and foreign countries have become increasingly frequent. Under the background of globalization, the domestic IT industry is trying to integrate with the world, and a large number of talents who can understand professional technology and have fluent communication with foreign IT professionals are needed. This requires higher vocational colleges to do a good job in IT English teaching, so as to create a solid foundation for students' career development. In short, in recent years, there has been a great demand for talents who are proficient in IT industry knowledge and have good English skills. Meanwhile, with the development of technology, the requirements for these talents are getting higher and higher. In order to improve the quality of IT English teaching, this paper constructs a new flipped classroom teaching mode of IT English based on SPOC and verifies the effectiveness and superiority of this model through experiments. Therefore, this model can effectively improve the effectiveness of IT English teaching and cultivate more outstanding international talents for the development of the IT industry.

Data Availability

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

Conflicts of Interest

The author declares that there are no conflicts of interest.

References

[1] F. Zhang, L. L. Dong, C. X. Wang, N. Ye, and Q. He, "Research on project driven it English teaching mode for software engineering majors," *Education Teaching Forum*, vol. 43, no. 5, pp. 192–194, 2019.

[2] F. Yu, "Research on the innovation of ESP English blended teaching under the background of informatization-taking it English as an example," *Campus English*, vol. 29, no. 1, pp. 36-37, 2019.

[3] M. Wu, H. H. Wang, J. Zhang, J. Liu, and T. Lu, "Research on blended learning teaching method of it English course software guide," *Educational Technology*, vol. 25, no. 10, pp. 64-65, 2017.

[4] L. Shao and X. Liu, "The application of ESP in college English teaching-taking it English as an example," *Science and Education*, vol. 56, no. 4, pp. 192-193, 2019.

[5] F. Zhang, "Task-based IT vocational English project oriented teaching model," *University Education*, vol. 18, no. 2, pp. 22–24, 2017.

[6] X. M. Gao, "Discussion on it English teaching method," *Shanxi Youth*, vol. 21, no. 1, pp. 81–87, 2019.

[7] X. Q. Zhang, T. G. Tang, H. Y. Wang, J. R. Zhou, and H. Yang, ""Internet +" research on the innovation and entrepreneurship education mode of IT majors," *The Science Education Article Collects*, vol. 15, no. 10, pp. 59–61, 2017.

[8] F. Jiang, "On the teaching strategies of IT professional English in higher vocational colleges based on vocational ability training," *Times Education*, vol. 25, no. 5, pp. 213-214, 2017.

[9] Y. Zhang, "The integration of IT professional English and information professional English education-a case study of

dalian neusoft information college," *Chinese Journal of Education*, vol. 48, no. S1, pp. 161–164, 2017.

[10] Z. Yuan, "Discussion on the curriculum reform of IT vocational English based on post demand-taking Wuhan vocational college of software engineering as an example," *Journal of Seeking Knowledge Guide*, vol. 30, no. 1, pp. 117-118, 2018.

[11] W. J. Cui, "Tourism English teaching mode under modern audio visual education technology," *Guangdong Canye*, vol. 52, no. 5, pp. 93-94, 2018.

[12] J. Zhu and G. J. Jia, "A study of college English teaching model based on cognitive linguistics," *Foreign Language World*, vol. 21, no. 3, pp. 30–37, 2018.

[13] L. Cheng, "Research on nursing professional English teaching mode based on content-based instruction idea," *Chinese Nursing Research*, vol. 32, no. 9, pp. 1449–1451, 2018.

[14] J. N. Shan, H. B. Zhu, B. Pang, Y. W. Li, and J. B. Luan, "Exploration on English teaching modes in TCM universities under humanistic education concept," *Western Journal of Traditional Chinese Medicine*, vol. 31, no. 10, pp. 40–42, 2018.

[15] S. N. Chu, J. X. Chen, and Y. H. Liu, "Research on the practice of science and trade English translation teaching under the "translation studio" mode," *Educational Research*, vol. 3, no. 11, pp. 108-109, 2020.

[16] Y. Dou, "Research on teaching mode of electrical engineering English," *Transformer*, vol. 57, no. 4, pp. 94–99, 2020.

[17] J. G. Liu, "A study on the teaching mode of college English under the background of modern science and technology development," *Food Research and Development*, vol. 41, no. 21, pp. 258-259, 2020.

[18] X. X. Cui, "Exploration of English teaching mode for electroplating specialty," *Electroplating & Pollution Control*, vol. 40, no. 2, pp. 114-115, 2020.

[19] J. Y. Zhang, "Exploration of English teaching mode for electric power specialty," *Electric Drive*, vol. 49, no. 11, pp. 123–126, 2019.

[20] R. S. Zhou, "English teaching model based on constructivism," *Education Review*, vol. 25, no. 8, pp. 147–149, 2018.

Hindawi
Scientific Programming
Volume 2021, Article ID 9961963, 8 pages
https://doi.org/10.1155/2021/9961963

Research Article

Bridge Extraction Algorithm Based on Deep Learning and High-Resolution Satellite Image

Wenbing Yang,[1] Xiaoqi Gao,[2] Chunlei Zhang,[2] Feng Tong,[3] Guantian Chen,[2] and Zhijian Xiao[4]

[1]*Yiwu Industrial and Commercial College, Yiwu, Zhejiang 322000, China*
[2]*Zenghe Packaging Co., Ltd., Wenzhou, Zhejiang 325000, China*
[3]*Wenzhou Heshun Packaging Machinery Co., Ltd., Wenzhou, Zhejiang 325000, China*
[4]*College of Digital Engineering, Zhejiang Dongfang Polytechnic, Wenzhou, Zhejiang, China*

Correspondence should be addressed to Zhijian Xiao; xiaozhijiandigital@126.com

Received 20 March 2021; Revised 23 April 2021; Accepted 7 May 2021; Published 2 June 2021

Academic Editor: Shah Nazir

This paper proposes a novel method of extracting roads and bridges from high-resolution remote sensing images based on deep learning. Edge detection is performed on the images in the road area along with the road skeleton line, and the result of the detected binary edge is vectorized. The interference of protective belts on both sides of the road, road vehicles, road green belts, traffic signs, etc. and the shadow interference of the bridge itself are eliminated to determine the parallel sides of the road. The bridge features on the road are used to locate the detected bridge and obtain information such as the location, length, width, and direction of the bridge, verifying the experimental results of the Shaoguan Le point images. In addition, in order to learn higher-level road feature information, the algorithm in this paper introduces the hollow convolution and multicore pooling modules. Secondly, the residual refinement network further refines the output of the prediction network to improve the ambiguity of the prediction network results. In addition, in view of the small proportion of road pixels in remote sensing images, the network also integrates binary cross entropy, structural similarity, and intersection ratio loss function to reduce road information loss. The applicability of the proposed study was tested, and the results show that the algorithm is very effective for the extraction of road and bridge targets.

1. Introduction

The bridge is one of the important artificial buildings. Its detection and identification are of great significance for GIS data acquisition, map update, and auxiliary supervision of bridge construction. Bridges are divided into water bridges and road bridges. At present, most of the existing algorithms are carried out for the target of bridges on the water. These algorithms mainly adopt a top-down knowledge-driven recognition method, that is, first use prior knowledge to establish a target recognition model. According to the hypothesis, segmentation, marking, and feature extraction are carried out on a purposeful basis, and then, further target detection is carried out [1–3]. These algorithms have certain reference significance for the extraction of road bridges on high-resolution images [4], but the surface conditions of the images of bridges on the road are complex and changeable. Buildings, protective belts on both sides of the road, vehicles on the road, road green belts, traffic signs, signs, billboards, and other factors will interfere with the extraction of bridges. At present, there is no complete algorithm for bridge extraction on the road [5–7].

In order to effectively identify the road bridge target in the remote sensing image, the key is how to extract the characteristics of the road bridge in the remote sensing image [8, 9]. It is concluded that the bridge on the road mainly has the following characteristics: (1) the bridge body is suspended in the air, except for the supporting piers and cables, and the bridge body has no contact with other objects on the ground. What is shown in the image is that the bridge body and the surrounding area have a clear boundary line. Even the overpass between the upper and lower highways

has a difference in brightness due to the influence of light. (2) There is basically no big change in the gray value inside the bridge, and it has local averageness, and the gray level of the whole bridge does not change much and it is the same. (3) The shape of the bridge imaged on the image is a long rectangle, with a pair of approximately parallel sides that intersect the road but are not necessarily perpendicular to the road. (4) The bridge has a certain width on the high-resolution image, usually a few pixels. (5) Across the road. Among them, (1) and (2) are the radiation characteristics of road bridges; (3) and (4) are the geometric characteristics of on-road bridges; and (5) is the functional characteristics of on-road bridges. We make full use of these features, formulate different criteria pertinently, and propose a reasonable method of extracting bridges on the road [10].

The organization of this paper is as follows: Section 2 briefly depicts the algorithm and realization of road bridge recognition. Section 3 shows the details of the deep-learning-based method. The experiments of the proposed study are presented in Section 4. The conclusion and future work are given in section 5 of this paper.

2. The Algorithm and Realization of Road Bridge Recognition

In this paper, aiming at the imaging characteristics and resolution of high-resolution visible light images [11] and the main characteristics of road bridge targets, a method for extracting road bridges based on multiple information is proposed. First, a certain size of sliding window is used to track the road skeleton line for Canny edge detection, and the edge information within the road domain is detected. Then, the detected edges are vectorized to obtain a set of vector lines, the criteria are used to remove invalid line segments, and finally, the bridge parallel lines are detected, the bridge target is identified, and the parameters of the bridge are determined. The overall framework of the algorithm is shown in Figure 1.

2.1. Edge Detection. The edge of the image is the most basic feature of the image. The most notable feature of the bridge is a pair of parallel lines. This article attempts to analyze and determine the outer edge line of the bridge body from edge detection [12, 13]. Edges are usually detected by using first-order and second-order derivatives to calculate the abruptness of the brightness value. Sobel, Roberts, LOG, and Canny are mainstream operators for calculating the abruptness of the brightness value.

Most of the bridges on the road have low contrast with the road, and the edges are relatively weak. Compared with Sobel, Roberts [14], and LOG, Canny edge detection uses high and low thresholds to extract edges and then connects the edges in the high-threshold image into contours. When the end point of the contour is reached when connecting, the edges that can be connected on the low-threshold image are found and collected until all the gaps are connected. As shown in Figure 2, the Canny operator can accurately locate the abrupt signal of the bridge edge line. Therefore, the Canny operator is selected as the edge detection operator in this paper.

In order to avoid the impact of buildings, vegetation, water system, etc. other than the road and because the road bridge is located within the road area, we can track the existing road vector data through the sliding window method and perform edge detection on the image within the road area. To ensure that bridges on the road can be fully detected, the length and width of the sliding window here are 1.5 times the width of the road. This not only further restricts the scope of bridge detection but also detects the edge of the bridge more quickly and accurately.

2.2. Edge Vectorization. In order to better analyze the geometric feature lines of the bridge, the binarized edge image can be vectorized to obtain a series of edge vector lines. For the removal of invalid line segments later, the straight line fitting of bridge lines, etc. provide edge point collections of edge line segments. The edge vectorization in this article is based on the ArcGIS Engine development environment, using the GP tool under development to vectorize the detected edge binarization image to obtain the corresponding vector file [15].

2.3. Invalid Line Segment Removal. The types of image features in the road area are more complicated. After edge extraction, a large number of invalid edge lines that interfere with the identification of bridges on the road will be obtained. Therefore, those invalid line segments that affect bridge detection must be removed before detecting bridge parallel lines. Specific interference factors include vehicles, trees, traffic signs, signs, billboards, railings on individual bridges, and shadows of the bridge itself.

2.3.1. Interference Removal Caused by Factors such as Vehicles, Trees, Traffic Signs, and Billboards. In the image after edge extraction, the vehicles and trees on the road mostly correspond to relatively small discontinuities and relatively short discontinuous edge lines, and the edges are mostly rough and the local curvature is larger. According to this feature, in order to avoid the impact of vehicles, road green belts, traffic signs, billboards, etc. on bridge recognition, this paper formulates guidelines to remove those very short or relatively large curvature edge lines. The specific algorithm guidelines are as follows:

Rule 1: all the edge lines are recorded as set N, each edge vector line of the edge vector file is traversed, and the curve length li of the edge vector line is calculated. Suppose the length threshold in the rule of removing line segments is lmin, then the set after removing the too short broken lines is $T = \{\mathrm{ti} | \mathrm{ti} \in N, \mathrm{li} > \mathrm{lmin}\}$.

Criterion 2: for the edge line in the set T, the distance between the end points of the line segment is di, the curve length of the line segment is li, then the curve degree parameter ci = di/li of the line segment is set, and cmin is the curvature threshold. Then, most of the

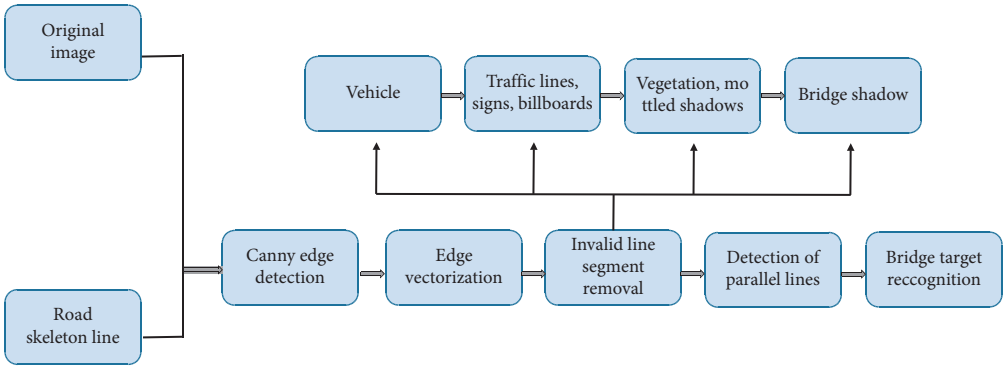


Figure 1: Overall framework of the bridge to identify.

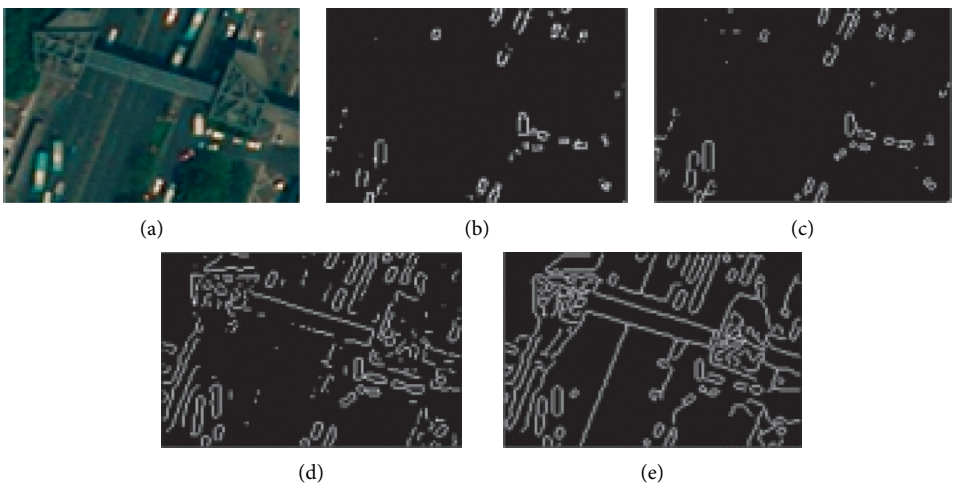


Figure 2: Result of edge detection. (a) Original image. (b) Roberts. (c) Sobel. (d) Log. (e) Canny.

invalid edge lines corresponding to the larger curvature of the vehicles on the road are removed, and the set $T' = \{ti|ti \in T,\ li > lmin,\ ci > cmin\}$ is obtained.

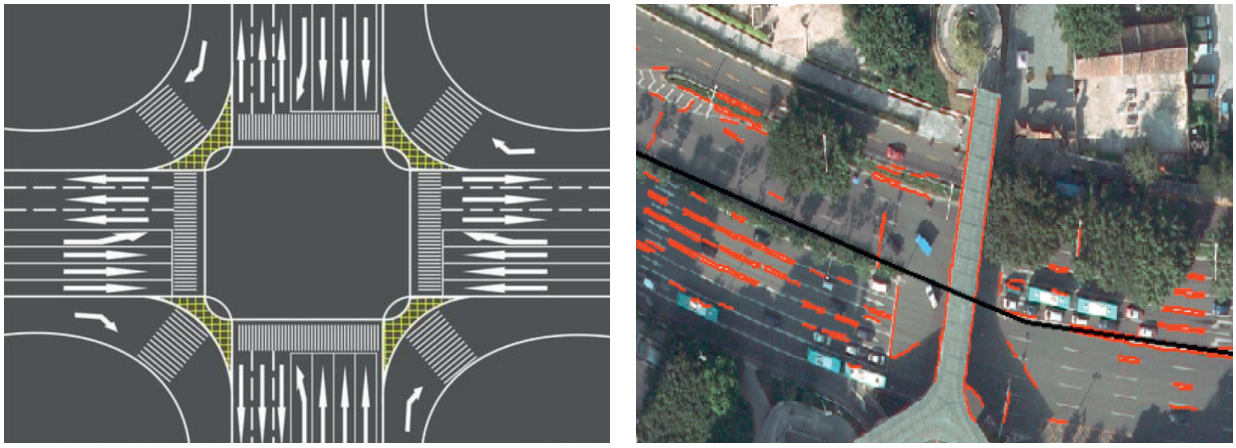
2.3.2. Removal of Interference Caused by Traffic Signs. The types of traffic signs on the road include guide lane lines, lane dividing lines, center double solid lines, sidewalk lines, and stop lines. Among them, the guide lane line, the lane dividing line, the center double solid line, and the sidewalk line are imaged on the high-resolution remote sensing image. After edge detection, the edge line formed is almost parallel to the road centerline. In the image, the bridge and the road center line intersect. Figure 3(a) shows a schematic diagram of a traffic sign line, and Figure 3(b) shows it on a remote sensing image with a resolution of 0.2 m. The spatial relationship between the traffic sign line and the road centerline is approximately parallel and does not intersect, and the bridge edge line is close to the road centerline. Therefore, in order to obtain more reliable results before identification, the article adopts criterion 3 to restrict.

Criterion 3: due to the imaging method of the central projection, the bridge and the road across it cannot be parallel. By setting the angle between the sideline of the bridge and the centerline of the road to be greater than a certain threshold α, after experiments, setting the threshold $\alpha = 30^{\circ}$ can remove most of the virtual scene information caused by the traffic sign line. The interference caused by the stop line needs to be further eliminated by the following parallel line detection criterion.

2.4. Interference Removal of the Shadow of the Bridge Itself. For shadows on remote sensing images, it is mainly caused by insufficient illumination, which leads to low brightness in shadow areas and lack of color information [7].

After the Canny edge detection, the shadows of the bridge and itself often form three approximately parallel edge lines. Based on the intersection of the line in the middle and the center line of the road, a certain range of data is taken on both sides to take the average. The side with the smaller mean value is judged to be the shadow part, and the corresponding edge line is regarded as the shadow edge line and deleted to remove the shadow interference of the bridge itself. The schematic diagram of shadow interference removal is shown in Figure 4.

2.4.1. Interference Removal Caused by Factors such as Bridge Railings. If there is a bridge on the road, there must be intersections between the center line of the road and the edge



(a) (b)

Figure 3: Traffic signs line. (a) Schematic diagram of traffic lines. (b) Spatial relationship among traffic sign lines, bridge edges, and road center lines on remote sensing images.

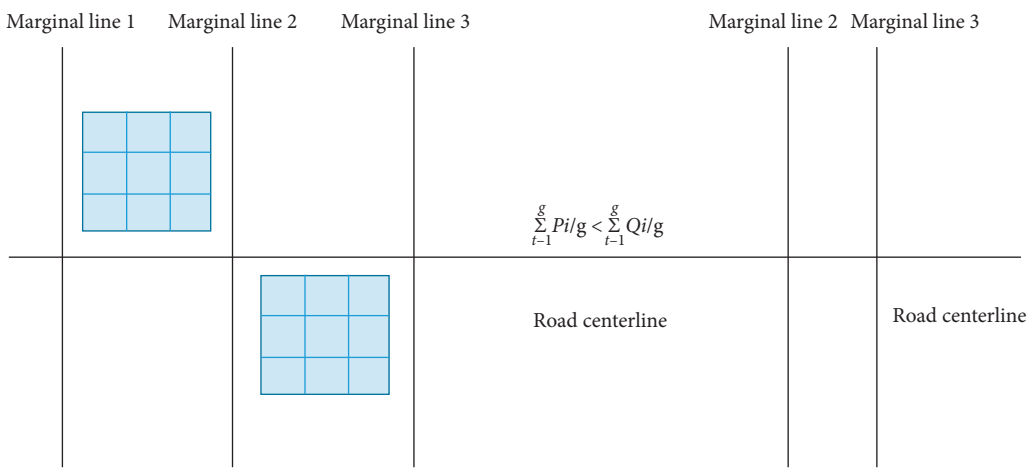


Figure 4: Schematic of bridge shadow interference removal.

of the bridge [3], and these intersections appear in pairs under normal circumstances. However, due to the angle between the normal of the sun and the angle of incidence, the image of the railing of the bridge may fall within the range of the bridge, and the intersection point with the center line of the road will also be recorded. Therefore, two or more points may be recorded at and near a certain edge of the bridge when detecting along the center line. Therefore, redundant points need to be deleted to avoid interference caused by factors such as bridge railings. The processing method is to traverse all the intersections, and if there is an intersection within the 8 neighborhood of another point, one of them is recorded [3].

2.5. Detect Parallel Edges. Through the deletion of invalid line segments, we obtain potential bridge parallel line pairs, but there may still be virtual scenes in these lines, so the obtained edge lines need to be further confirmed before identification. Because the two edges of the bridge are not ideal parallel lines and the bridge has a certain width, the virtual scene can be further eliminated by extracting approximately parallel edge line segments within a certain range [6], as shown in Figure 5.

The line segments AB and CD are the suspected edge line segment pairs of the bridge. A1 and B1 are the projections of AB on CD. C1 and D1 are the projections of CD on AB, respectively. d1–d4 are the distances from the end of the straight line to another straight line. The angle between the two straight lines is θ.

Criterion 4: (1) the bridge has a certain width, so d1–d4 is less than a specified threshold; (2) the intersection of the detected bridge edge and the road centerline is taken as the starting point, and the road width is used as the limit to track the edge line of the bridge, record the coordinates of the tracking point, and perform least-squares fitting on it to obtain the slope of the edge. Because the bridge has two

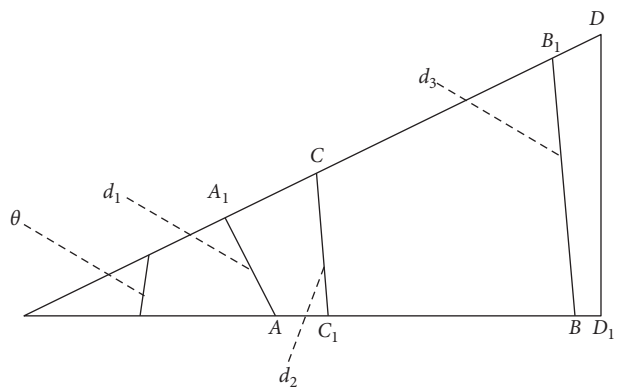


Figure 5: Criteria for detection of parallel lines.

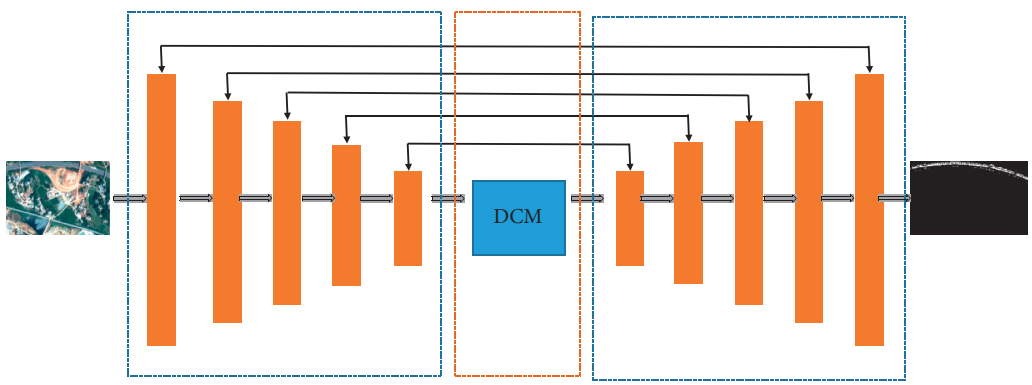


Figure 6: Encoder-decoder structure overall frame diagram.

sides, two slope values can be obtained. According to the angle formula, the angle θ corresponding to the two sides can be obtained. If θ is less than a certain threshold, it is considered as a pair of sides of the bridge.

2.6. Bridge Recognition. After the invalid line segments are removed, the edges of the bridge can be paired. Starting from the recorded first edge line, two adjacent lines can determine a bridge [16], and the intersection points with the center line of the road are $B1$ $(x1, y1)$ and $B2$ $(x2, y2)$.

(1) Determination of the location of the bridge:

$$B_c(x, y) = \left(\frac{(x_1 + x_2)}{2}, \frac{(y_1 + y_2)}{2}\right), \quad (1)$$

(2) Determination of the bridge position: the azimuth angle of the bridge sideline pair is averaged as the bridge position

(3) Determination of the width of the bridge: after obtaining the slope of the side line of the bridge, a straight line equation of two parallel lines is constructed according to the midpoint of the two side lines of the bridge, and the distance between the two lines is the width of the bridge

(4) Determination of the length of the bridge: the calculation principle of the length of the bridge is to take the position of the bridge as the starting point and grow on the image between the two bridge edges according to the consistency of the area

3. Deep-Learning-Based Method

This paper also uses deep-learning methods [16, 17] to assist road and bridge extraction tasks. In the image segmentation task, the more common structure is the encoder-decoder structure. The encoder is responsible for downsampling and reducing the image dimensions, and the decoder is responsible for upsampling and reducing the image dimensions. As shown in Figure 6, the method proposed in this paper chooses to load the ResNet-34 pretraining model. The depth of the entire network is designed as five layers, including one input convolutional layer and four encoder layers. When using the Massachusetts dataset for training, the image size of the input into the network is $500 \times 500 \times 3$; then, it will undergo a bilinear interpolation operation to interpolate the image size to $512 \times 512 \times 3$ in the input convolutional layer, and the feature map of the size of $512 \times 512 \times 3$ that will pass the step size is 2, the convolution kernel is a 7×7 convolution process, and the output feature map size at this time is $2\_56 \times 2\_56 \times 64$; then, through the maximum pooling process with a convolution kernel of 2×2 and a step size of 2, the feature map obtained the size $128 \times 128 \times 64$, and then, it is used as the input of the first-

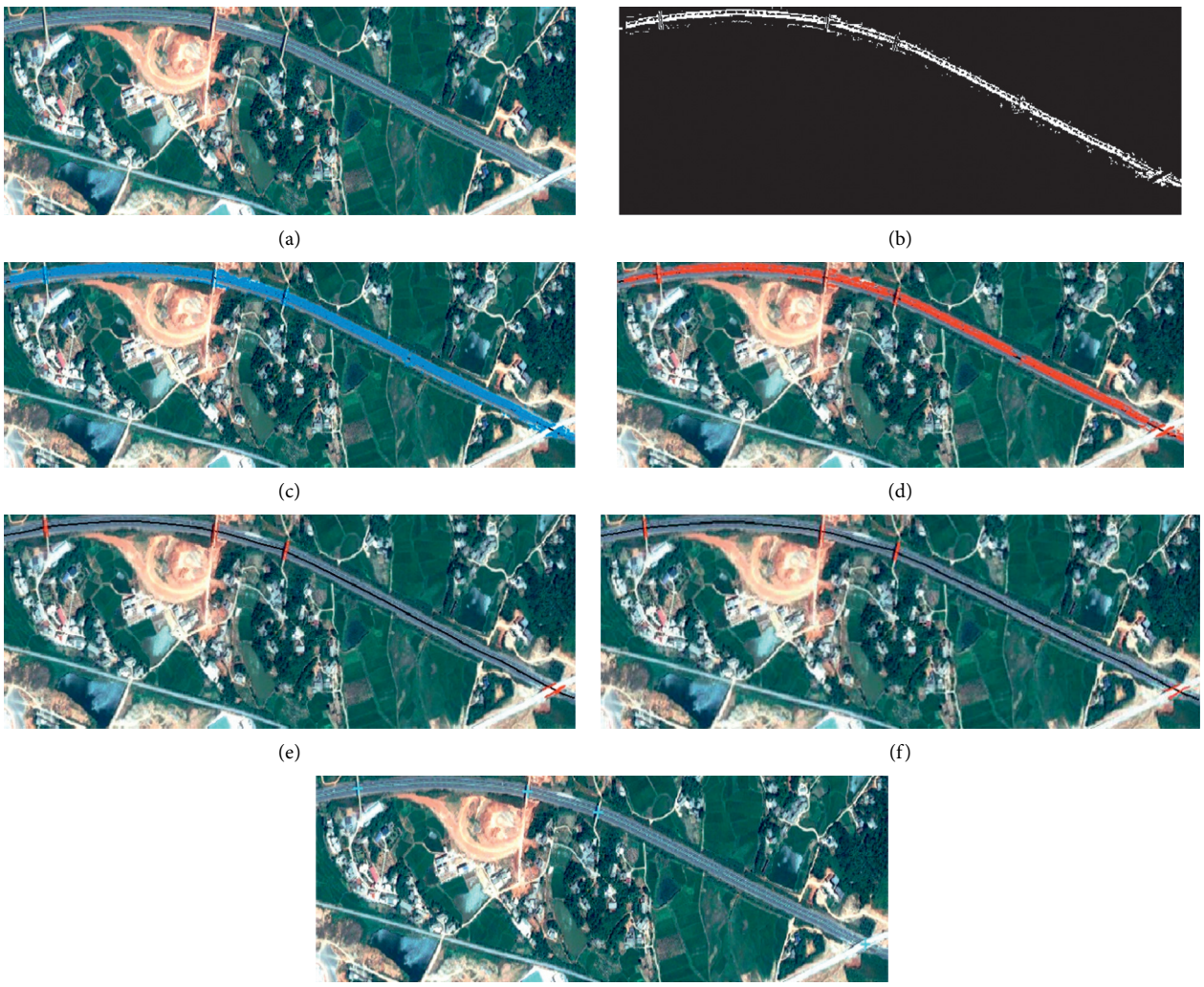


Figure 7: Original and extraction results of the experimental zone. (a) Original image. (b) Binary image after edge detection in the road area. (c) Edge vectorization results within the road area. (d) Result after removing interference from vehicles and road green belts. (e) The result of removing interference from traffic sign lines. (f) The result after removing the shadow. (g) Bridge extraction results.

layer decoder. In the abovementioned four-layer encoder layer, each layer of encoder is composed of a different number of residual blocks, and the residual block includes two 3×3 convolution kernels and a combined jump branch. The output generated by the four-layer encoder through the downsampling convolution operation is $128 \times 128 \times 64$, $64 \times 64 \times 128$, $32 \times 32 \times 256$, and $16 \times 16 \times 512$, and the output of the last encoder layer will be used as the input of the transition layer. Figure 6 represents the encoder-decoder structure overall frame diagram.

4. Experiment

The recognition method proposed in this paper is implemented in the VS2008 + ArcGIS Engine environment. Figure 7 shows the recognition results of part of the experimental area of Pleiades remote sensing image in Shaoguan.

Figure 7 shows part of the recognition results and intermediate results. Figure 7(a) is the original image; Figure 7(b) is the binarized image after the edge is extracted along the center line of the road using the Canny operator; Figure 7(c) is the edge raster image converted into vector lines by ArcGIS Engine, for the subsequent removal of invalid line segments and rapid topological calculation for detecting parallel lines; Figure 7(d) is the removal of interference factors such as vehicles and green belts, the threshold of curvature is set to 0.9, and the threshold of length is set to 4 pixels, and good results are obtained; Figure 7(e) is the removal of other interference factors such as traffic sign lines, and the threshold a of the angle with the road centerline is 30°; Figure 7(f) is the result of removing the shadow interference of the bridge itself, and further removing the virtual scene by detecting the parallel lines, the bridge linear edge pair is basically recognized [16]. According to that the pixel value of the bridge shadow is

Table 1: Bridge parameter identification results of the experimental zone (from left to right).

Serial number	Bridge center position (latitude and longitude)	Bridge length (m)	Bridge width (m)	Bridge orientation (°) (angle with true north)
1	(112.557886, 27.907211)	39.63964	4.951638	−7.569795
2	(112.562436, 27.907155)	39.63964	8.250913	9.733668
3	(112.564332, 27.906753)	39.63964	5.392012	17.764025
4	(112.571607, 27.904112)	39.63964	9.117343	69.819042

smaller than the pixel value of the bridge itself, the shadow effect is eliminated. When detecting parallel lines, the threshold of the angle θ between the two sides is set to 5°, and the distance threshold is set to 40 m; Figure 7(g) is the final recognition result, and the light blue cross represents the bridge target. Table 1 shows the position, orientation, length, width, and other parameters of the bridge detection results from left to right in Figure 7(g), and it can be seen that bridges on the road can be better identified. It can be seen from Table 1 that the position, orientation, length, and width of the bridge can be obtained relatively correctly using the algorithm of this article.

5. Conclusions and Further Work

This paper proposes a set of high-resolution remote sensing image road bridge extraction technologies based on multiple information. Based on the main characteristics of the bridge, the algorithm establishes a criterion method for bridge recognition, using the road area as the recognition area and detecting along the center line of the road. Limiting the detection range to the vicinity of the center line of the road is conducive to the improvement of detection efficiency and results; the vectorization of the edge after Canny edge detection is beneficial to the deletion calculation of the virtual scene line, and the bridge edge line is quickly and accurately determined. The identification of bridges on the road provides direct and rapid dynamic updates for the map and provides auxiliary reference for municipal planning.

The algorithm in this paper performs detection along the center line of the road, which depends on the accuracy of the road network information provided. If the road network is not accurately extracted, it will lead to misdetection and missed detection of bridge extraction; in addition, the detection algorithm in this paper is aimed at parallel bridges on the road, but it cannot be extracted correctly for ring-shaped and irregular bridges. This requires us to conduct further research.

Data Availability

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

Conflicts of Interest

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

References

[1] J. Han, K. Liang, B. Zhou, X. Zhu, J. Zhao, and L. Zhao, "Infrared small target detection utilizing the multiscale relative local contrast measure," *IEEE Geoscience and Remote Sensing Letters*, vol. 15, no. 4, pp. 612–616, 2018.

[2] J. Nie, S. Qu, Y. Wei, L. Zhang, and L. Deng, "An infrared small target detection method based on multiscale local homogeneity measure," *Infrared Physics & Technology*, vol. 90, pp. 186–194, 2018.

[3] Y. Shi, Y. Wei, H. Yao, D. Pan, and G. Xiao, "High-boost-based multiscale local contrast measure for infrared small target detection," *IEEE Geoscience and Remote Sensing Letters*, vol. 15, no. 1, pp. 33–37, 2017.

[4] H. Zhao, X. Qi, X. Shen, J. Shi, and J. Jia, "Icnet for real-time semantic segmentation on high-resolution images," in *Proceedings of the European Conference on Computer Vision (ECCV)*, pp. 405–420, Ithaca, NY, USA, August 2018.

[5] C. Tan, N. Uddin, E. J. OBrien, P. J. McGetrick, and C. W. Kim, "Extraction of bridge modal parameters using passing vehicle response," *Journal of Bridge Engineering*, vol. 24, no. 9, Article ID 04019087, 2019.

[6] T. Sun, Z. Di, and Y. Wang, "Combining satellite imagery and gps data for road extraction," in *Proceedings of the 2nd ACM SIGSPATIAL International Workshop on AI for Geographic Knowledge Discovery*, pp. 29–32, New York, NY, USA, November 2018.

[7] D. Wang, X. Meng, C. Gao, S. Pan, and Q. Chen, "Multipath extraction and mitigation for bridge deformation monitoring using a single-difference model," *Advances in Space Research*, vol. 60, no. 12, pp. 2882–2895, 2017.

[8] X. Zhang, Y. Yang, Z. Li, X. Ning, Y. Qin, and W. Cai, "An improved encoder-decoder network based on strip pool method applied to segmentation of farmland vacancy field," *Entropy*, vol. 23, no. 4, p. 435, 2021.

[9] W. Cai, B. Liu, Z. Wei, M. Li, and J. Kan, "TARDB-net: triple-attention guided residual dense and BiLSTM networks for hyperspectral image classification," *Multimedia Tools and Applications*, vol. 80, no. 7, pp. 11291–11312, 2021.

[10] Z. Shi and N. Uddin, "Extracting multiple bridge frequencies from test vehicle—a theoretical study," *Journal of Sound and Vibration*, vol. 490, Article ID 115735, 2021.

[11] A. Lichtenegger, D. J. Harper, M. Augustin et al., "Spectroscopic imaging with spectral domain visible light optical coherence microscopy in Alzheimer's disease brain samples," *Biomedical Optics Express*, vol. 8, no. 9, pp. 4007–4025, 2017.

[12] Y. Liu, M. M. Cheng, X. Hu, K. Wang, and X. Bai, "Richer convolutional features for edge detection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 41, no. 8, pp. 3000–3009, 2017.

[13] P. Ganesan and G. Sajiv, "A comprehensive study of edge detection for image processing applications," in *Proceedings of the 2017 International Conference on Innovations in*

Information, Embedded and Communication Systems (ICIIECS), pp. 1–6, IEEE, Coimbatore, India, March 2017.

[14] H. A. Sholeh, Y. Mulyani, and H. D. Saptama, "Studi perbandingan pengenalan karakter aksara lampung dengan metode deteksi tepi roberts dan sobel," *Inovasi Pembangunan: Jurnal Kelitbangan*, vol. 6, no. 3, pp. 261–272, 2018.

[15] B. Dahlberg and M. Versen, "Loop transformation algorithm for test vector accessing at high speed," in *Proceedings of the 45th International Symposium for Testing and Failure Analysis (ISTFA)*, p. 434, ASM International, Chardon, Ohio, December 2019.

[16] X. Ning, W. Li, and W. Liu, "A fast single image haze removal method based on human retina property," *IEICE Transactions on Information and Systems*, vol. E100.D, no. 1, pp. 211–214, 2017.

[17] X. Ning, X. Wang, S. Xu et al., "A review of research on co-training. concurrency and computation: practice and experience," *Concurrency and Computation Practice and Experience*, p. e6276, 2021, in Press.

Hindawi
Scientific Programming
Volume 2021, Article ID 1783384, 9 pages
https://doi.org/10.1155/2021/1783384

Research Article

A Cultural and Functional Approach to the Assessment of Logical Thinking Ability in English Writing

Cao Juan

Zunyi Medical and Pharmaceutical College, Zunyi 563006, Guizhou, China

Correspondence should be addressed to Cao Juan; 2004030328@st.btbu.edu.cn

Received 16 April 2021; Revised 27 April 2021; Accepted 28 April 2021; Published 1 June 2021

Academic Editor: Shah Nazir

The analysis of influencing factors of logical thinking ability in English writing is the key effective factor of evaluating logical thinking ability in English writing. In order to accurately evaluate logical thinking ability in English writing, this paper studies the evaluation method of logical thinking ability in English writing from the perspective of culture and function. This paper analyzes the relationship between the influencing factors of logical thinking ability in English writing. The factors of text structure and language expression reflect the culture and function of logical thinking ability in English writing, respectively, which have a direct impact on logical thinking ability in English writing. From these two aspects, 15 evaluation indexes are selected to construct the evaluation system of logical thinking ability in English writing. Considering the significant fuzziness of logical thinking ability in English writing from the perspective of culture and function, the comprehensive evaluation method of fuzzy mathematics is used for the process, the evaluation criteria are determined, the evaluation matrix is constructed, and the membership function is calculated, to complete the comprehensive evaluation of fuzzy mathematics based on the membership function and weight matrix. The experimental results show that this method can accurately evaluate the logical thinking ability of English writing and can be effectively used in the area of research.

1. Introduction

The word "thinking" in Chinese is synonymous with "deliberating" and "speculating." Thinking science believes that thinking is the process of people receiving information, storing information, processing information, and outputting information and generally reflects the objective reality [1]. From the essence of thinking, thinking is the conscious, indirect, and general reflection of the conscious human brain on the essential attributes and internal laws of objective reality [2]. Writing ability thinking, as far as physiological mechanism is concerned, is a writing thinking activity actively participated by trillions of cells under the action of human brain integrity; as far as the process is concerned, it is two kinds of transformation movement with opposite direction and internal unity of spiritualization of life (subjectivization of object) and lifestyle of spirit (objectivization of subject) [3]. It is a kind of directional creative thinking activity, which is the only way to sublimate understanding, dredge ideas, sort out materials, and design blueprints.

Logical thinking refers to the way of thinking that conforms to some man-made thinking rules and forms [4]. Logical thinking, also known as abstract thinking, is an advanced form of thinking. It is "a way of thinking in which people reflect reality by means of concept, judgment, and reasoning in the process of cognition. It is a thinking activity characterized by abstraction, leaving aside specific images and revealing the essential attributes of things". From the way of logical thinking, logical thinking can be divided into forward thinking, reverse thinking, divergent thinking, inductive thinking, and deductive thinking.

As a form of language output, English writing plays an important role in English learning, but it is also influenced by logical thinking to a great extent [5]. As we all know, English writing ability is one of the most difficult abilities in English learning. English writing, that is, the written expression of English, is to express certain thoughts and

feelings through certain ideas and symbols, so as to be accepted by others. Both conception and expression have a process of logical thinking. Words are the smallest unit in language use, and concepts are called thinking cells. Notional words and phrases generally express concepts. Because words produce sentences, declarative sentences and rhetorical sentences can express statements, and product sentences can become sentence groups or logical segments. In fact, this is the process of logical proposition composed of concepts and reasoning formed by propositions [6]. Just as the famous Romanian logician Olli Wald said, logic is the common grammar of all languages and the product of language's abstract ability and generalization ability.

Chinese students are usually used to inductive thinking, which is different from deductive thinking in western society. It can be said that the logical way of thinking in English writing can, to a certain extent, reflect the writing ability, writing effect, and intelligence level of English majors [7]. In the teaching of English writing, people often only pay attention to language training but ignore the more important and fundamental training of logical thinking. As a result, there are some logical problems in writing, such as language barrier, unclear level, and unclear theme. Good English works have rigorous thinking, clear thinking, and clear context. The author must strictly abide by the basic laws of logical thinking: the law of identity, the law of contradiction, the law of excluded middle, and the law of sufficient reason. These laws require writing thinking to be definite, distinct, and argumentative from different angles.

According to the survey results of students' English writing [8], only 20% of the students who participated in the interview liked English writing, and the other 80% did not like it to some extent. The reasons they do not like include teaching methods, contents, and the exams they have to take in, but the main reason is the contents. They think that they have learned a lot, but they have "nothing to say" every time they write. Their "speechless" is that their logical way of thinking has not been well developed and guided. Their logical way of thinking has not been well guided and developed, they have not formed an effective logical way of thinking and injected appropriate ideological connotation into this way, and it is difficult to sublimate their own understanding, dredge the thinking of writing, sort out the writing materials, and design the blueprint of the article.

Lu's linear structure causality model in "Research on the influencing factors of college students' English writing ability" [9] shows that Chinese writing ability, oral English ability, and English vocabulary level have a direct impact on English writing ability, while English comprehension knowledge and ability have no direct impact on English writing ability. Marc and Alessandra put forward the "Analysis of the causes and countermeasures of college English writing function words failure" to analyze the causes of students' high pragmatic failure in the process of writing [10]. This paper argues that the main reasons are the wrong way of vocabulary acquisition, the lack of positive vocabulary, and the lack of cultural background knowledge. Arifin and Alkadri analyze "The current situation and influencing factors of college students' weariness of learning practical writing" [11] and find that, in the actual teaching, college students are generally tired of learning writing, such as fear of difficulty, burnout, anxiety, and weak interest. This paper argues that this phenomenon seriously affects the teaching effect of writing course and the quality of talent training in colleges and universities. They think that students, schools, teachers, and other factors are the main causes of this kind of weariness. The above researches in literature fully show that the analysis of influencing factors of logical thinking ability in English writing is the main influencing factor of evaluating logical thinking ability in English writing. Based on this study, the evaluation method of logical thinking ability in English writing from the perspective of culture and function is analyzed from the perspective of culture and function, the evaluation index is selected, and the comprehensive evaluation method based on fuzzy mathematics is used to obtain effective evaluation results. The following are the contributions of the study conducted:

(i) The evaluation method of logical thinking ability in English writing is studied based on culture and function perspectives.

(ii) The relationship between the influencing factors of logical thinking ability in English writing is analyzed.

(iii) 15 evaluation indexes are selected to construct the evaluation system of logical thinking ability in English writing.

(iv) Application of fuzzy mathematics is used for the process, the evaluation criteria are determined, the evaluation matrix is constructed, and the membership function is calculated, to complete the comprehensive evaluation of fuzzy mathematics based on the membership function and weight matrix.

(v) To show the effectiveness of the proposed study, experimental work has been done.

The paper is organized as follows: Section 2 is the representation of the materials and methods followed by details in subheadings. Section 3 briefly discusses the results of the study done. How to improve the thinking ability of English writing is given in Section 4 with details in subheadings. The conclusion of the paper is given in Section 5.

2. Materials and Methods

2.1. The Relationship between the Influencing Factors of Logical Thinking Ability in English Writing. According to the results of relevant surveys and interviews [12], it can be found that, among many factors influencing logical thinking ability in English writing, textual structure, language expression, mode of thinking, writing psychology, and writing environment are indeed the main factors influencing English writing ability. The relationship between these main factors is shown in Figure 1.

As the previous analysis shows, among the five main factors, the mode of thinking, writing psychology, and writing environment is the "software" factors; the textual

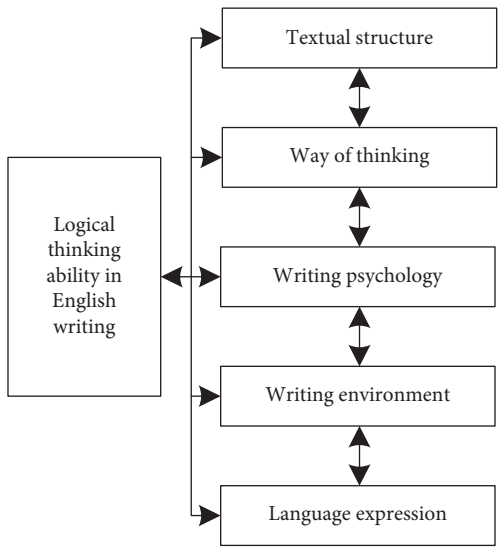


Figure 1: Influencing factors of logical thinking ability in English writing.

structure and language expression are the "hardware" factors. These factors are interrelated and inseparable [13]. Among them, "hardware" factors, textual structure factors, and language expression factors, respectively, reflect the culture and function of logical thinking ability in English writing, which have a direct impact on logical thinking ability in English writing. Based on these two aspects, this paper constructs an evaluation system of logical thinking ability in English writing. Considering the significant fuzziness of logical thinking ability in English writing from the perspective of culture and function, the comprehensive evaluation method of fuzzy mathematics is selected to evaluate the ability.

2.2. Construction of Index System

2.2.1. Selection Principle of Evaluation Index. The construction of the evaluation system of logical thinking ability in English writing should be based on promoting students' understanding of English, English writing ability, and English writing evaluation criteria [14]. In the process of evaluation, the cultural and functional evaluation of the writing process and writing results should be effectively balanced, the evaluation of the writing process should be highlighted, and the "for learning" and "as learning" of evaluation function should be given full play [15], highlighting the central position of students in the evaluation and their subjective initiative, stimulating students' self-esteem and self-confidence, enhancing their interest in learning English writing, collecting curriculum and teaching information in multilevel, multidimensional, and multichannel, scientifically analyzing, explaining, and giving feedback on evaluation results, improving students' logical thinking ability and autonomous learning ability in English writing, and effectively promoting the cultivation of students' practical ability of English academic discourse.

2.2.2. Construction of Index System. Based on the above principles, from the perspective of culture and function, this paper selects 15 evaluation indexes to construct the evaluation system of logical thinking ability in English writing, such as English vocabulary knowledge and English grammar knowledge, obeying the law of excluded middle and identity, as shown in Table 1.

Specifically speaking, from the perspective of culture, the influence of logical thinking on English writing includes four aspects: logical thinking and English writing word selection, logical thinking and English sentence making, logical thinking and paragraph construction, and logical thinking and discourse planning; from the perspective of function, the ability of logical thinking in English writing should abide by the law of excluded middle, identity, contradiction, and sufficient reason.

2.3. Comprehensive Evaluation Method Based on Fuzzy Mathematics.

A comprehensive evaluation method based on fuzzy mathematics is used to evaluate the logical thinking ability of English writing from the perspective of culture and function. The specific analysis process is as follows.

2.3.1. Construction of Fuzzy Factor Set. The construction of a comprehensive evaluation factor set based on fuzzy mathematics describes the selection of representative factors that affect the logical thinking ability of English writing (i.e., the evaluation indexes in Table 1) as evaluation factors. $P = \{p_1, p_2, \ldots, p_i, \ldots, p_n\}\,(i = 1, 2, 3, \ldots, n)$ is used as the evaluation factor set based on fuzzy comprehensive of logical thinking ability of English writing and p_i as the i-th evaluation factor of logical thinking ability of English writing.

2.3.2. Determination of Evaluation Criteria. B is the set of logical thinking ability in English writing. The formula is $B = \{b_1, b_2, \ldots, b_j, \ldots, b_m\}\,(j = 1, 2, 3, \ldots, m)$, where b_j is the j-th evaluation standard. According to the standard of logical thinking ability in English writing, the j value is 5.

2.3.3. Construction of Evaluation Matrix. As the basis of constructing a comprehensive evaluation matrix of fuzzy data, membership function construction should be based on "triangular fuzzy number," "descending half trapezoidal fuzzy number," and "ascending half trapezoidal fuzzy number" [16] and refer to the ability division norms in the relevant standards of logical thinking ability in English writing to construct the membership function of logical thinking ability index factors in English writing. Based on the relevant standards of logical thinking ability in English writing [17], we can get that, in the relevant standards, the standard value of level 4 and level 5 is a critical value, which is consistent with the standard value of Z_{i4} and Z_{i5}. The critical value is essentially a distinguishing value, lacking the transition between level 4 and level 5, which makes the membership function have the problem of "discontinuity" or "repetition." Some of them are actually the assessment results of logical thinking ability in English writing. The

Table 1: Evaluation system of logical thinking ability in English writing.

Target layer	First level index layer	Secondary index layer	Index description
Assessment of logical thinking ability in English writing	From the perspective of culture	Understanding vocabulary knowledge in English	CET-4 basic vocabulary and CET-4 vocabulary
		Understanding grammar in English	Test scores of tense, voice, congruence, structure, and nonfinite verb
		Understanding discourse knowledge in English	Test scores of grammatical cohesion, lexical cohesion, and discourse reconstruction
		English reading comprehension ability	CET4 reading comprehension test scores
		English listening comprehension	CET4 listening comprehension test scores
		English vocabulary reuse ability	CET-4 basic vocabulary and CET-4 vocabulary
		English grammar ability	Test scores of tense, voice, congruence, structure, and nonfinite verb
		The ability of English discourse reuse	Test scores of grammatical cohesion, lexical cohesion, and discourse reconstruction
		Oral English ability	Reading, oral test scores, and teacher impression scores
		English writing ability	Composition score, number of words, and superstructure score
		Chinese writing ability	Composition score, number of words, and superstructure score
	From the perspective of function	Abide by the law of exclusion	The theme of the paper is distinct and the material selection is typical
		Abide by the same principles	The material and theme of the article are coordinated and unified
		Abide by the law of contradiction	The level of the article is clear and the discussion is consistent
		Follow the law of good reason	Convincing people by reason, conclusive evidence, full exposition, and satisfactory performance of the theme

evaluation of logical thinking ability in English writing with the fifth level ability standard is obviously unable to construct the membership function (including $Z_{i5} - Z_{i4} = 0$) of the fifth level ability, which is prone to the problem of deviation in the analysis results. Considering that the membership function of the fourth level ability is symmetrical to $x_i = Z_{i4}$, based on the triangular fuzzy number optimization, the constraint condition $Z_{i5} - Z_{i4} = Z_{i4} - Z_{i3}$ is set, thus to construct the membership function of the fourth level ability and the fifth level ability of logical thinking in English writing.

The optimized membership function of logical thinking ability evaluation in English writing can ensure the continuity of the membership relationship between the first level ability and the fifth level ability in the process of logical thinking ability evaluation in English writing [18], which can better meet the evaluation standard of logical thinking ability in English writing. The membership function formula of the optimized logical thinking ability in English writing is described as follows.

Under the condition that p_i belongs to the first level ability, the membership function formula is described as follows:

$$r_{i1} = \begin{cases} 1, & x_i \le Z_{i1}, \\ \dfrac{Z_{i2} - x_i}{Z_{i2} - Z_{i1}}, & Z_{i1} < x_i < Z_{i2}, \\ 0, & x_i > Z_{i2}. \end{cases} \tag{1}$$

Under the condition that p_i belongs to the second level ability, the membership function formula is described as follows:

$$r_{i2} = \begin{cases} 0, & x_i \le Z_{i1} \cup x_i > Z_{i3}, \\ \dfrac{x_i - Z_{i1}}{Z_{i2} - Z_{i1}}, & Z_{i1} < x_i \le Z_{i2}, \\ \dfrac{Z_{i3} - x_i}{Z_{i3} - Z_{i2}}, & Z_{i2} < x_i \le Z_{i3}. \end{cases} \tag{2}$$

Under the condition that p_i belongs to the third level ability, the membership function formula is described as follows:

$$r_{i3} = \begin{cases} 0, & x_i \le Z_{i2} \cup x_i > Z_{i4}, \\ \dfrac{x_i - Z_{i2}}{Z_{i3} - Z_{i2}}, & Z_{i2} < x_i \le Z_{i3}, \\ \dfrac{Z_{i4} - x_i}{Z_{i4} - Z_{i3}}, & Z_{i3} < x_i \le Z_{i4}. \end{cases} \tag{3}$$

Under the condition that p_i belongs to the fourth level ability, the membership function formula is described as follows:

$$r_{i4} = \begin{cases} 0, & x_i \le Z_{i3} \cup x_i > Z_{i5}, \\ \dfrac{x_i - Z_{i3}}{Z_{i4} - Z_{i3}}, & Z_{i3} < x_i \le Z_{i4}, \\ \dfrac{Z_{i5} - x_i}{Z_{i5} - Z_{i4}}, & Z_{i4} < x_i \le Z_{i5}. \end{cases} \tag{4}$$

Under the condition that p_i belongs to the fifth level ability, the membership function formula is described as follows:

$$r_{i5} = \begin{cases} 0, & x_i \le Z_{i4}, \\ \dfrac{Z_{i5} - x_i}{Z_{i5} - Z_{i4}}, & Z_{i4} < x_i \le Z_{i5}, \\ 1, & x_i > Z_{i5}. \end{cases} \tag{5}$$

In formulas (1) to (5), Z_i, x_i, and r_i denote the limits of all levels of ability standards, the actual values of logical thinking ability in English writing, and the corresponding membership degrees of different ability levels.

According to the fuzzy data, the membership degree of the evaluation factor u_i to the evaluation set B is determined. The membership degree can be expressed by $R_i = (r_{1i}, r_{2i}, r_{3i}, r_{4i}, r_{5i})$, and the fuzzy relationship from P to B can be expressed by $R = \{R_{i1}, R_{i2}, \ldots, R_{ij}\}^T$ $(j = 1, 2, 3, \ldots, n)$. By describing R_i in the form of matrix, the following results are obtained:

$$R = \begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1j} \\ r_{21} & r_{22} & \cdots & r_{2j} \\ \vdots & \vdots & \vdots & \vdots \\ r_{i1} & r_{i2} & \cdots & r_{ij} \end{bmatrix}. \tag{6}$$

2.3.4. Building Weight Matrix. In the process of evaluating the ability of logical thinking in English writing, the importance (role, contribution, impact, etc.) of different factors $p_i\,(i = 1, 2, 3, \ldots, n)$ in factor set P is different, so it is necessary to empower the factors involved in the evaluation of the indexes [19]. There are differences in the weight values given and the level of influence on logical thinking ability in English writing. In the practical application process, the objective weighting method is often used, so the intensity contribution rate method is selected for weighting. The weighting method is a dynamic/variable weighting method, which can not only highlight the influence of the main intensity indexes but also take into account the overall influence of a single intensity indicator on the logical thinking ability of English writing. Formulas (7)–(9) are the detailed weight calculation formulas.

$$w_i = \frac{x_i}{1/5 \sum_{j=1}^{5} Z_{ij}}, \quad (i = 1, 2, \ldots, n;\ j = 1, 2, \ldots 5), \tag{7}$$

$$\overline{w_i} = \frac{w_i}{\sum_{i=1}^{n} w_i}, \quad (i = 1, 2, \ldots, n), \tag{8}$$

$$W = \{\overline{w_1}, \overline{w_2}, \ldots, \overline{w_n}\}, \sum_{i=1}^{n} \overline{w_i} = 1. \tag{9}$$

In the above formula, Z_{ij} and w_i, respectively, represent the limit values of each level of competence standard in the standard of logical thinking ability in English writing and the weight value of the normalized processing of the logical thinking ability indexes in English writing. Because w_i can be defined as the important membership of the corresponding factors, the weight matrix W can be defined as a fuzzy subset of the evaluation factor set U.

2.3.5. Comprehensive Evaluation of Fuzzy Mathematics. After determining P, B, and R, the rules of operation are constructed by using the weight matrix W and the membership relation matrix R, and the membership matrix G of the fuzzy comprehensive relationship between the objects of the logical thinking ability in English writing and the fuzzy comprehensive relationship of each evaluation level can be constructed.

$$G = W \odot R = \left(a_1, a_2, \ldots, a_j\right). \tag{10}$$

In formula (10), $\odot$ and a represent the degree values of fuzzy operation rules and the influencing factors of logical thinking ability in English writing belonging to different evaluation grades. The commonly used fuzzy operation rules include main factor determining type, weighted average type (multiplication and addition), and other operation rules [20].

3. Results

In order to verify the application effect of the proposed assessment method of logical thinking ability in English writing from the perspective of culture and function, 50 sophomores in an English Department of a university are randomly selected as the research objects, and the logical thinking ability in English writing of the selected subjects is evaluated by using the method of this paper. The evaluation results are as follows.

3.1. Construction of Fuzzy Factor Set. Taking the evaluation indexes in Table 1 as the main influencing factors of the

evaluation of logical thinking ability in English writing, the fuzzy factor set is constructed.

3.2. Determination of Evaluation Criteria. The logical thinking ability of English writing is divided into five levels, as shown in Table 2.

Table 2: Results of logical thinking in English writing.

Gradation	Classification criteria
Level 1 capability	The composite index is above 80
Level 2 capability	The composite index is between 60 and 80
Level 3 capability	The composite index is between 40 and 60
Level 4 capability	The composite index is between 20 and 40
Level 5 capability	The composite index is below 20

3.3. Weight Calculation. The difference in logical thinking ability in English writing and the fuzzy relationship matrix between fuzzy factor sets are the two important assessment indicators. Based and based on this root, the weight value of each evaluation criterion of logical thinking ability in English writing is calculated. The results are shown in Table 3.

The analysis of Table 3 shows that, in the evaluation index system constructed for logical thinking ability in English writing, English writing ability is the most important factor affecting logical thinking ability in English writing; English understanding of vocabulary knowledge, English understanding of grammar knowledge, Chinese writing ability, abiding by the law of excluded middle, abiding by the law of sameness, abiding by the law of contradiction, and abiding by the law of sufficient reason are of great significance to the logical thinking ability in English writing. The weight of English vocabulary reuse ability is the lowest among the evaluation indexes; the weight of the remaining evaluation indexes is more significant.

Table 3: Calculation results of weight values of different influencing factors.

Evaluating indicator	Weight
Understanding vocabulary knowledge in English	0.072
Understanding grammar in English	0.093
Understanding discourse knowledge in English	0.049
English reading comprehension ability	0.047
English listening comprehension	0.041
English vocabulary reuse ability	0.039
English grammar ability	0.040
The ability of English discourse reuse	0.041
Oral English ability	0.069
English writing ability	0.132
Chinese writing ability	0.083
Abide by the law of exclusion	0.075
Abide by the same principles	0.072
Abide by the law of contradiction	0.071
Follow the law of good reason	0.076

3.4. Evaluation Results. After the evaluation of this method, we can get the comprehensive evaluation results of the logical thinking ability of the selected subjects in English writing. The comprehensive evaluation results of 10 subjects are compared with the evaluation results based on the Nemerow index and principal component analysis. The results are shown in Figure 2 and Table 4.

Figure 2 can reflect the change trend of the evaluation results of the selected study area. Because the evaluation is a quantitative method based on rank analysis, in order to describe the correlation degree between each evaluation result, rank correlation coefficient (sequential correlation coefficient) is introduced to analyze the correlation of evaluation results. The definition formula of the rank correlation coefficient between the two elements is as follows:

$$r_{xy}{}' = 1 - 6\sum_{i=1}^{n}\frac{d_i^2}{n\left(n^2-1\right)d_i^2} = \left(R_{1i} - R_{2i}\right)^2. \tag{11}$$

In formula (11), $r_{xy}{}'$ is the rank correlation coefficient of the two indexes and $d_i^2 = (R_{1i} - R_{2i})^2$ is the square of the difference between the two indexes in the same group of samples.

According to the results of rank correlation analysis in Table 4, it is found that the rank correlation coefficients of the proposed method with fuzzy operation rule determined by main factors, the proposed method with fuzzy operation rule determined by weighted average, and the evaluation method based on Nemerow index and the evaluation method based on principal component analysis are 0.719, 0.690, and 0.550, respectively (when the confidence level (both sides) is 0.01, there is a very significant correlation). This shows that the above four evaluation methods can be used to evaluate the logical thinking ability of English writing.

Table 5 shows the statistical table of logical thinking evaluation results of 50 subjects' English writing ability.

It can be seen from Table 5 that the number of research objects in level 4 and level 5 is consistent according to the evaluation results of the proposed method under two different fuzzy operation rules, while the number of research objects in level 1 ability evaluation is gradually increased as the following order: the fuzzy operation rule of the proposed method determined by the main factor ⟶ the fuzzy operation rule of the method determined by the weighted average ⟶ the evaluation method based on principal component analysis ⟶ the evaluation methods based on Nemerow index; the number of research objects in level 2 ability evaluation is gradually decreased as the following order: the fuzzy operation rule of the proposed method determined by the main factor ⟶ the fuzzy operation rule of the method determined by the weighted average ⟶ the evaluation methods based on principal component analysis ⟶ the evaluation methods based on Nemerow index; the number of research objects in level 3 ability evaluation is gradually decreased as the following order: the fuzzy operation rule of the proposed method determined by the main factor ⟶ the fuzzy operation rule of the method determined by the weighted average ⟶ the evaluation method based on Nemerow index ⟶ the evaluation method based on principal component analysis. The results show that the

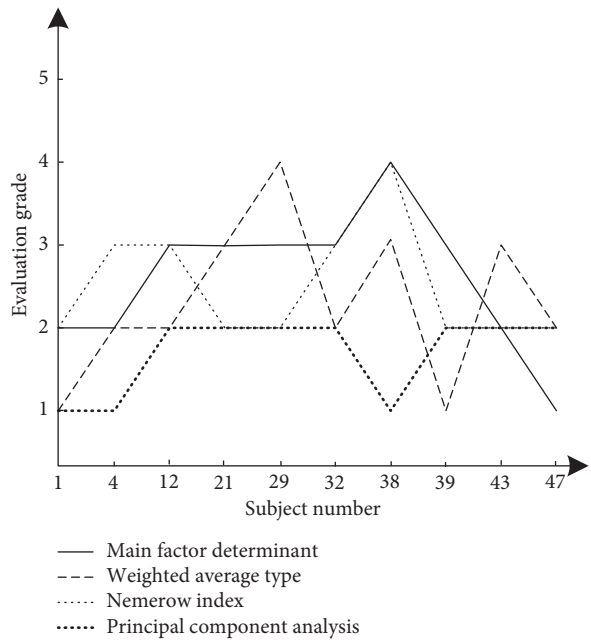


Figure 2: Membership chart of different research objects.

Table 4: Rank correlation coefficient of evaluation type ($n = 50$).

Evaluation method	The method of this paper (main factor determining type)	The method of this paper (weighted average type)	Evaluation method based on Nemerow index	Evaluation method based on principal component analysis
The method of this paper (main factor determining type)	1.000			
The method of this paper (weighted average type)	0.719**	1.000		
Evaluation method based on Nemerow index	0.657**	0.690**	1.000	
Evaluation method based on principal component analysis	0.625**	0.591**	0.550**	1.000

Table 5: Statistical table of evaluation results ($n = 50$).

Evaluation grade	The method of this paper (main factor determining type)	The method of this paper (weighted average type)	Evaluation method based on Nemerow index	Evaluation method based on principal component analysis
Level 1 capability	6	9	20	10
Level 2 capability	28	24	15	19
Level 3 capability	9	8	6	5
Level 4 capability	6	6	7	11
Level 5 capability	1	1	2	5

evaluation accuracy of the method based on two different fuzzy operation rules and the evaluation method based on the Nemerow index is better than that based on principal component analysis. The evaluation results of the proposed method under two different fuzzy operation rules are different from those of the Nemerow index evaluation method due to the improved membership function. The reason is that the Nemerow index evaluation method has

discontinuous critical values, which cannot represent the real logical thinking ability of English writing. On the whole, the evaluation results obtained by the proposed method are better than those obtained by the two comparative methods.

It can be seen from Table 4 that the rank correlation coefficient of the proposed method with weighted average as fuzzy operation rule reaches 0.719, which indicates that this method is superior to other methods in the evaluation of logical thinking ability in English writing. However, due to the existence of overlapping budget in this method with the main factor as fuzzy operation rule, it is easy to cover up the influence of English writing ability on logical thinking ability in English writing. When some indexes are especially large, it is easy to cause unreasonable comprehensive evaluation results, while the weighted average type can retain all or most of the information in the index vector as far as possible, which solves the problem of the failure of the maximum clustering coefficient principle to a certain extent, and has the characteristics of positive laterality, single increment, continuity, and additivity, comprehensively reflecting the effect of a single index on the overall evaluation. The effect of fruit can reasonably reflect the comprehensive influence of all indexes.

4. How to Improve the Thinking Ability of English Writing?

4.1. Typed Writing Training. In order to cultivate students' good English writing habits, the first step is to improve students' ability of writing layout and strengthen their understanding of the differences between English and Chinese texts. There are two kinds of text exercises: one is outline writing, in which students write a detailed outline of the given topic; the other is sorting exercise, in which students choose a passage written by British and American people, delete the conjunctions, disrupt the sentence order, and ask students to reorder and add conjunctions to indicate logical relations. This exercise helps students to understand the logical habits of English writing. Contrast exercise is a combination of sentence construction and text training. Firstly, excellent English model articles and extract Chinese outline are selected. On this basis, students write independently. Then, students are asked to compare the model articles and find out the logical defects of their compositions.

4.2. Specific Writing Topics. The number of words required in College English writing practice is generally about 200 words, which determines that the content is extremely limited and the topic is too wide, which will directly affect the quality of writing. Therefore, teachers should select topics to make them more suitable for students' writing practice. Taking the writing exercise "My Thoughts on Environmental Protection" as an example, there are too many contents that can be written, and the scope should be narrowed. It can be modified to "My Thoughts on Environmental Protection on Campus/of Yang zi River/in the Age of Internet" and so on. For example, the composition of Unit 3 in Volume 1 is "How Science Changes Our Lives," and we may as well refine the writing requirements and let students choose a specific scientific invention and explore its impact on life more deeply. It can also effectively avoid the common problems of students' composition, such as scattered center and empty content.

4.3. Modular Evaluation Method. The main content of the College English writing course consists of six steps: making clear the goal of practice, determining the proposition and requirement, assigning writing tasks, teaching writing strategies and skills, guiding students' practice, and reviewing compositions and teachers' feedback. The last two steps are easy to ignore. The feedback that students get often only contains the overall impression scores given by teachers and the grammar problems that they randomly tick out. They cannot fundamentally understand the problems in their compositions, and the practice effect cannot be guaranteed. In recent years, great progress has been made in the computerization and networking of composition evaluation methods in foreign countries. PEG, e-rater, IEA, and other automatic scoring systems have been well applied. However, there is no mature, for English composition evaluation in Chinese context. Clear scoring rules and advanced means of evaluation can smooth the communication between teachers and students and improve the effect of writing teaching. In terms of scoring rules, it is suggested to adopt the six-point method to classify and consider logic, content, sentence construction, grammar, vocabulary, and operation details, score items, and give specific feedback. Office software and e-mail system provide favorable tools for modularization of writing assessment, which can make the steps of marking clearer and more fluent. Students use word documents to write compositions. The error correction function of the software helps to correct simple spelling and punctuation errors. The self-rating reflects the quality of operation details. The preliminary revised articles are sent to the students in the study group by e-mail. The second step is to evaluate the use of basic grammar and vocabulary, encourage the students in the same group to evaluate each other after group discussion, score each item, and make corrections with comments. After the completion of the mutual evaluation, it is sent to the teacher for review above the language level, giving specific modification opinions and scores in logic, content and sentence construction, and then returning to the evaluation group and the author; the fourth step of the evaluation is the students' rewriting and second submission; the fifth step is to review the final draft, evaluate whether the modification is appropriate, score and establish the learner file from six angles again, record the scores of two rounds of each exercise in detail, save the copy of the composition, and master the learning effect of the students; finally, give feedback on the comments, final draft, and model to the students.

5. Conclusion

English writing requires the expression of clear and complete ideas. The purpose of writing teaching is to improve students' ability to express ideas in English. And "thought

poverty" and "critical crisis" have become the true portrayal of English majors' writing. Under this background, this paper studies the evaluation method of logical thinking ability in English writing from the perspective of culture and function, selects the evaluation index from the perspective of culture and function, constructs the evaluation system, and obtains the effective evaluation results by using the comprehensive evaluation method of fuzzy mathematics. It is expected that this method can effectively improve students' English writing ability.

Data Availability

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Conflicts of Interest

The author declares that he has no conflicts of interest.

References

[1] B. María-José, G. D. Héctor, and M. Poblete, "Importance and possibilities of development of critical thinking in the university: the teacher's perspective," *Multidisciplinary Journal of Educational Research*, vol. 11, no. 1, 20 pages, 2021.

[2] A. Rakhmawan, H. Firman, S. Redjeki, and S. Mulyani, "Contribution of logical thinking ability to students' achievement in three level of representations in chemical dynamic materials," *Jurnal Penelitian Dan Pembelajaran IPA*, vol. 4, no. 2, 116 pages, 2018.

[3] S. Andriani, E. Nurlaelah, and K. Yulianti, "The effect of process oriented guided inquiry learning (pogil) model toward students' logical thinking ability in mathematics," *Journal of Physics Conference Series*, vol. 1157, no. 4, Article ID 04, 2019.

[4] S. Murni, S. Ruqoyyah, and L. H. Fasha, "Analysis of mathematical thinking skills in multiple intelligence perspectives of primary school students," *Journal of Physics: Conference Series*, vol. 1657, no. 1, Article ID 012013, 2020.

[5] H. Riyanti, S. Suciati, S. Suciati, and P. Karyanto, "Enhancing student' logical-thinking ability in natural science learning with generative learning model," *Biosaintifika: Journal of Biology & Biology Education*, vol. 10, no. 3, pp. 648–654, 2018.

[6] D. Cabral and C. Sales, "Contributions and implications of the dialogical perspective: the therapist's self in the family/couple therapy," *Nova Perspectiva Sistêmica*, vol. 28, no. 63, pp. 21–41, 2019.

[7] R. Rahmawati, N. Y. Rustaman, I. Hamidah, and D. Rusdiana, "The profile of cognitive style, logical thinking ability, and conceptual knowledge of electricity and magnetism topic based on prospective physics teachers' grade level," *Journal of Physics: Conference Series*, vol. 1157, no. 3, Article ID 032036, 2019.

[8] S. Gröschl and P. Gabaldon, "Business schools and the development of responsible leaders: a proposition of edgar morin's transdisciplinarity," *Journal of Business Ethics*, vol. 153, no. 1, pp. 185–195, 2018.

[9] Y. X. Lu, "The exploration of university practical teaching and supporting evidence from the perspective of ability training—illustrated in the case of Heilongjiang provincial h university social work major," *Journal of Heilongjiang Institute of Technology*, vol. 33, no. 3, pp. 73–76, 2019.

[10] B. Marc, C. Alessandra, C. Alessandra, and W. Kerstin, "The function of personality in suicidal ideation from the perspective of the interpersonal-psychological theory of suicide," *International Journal of Environmental Research & Public Health*, vol. 15, no. 4, pp. 636–638, 2018.

[11] R. Arifin, R. Alkadri, D. P. Sari, L. Resthiningsih, and A. M. Holish, "Improving law student ability on legal writing through critical and logical thinking by irac method," *Indonesian Journal of Advocacy and Legal Services*, vol. 1, no. 1, pp. 107–128, 2019.

[12] 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.

[13] N. Wang, "A comparative study of folk culture translation from the perspective of cultural presupposition—taking four English versions of biancheng as an example," *Open Access Library Journal*, vol. 7, no. 5, pp. 1–9, 2020.

[14] M. Khoroshukha, A. Bosenko, S. Prysiazhniuk, O. Tymchyk, and J. Nevedomsjka, "Influence of sexual dimorphism on the development of the logical thinking function in young athletes aged 13–15 years with different blood groups," *Georgian Medical News*, vol. 309, pp. 108–114, 2020.

[15] C. H. Peng, "Simulation of optimal feature detection and location method under chaotic data structure," *Computer Simulation*, vol. 35, no. 3, pp. 345–380, 2018.

[16] J. Yuan, W. Li, B. Xia, Y. Chen, and M. J. Skibniewski, "Operation performance measurement of public rental housing delivery by ppps with fuzzy-ahp comprehensive evaluation," *International Journal of Strategic Property Management*, vol. 23, no. 5, pp. 328–353, 2019.

[17] H. Quanhua, Z. Wenting, L. Xuanye, and Z. Xuan, "Urban ecosystem health assessment inside Xi'an by multivariate data," *Journal of Environmental Protection and Ecology*, vol. 19, no. 3, pp. 1016–1025, 2018.

[18] S. R. Manurung and D. D. Panggabean, "The identification of logical thinking ability based on gender in physics teachers prospective students," *Journal of Community Research and Service*, vol. 2, no. 1, 192 pages, 2018.

[19] H. Riyanti, S. Suciati, and P. Karyanto, "The effectiveness of generative learning model to enhance students' logical-thinking ability in science learning," *EDUSAINS*, vol. 10, no. 2, pp. 309–318, 2018.

[20] A. Bahtiar, A. Syamsuddin, and I. Akib, "Description of mathematical communication skills, logical thinking and its influence on the ability of mathematical literacy for students of grade v elementary school," *International Journal of Scientific & Technology Research*, vol. 9, no. 4, pp. 1075–1078, 2020.

Hindawi
Scientific Programming
Volume 2021, Article ID 9921960, 7 pages
https://doi.org/10.1155/2021/9921960

Review Article

Genetic Algorithm and BP Neural Network for College Physical Education Teaching Evaluation

Shuxin Wang

Sias University, Zhengzhou, Henan, China

Correspondence should be addressed to Shuxin Wang; 11138@sias.edu.cn

Received 19 March 2021; Accepted 25 May 2021; Published 31 May 2021

Academic Editor: Shah Nazir

Physical education is considered as the essential part of health plan and also reflects the higher education system. A teacher with physical education is accountable to teach physical education and/or health education to the students. This study has proposed a new approach for teaching evaluation with college physical education in universities. By combining neural network with genetic algorithm (GA), the approach that we proposed can improve the general traditional BP network, such as global convergence and training time. The model uses MATLAB software for empirical research work. According to MATLAB simulation and experiments, it was revealed that the combining neural network with GA algorithm has effective application prospects for the teaching evaluation with college physical education in universities.

1. Introduction

At present, we always do the teaching evaluation through different methods and approaches, as the teaching quality has always been the focus of social attention. In China, we have a huge amount of higher education system in which a large number of students are studying. The evaluation of training quality of a university is an important issue in front of us which need proper evaluation mechanism. The quality evaluation system has been considered as an important part of education [1–5]. At present, many teaching quality evaluation systems are established according to the standards of ordinary colleges and universities level. In the current evaluation system, there are some problems for the evaluation of university education quality which includes the following. (a) The evaluation index system is not reasonable which cannot reflect the characteristics for university physical education. (b) The teaching evaluation index often contains many no quantitative factors, so in the evaluation, there are some complex nonlinear relationships between the input and output. So, it is hard to build precise mathematical models through a relationship. (c) The teaching quality evaluation methods, such as the absolute evaluation method, the relative evaluation method, the rating method, the comment method, the realistic method, and the comprehensive scoring method, are either too subjective, or simple mathematical operations (such as addition, subtraction, multiplication, and division) are used to evaluate the teaching effect, ignoring the nonlinear relationship between the evaluation indexes and the teaching effect, and the results are difficult to reflect the teaching quality; it is also difficult to be accepted by people [6–13]. (d) It is difficult to make an accurate evaluation by using traditional methods to evaluate the results of some indicators, and the calculation is complex, the solution is cumbersome, and these algorithms also lack of self-learning ability [14–16].

To overcome these problems, this study has proposed a novel approach for teaching evaluation with college physical education in universities. By combining neural network with genetic algorithm, the current approach can progress the general traditional BP network, such as global convergence and training time. The approach has used MATLAB software for experimental work. Based on the simulation and experiments, it was shown that combining the neural

network with GA algorithm has operational applications for the teaching evaluation with college physical education in universities.

2. BP Neural Network with Genetic Algorithm

The following sections briefly show the details.

2.1. Genetic Algorithm. Genetic algorithm (GA) mainly includes selection, crossover, mutation, and other operations. By combining neural network with genetic algorithm, it not only makes use of the strong learning ability of neural network but also combines the global search characteristics of genetic algorithm.

In this paper, we introduce the GA into the BP neural network for the evaluation model, which is a further optimization application of BP neural network. Various applications of GA exist [17, 18].

2.2. BP Neural Network System Model Based on Genetic Algorithm. In BP neural network, we use the descent method as the learning rule. By adjusting the weight value of each layer, we try to minimize the square sum of network error. As shown in Figure 1, firstly, representative teaching quality evaluation indexes are selected as input information; then, we sent these evaluation indexes to the input layer, hidden layer, and output layer for training. We use the genetic algorithm to iteratively learn the network weights according to the training objective function to find the best weights. The network becomes a stable pattern evaluator, and then, the evaluation results can be output.

The number of input nodes of the model is n (evaluation index of college physical education quality), the number of middle layer nodes is n, and the number of output layer nodes is 1, which is the evaluation value of college physical education quality. The input layer node transmits the information directly to the middle layer node, so the output of the input layer node is equal to the input; the output information of the middle layer node is the input information of the output layer node, and the output layer has only one node, which receives the input of the middle layer node and outputs the evaluation results of the teaching quality:

(1) Input layer node $p_i, i = \{1, 2, \ldots, n\}$, that is, n represents teaching quality assessment

(2) Middle layer node

The input is

$$I_j = \sum_{j=1}^{n} \omega_{ij} p_i. \tag{1}$$

The output is

$$O_j = \frac{1}{\left\{1 + \left[\left(\sum_{i=0}^{n} \omega_{ij} p_i\right)^{-1} - 1\right]^2\right\}} = \frac{1}{\left[\left(I_j^{-1} - 1\right)^2\right]}, \tag{2}$$

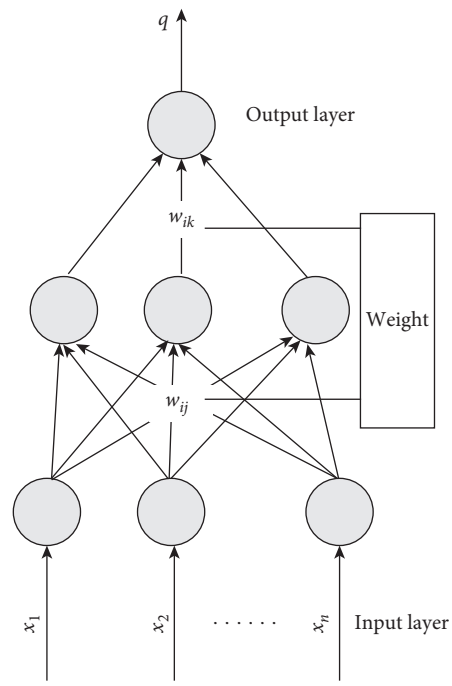


Figure 1: BP neural network training model with genetic algorithm.

where ω_i is the weight from the input layer node i to the middle layer node j and p_i represents the ith factor of the sample, i.e., the ith teaching quality evaluation index.

(3) Output layer node: the output layer has only one node, and the input is the output of the middle layer node:

$$q = \frac{1}{\left\{1 + \left[\left(\sum_{i=0}^{n} \omega_j O_j\right)^{-1} - 1\right]\right\}^2}, \tag{3}$$

where q is the actual output value of the sample, that is, the calculated teaching quality evaluation value and ω_j is the connection weight from the middle layer node i to the output layer node.

2.3. Genetic Algorithm

2.3.1. Coding Mode. The common genetic coding methods are real number encoding and binary encoding. Real number coding has high precision and is easy to search in a large space; binary coding is in line with the important biological characteristics of diploid chromosomes in higher organisms and has the function of long-term memory, which is convenient for various genetic operations. To facilitate genetic operation, binary coding is used for weight coefficient.

2.3.2. Training Objective Function. The training objective function is defined as the mean value of the squared sum of the error between the actual output and the evaluation value of P samples:

$$E = \left(\frac{1}{p}\right)\sum_{p=1}^{p}[\overline{q} - q]^2 = \left(\frac{1}{p}\right)\sum_{p=1}^{p}E_j, \tag{4}$$

where P is the total logarithm of the training sample pattern, q is the actual output value of the sample network, and q is the expected value of the sample network. The actual output Q of the identified object (the teaching quality evaluation model of Higher Vocational Colleges) is taken as the feedback signal, which is compared with the expected value of the neural network sample to make $E < e$. If e does not meet the requirements due to the small number, the weight coefficient is adjusted continuously to meet the corresponding requirements.

2.3.3. Weight Adjustment Model of BP Neural Network Evaluation System. The purpose of network learning is to minimize e by adjusting the connection weight in the network. Using the gradient descent method, the adjustment of connection weight is as follows:

$$\begin{cases} \omega_{ij} = -\eta\left(\dfrac{\partial E}{\partial v_{ij}}\right), \\ \\ \omega_j = -\eta\left(\dfrac{\partial E}{\partial w_{ij}}\right). \end{cases} \tag{5}$$

Among η, for the learning rate, $0 < \eta < 1$. Then, the connection weight adjustment amount between the input layer node and the middle layer node is

$$\omega_{ij} = x_i w_j O_j^2\left[1 - \sum_{i=1}^{n}\omega_{ij}x_j\right]\delta_j. \tag{6}$$

The adjustment amount of connection weight is as follows:

$$w_j = q^2 O_j\left[1 - \sum_{j=1}^{l}w_j O_j\right][\overline{q} - q]^2. \tag{7}$$

By using the above model, the connection weight of the neural network can be determined according to the iterative algorithm of the common neural network, and the error between the actual output and the sample value can be reduced.

2.3.4. Description of Connection Weight Optimization Problem Based on Optimal Neural Network.

$$\min(E) = f(w_1, \ldots, w_n), \tag{8}$$

where e is the total error of network training and $w_1, \ldots, w_n$ are the continuous weights after unified numbering which include the connection weights of input layer nodes and middle layer nodes and the connection weights of middle layer nodes and output layer nodes, n is the total number of connection weights. The constraints are $a_1 \le w_i \le b_1$. Among them, b_1 and a_1 are variables of w_i of the upper and lower bounds of change.

2.3.5. Generation of Initial Population. As for the selection of initial population size, according to the research, the larger the nonlinearity of the problem, the larger the P. In this paper, $P = 30$, and the initial population is composed of P randomly generated strings.

2.3.6. Calculation of Fitness. The calculation of population fitness function is always a bottleneck problem in the application of genetic algorithms. Because the selection probability of individuals is directly proportional to fitness, the definition of fitness function has a great influence on the genetic algorithms. There is a close mapping relationship between the fitness function and objective function, so the following fitness calculation method is adopted:

$$f_i = \begin{cases} C - E & E < C, \\ 0 & E \ge C, \end{cases} \tag{9}$$

where e is the training objective function and C is the maximum of all e in the current generation.

2.3.7. Genetic Manipulation Produces New Populations. Genetic manipulation mainly includes replication, crossover, and mutation. The purpose of copy operation is to improve the system global convergence and convergence efficiency. The main idea is that the replication probability PR of the string is proportional to its fitness, that is, the individuals with high fitness will be copied to the next generation with high probability. Cross operation can combine new individuals and search effectively in strand space to avoid effective gene deletion. Only in the case of effective gene deletion, the effect of mutation operation is more obvious. To take into account the efficiency of global convergence and avoid premature convergence caused by effective gene deletion, the following strategies are adopted in the selection of genetic operation parameters:

$$P_r = \begin{cases} 2\left(\dfrac{f_1}{C}\right)^2, & \dfrac{f_1}{C} \in [0, 0.5], \\ \\ \left[1 - 2\left(1 - \left(\dfrac{f_1}{C}\right)^2\right)\right], & \dfrac{f_1}{C} \in [0.5, 1], \end{cases} \tag{10}$$

where C is the parameter selected to define the fitness and f_1 is fitness. Equation (10) makes the replication probability change with the change of fitness, that is, the individuals with high fitness will be inherited to the next generation with a higher probability and the individuals with low fitness will be inherited to the next generation with a lower probability. The results show that the crossover probability PC is usually [0.65, 0.9], while the mutation probability PM is usually [0.001, 0.1].

2.4. The Steps of Genetic Algorithm Training Are Neural Network Algorithm

(1) The training sample data is selected, and the connection weights of input layer and middle layer and

middle layer and output layer are generated randomly: ω_{ij} and ω_j

(2) The actual output Q of the sample is calculated according to equations (1)–(3)

(3) According to equation (4), calculate the value function e of the square sum of the error between the actual output value and the expected value

(4) If $e \leq e$ meets the error requirements, the training is over, and go to step (11); otherwise, go to step (5)

(5) The weight adjustment values of the input layer and the middle layer and the middle layer and the output layer are calculated according to equations (6) and (7)

(6) Recalculate the new connection weight according to equations (9) and (10): ω_{ij} and ω_j

(7) According to the new weights and sample data, the average error e of the actual output Q network of samples is recalculated according to equations (1)–(4)

(8) If $E \leq \delta$ and the error requirement is met, the training is over, and go to step (11); otherwise, go to step (9)

(9) Set the population size n, weight change range, crossover probability PC, and mutation probability PM, and determine the replication probability PR according to equations (10) and (11)

(10) Copy crossover, mutation, and genetic algorithm calculation, according to the network weight generated by the genetic algorithm, as the initial weight value of the next training of the neural network model, and go to step (5)

(11) At the end of training, output the weight that meets the training accuracy: ω_{ij} and ω_j

(12) According to the determined weight values ω_{ij} and ω_j, put the experimental data to be evaluated and output the evaluation results

By fixing the weights of the neural network that has completed the training task, the network model constitutes a pattern evaluator.

3. Evaluation of Physical Education Teaching Quality in Universities

With the expansion of science and information technology, the applications of computer technology have been extensively adopted in all aspects of life. Class room teachings and various sports activities are enriched with the support of computer technology [19]. Various research studies have been conducted associated to physical education. A study has analyzed 35 physical education teachers' appreciation of a continuous professional development training embedded in self-determination theory, need-supportive teaching, and using quantitative and qualitative methods [20]. The application of virtual reality and web embedded system in the field of evaluation of physical education teaching is used [21]. The study has evaluated the efficacy of need-supportive teaching in physical education on daily moderate-to-vigorous physical activity of girls through the mixed method evaluation [22].

3.1. Establishing the Evaluation Index Teaching Quality. In order to evaluate the teaching quality more scientifically and effectively, we must first determine a set of scientific evaluation index systems of teaching quality. On the basis of interviews with some teachers and students of Sias University, Zhengzhou University of Industrial Technology, Huanghuai University, and Huanghe Science and Technology University, combined with the summary of the theory and practice of the research on the evaluation of the quality of physical education teaching in universities, this paper establishes the evaluation index system of the quality of physical education teaching in universities, as shown in Table 1. From Table 1, we know the index system determines the network structure of the teaching-quality evaluation model.

3.2. Normalization of Input Index. Because the secondary index input is obtained by using the percentile system and scoring by students, the magnitude of each component value is very different. If the original data is applied directly without any transformation, the absolute value of the original data may be too large and fall outside the effective processing range of neurons, resulting in the so-called "saturation phenomenon." Even if the absolute value of the original data is not too large, it is possible that one component is too large, and the impact on the network is far greater than that of other components so that other components lose the ability to regulate the network. Therefore, it is necessary to normalize the input samples of the neural network. In this paper, the maximum and minimum method is used for normalization processing because this method is a linear transformation for data processing, which can better retain its original meaning and will not cause information loss. For the input data, the normalization formula used in this paper is as follows:

$$X = \frac{I - I_{\min}}{I_{\max} - I_{\min}}, \tag{11}$$

where x is the input value of the neural network after normalization, I is the input value of the neural network without processing, $I_{\max}$ is the maximum input of neural network, and $I_{\min}$ is the minimum input of neural network.

3.3. Preparation of Sample Library. In the GABP neural network, based on the summary and analysis, according to the evaluation index of teaching quality, this paper formulates the teaching questionnaire of higher vocational colleges so that the students in our college can choose and score the indicators of teachers. The collected sample data were standardized according to formula (11). These data are transformed into data [0, 1]. The processed sample data is shown in Table 2.

Table 1: Evaluation index system of college physical education teaching quality.

First-level indicators	Secondary indicators	Input	First-level indicators	Secondary indicators	Input
Teaching attitude	Rigorous lesson preparation	X1	Content of courses	Large proportion of practical training	X10
	Operation wholesale	X2		Emphasis	X11
Lecture ability	Content system	X3	Teaching method	Teach students in accordance with their aptitude	X12
	Complex problems	X4		Focus on inspiration	X13
	Integrating theory with practice	X5	Teaching and educating	Interaction between learning and students	X14
	Heuristics	X6		Be a model for others	X15
	Key point and difficulty	X7		Strict requirements	X16
	Demonstration	X8	Teaching effectiveness	Students should master it well	X17
	Students' enthusiasm	X9		Students are good at doing things	X18

Table 2: Training sample set after data preprocessing.

No.	1	2	3	4	5	6	7	8	9	10	11	12	13	14
X1	0.51	0.69	0.65	0.66	0.68	0.55	0.56	0.98	0.98	0.70	0.88	0.50	0.51	0.71
X2	0.53	0.52	0.55	0.96	0.30	0.53	0.95	0.80	0.95	0.51	0.93	0.59	0.91	0.55
X3	0.70	0.96	0.88	0.65	0.31	0.45	0.80	0.6	0.72	0.68	0.67	0.69	0.30	0.65
X4	0.73	0.70	0.59	0.67	0.31	0.49	0.70	0.61	0.70	0.67	0.87	0.69	0.30	0.56
X5	0.52	0.96	0.97	0.51	0.55	0.45	0.51	0.34	0.95	0.48	0.93	0.48	0.31	0.59
X6	0.98	0.97	0.98	0.98	0.33	0.51	0.71	0.32	0.75	0.72	0.91	0.67	0.33	0.54
X7	0.70	0.72	0.68	0.69	0.51	0.52	0.51	0.51	0.69	0.53	0.69	0.75	0.49	0.49
X8	0.51	0.69	0.93	0.75	0.69	0.48	0.68	0.48	0.71	0.67	0.76	0.71	0.48	0.49
X9	0.96	0.95	0.95	0.95	0.53	0.51	0.72	0.69	0.92	0.75	0.95	0.65	0.51	0.67
X10	0.71	0.98	0.96	0.65	0.48	0.69	0.72	0.69	0.92	0.75	0.95	0.65	0.51	0.68
X11	0.78	0.68	0.69	0.72	0.66	0.65	0.60	0.68	0.70	0.82	0.98	0.64	0.54	0.69
X12	0.96	0.93	0.99	0.89	0.81	0.49	0.50	0.56	0.89	0.47	0.93	0.51	0.33	0.49
X13	0.96	0.97	0.91	0.92	0.31	0.33	0.71	0.31	0.71	0.71	0.71	0.71	0.31	0.71
X14	0.72	0.69	0.68	0.69	0.53	0.72	0.65	0.51	0.71	0.67	0.69	0.73	0.51	0.51
X15	0.70	0.71	0.88	0.71	0.59	0.68	0.55	0.75	0.72	0.65	0.41	0.71	0.31	0.69
X16	0.70	0.02	0.65	0.68	0.66	0.71	0.89	0.65	0.75	0.73	0.59	0.49	0.28	0.69
X17	0.93	0.70	0.89	0.73	0.67	0.65	0.67	.071	0.71	0.75	0.68	0.72	0.27	0.72
X18	0.71	0.92	0.74	0.98	0.30	0.32	0.63	0.58	0.89	0.71	0.70	0.69	0.31	0.69

3.4. Simulation Process and Results. The neural network toolbox in MATLAB contains the realization of various neural network algorithm programs. The realization of BP neural network by MATLAB generally has four steps: the first is the initialization of the network, the second is the creation of the network, the third is the simulation of the network, and the fourth is the training of the network. To test the performance of GA neural network training model, read the training data and teacher data (the first 10 groups of data in Table 2), generate the corresponding network model according to the neural network model structure proposed in Section 2.2, and set the error precision of network training as $e = 0.001$. The operation parameters of the neural network are set as follows: learning efficiency 0.9 and momentum coefficient 0.9, and the operation parameters of genetic algorithm are set as follows: weight variation range 0.5, crossover probability 0.75, and mutation probability 0.005. The simulation process is as follows.

3.4.1. Simulation Process

```
Fid = fopen(train_data.txt);
P = fscanf(fid,%g, [10, 18]);
%P = mydata
T = fscanf(fid,%g, [1, 10]);
Net = = newff([0,1], [1, 7],{'tansig','tansig'},traingdm');
Net.trainParam.show = 50;
. . .
net.trainParam.goal = 0.001;
fidt = fopen('test_data.txt');
while(feof(fidt))
Q = fscanf(fidt,%g, [4, 18]);
B = sim(net,Q);
End
```

Table 3 shows the results with BPGA, and the simulation evaluation results and expert evaluation results of the four test sets (the last four groups of data in Table 2 are shown in Table 4). From Tables 3 and 4, we know that it is not only the training samples which are very close to the expert evaluation results but also the simulation evaluation results of the four test sets are very close to the expert evaluation results.

Table 3: Comparison of neural network training results based on genetic algorithm and actual evaluation results.

Number	1	2	3	4	5	6	7	8	9	10
Expert evaluation	0.700	0.750	0.910	0.710	0.450	0.571	0.680	0.520	0.850	0.650
Network test	0.710	0.740	0.711	0.720	0.464	0.573	0.670	0.0490	0.860	0.651

Table 4: Comparison between test results of the test set and actual evaluation results.

Number	11	12	13	14
Expert evaluation	0.825	0.630	0.511	0.630
Network test	0.810	0.651	0.535	0.610

4. Conclusions

The contribution of the paper is to improve the traditional BP neural network using the GA. As the GA can iterative the weight factor to change the input of BP neural network. The basic idea is to use GA to optimize the input parameters of the neural network and speed the system convergence, finally, taking the optimization result as the initial value of BP algorithm, and then, train the network with BP algorithm. In this way, BP algorithm with GA is alternately run until the required accuracy of the problem is reached. In the experiment, we use many groups of sample data of our college to test and establish a BP neural network model based on genetic algorithm. Through the optimization of neural network connection weights, that is, the genetic evolution of the initial population generation by generation, the evaluation of college physical education teaching quality is realized. Results show the effectiveness of the study.

Data Availability

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

Conflicts of Interest

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

Acknowledgments

This study was supported by the project of Science and Technology of Henan Province (Social Development), Research on the Integrated Development of Physical Exercise and Physical Therapy from the Perspective of Multiagent Collaboration (no. 202102310651), in 2020.

References

[1] L. Song and J. Huang, "Design of teaching quality evaluation model based on BP neural network," *Electromechanical technology*, vol. 19, no. 5, pp. 1–10, 2020.

[2] L. Gu, "Evaluation of multimedia teaching quality based on grey relational analysis and neural network," *Modern Electronic Technology*, vol. 43, no. 9, pp. 191–194, 2020.

[3] J. Zhang, "Evaluation of english teaching quality based on GA optimized RBF neural network," *Computer System Application*, vol. 29, no. 3, pp. 171–176, 2020.

[4] Y. Wu, C. Wu, and Y. Yin, "Internet teaching quality evaluation model based on evidence theory and neural network," *Modern Electronic Technology*, vol. 43, no. 19, pp. 183–186, 2020.

[5] Y. Li, "Design of university teaching quality evaluation model based on data mining algorithm," *Modern Electronic Technology*, vol. 43, no. 17, pp. 127–130, 2020.

[6] L. Zhang, F. Zheng, L. Wang et al., "Research and application of curriculum quality evaluation system based on BP neural network," *Journal of Chifeng University (Natural Science Edition)*, vol. 23, no. 5, pp. 13–18, 2020.

[7] Z. Zhang, Y. Lang, and Z. Xie, "Research on real time emotion recognition based on exception in classroom quality analysis," *Automation Technology and Application*, vol. 39, no. 6, pp. 52–57, 2020.

[8] G. Zuo, X. Su, H. Wang et al., "Research on classroom behavior detection of college students based on CNN deep learning model," *Intelligent Computer and Application*, vol. 10, no. 2, pp. 158–160, 2020.

[9] H. Cui, *Research on the Application of K-Modes Algorithm and Neural Network in Teaching Evaluation and Learning Prediction*, vol. 15, no. 2, pp. 56–63, 2020.

[10] L. Zhang, F. Zheng, L. Wang et al., "Research and application of curriculum quality evaluation system based on BP neural network," *Journal of Chifeng University (Natural Science Edition)*, vol. 036, no. 1, pp. 102–104, 2020.

[11] Q. Wu, L. Qiang, and X. Guan, "Optical score recognition method based on multi-scale residual convolution neural network and bidirectional simple cycle unit," *Progress in Laser and Optoelectronics*, vol. 57, no. 8, 2020.

[12] X. Lian, J. Huang, and J. Wu, "Case study on practical teaching of neural network theory and application," *Computer and Information Technology*, vol. 28, no. 6, pp. 71–74, 2020.

[13] L. Shi, C. Cheng, and S. Yi, "Research on the evaluation of college students' practical teaching effect based on BP neural network," *Journal of Yangzhou University (Higher Education Research Edition)*, vol. 24, no. 2, pp. 115–121, 2020.

[14] P. Su, "A feedback teaching mode based on machine learning," *Curriculum Education Research*, vol. 22, no. 14, 2020.

[15] G. Cheng, Y. Xu, and Y. Zheng, "Exploration of metalworking practice teaching reform based on machine vision to improve practice safety," *Experimental Technology and Management*, vol. 37, no. 5, pp. 265–269, 2020.

[16] Z. Yuanyi and L. Zheng, "Interactive text classification method of online collaborative learning based on deep neural network," *Modern Distance Education Research*, vol. 32, no. 3, pp. 106–114, 2020.

[17] A. Semnani, M. Ostadhassan, Y. Xu, M. Sharifi, and B. Liu, "Joint optimization of constrained well placement and control parameters using teaching-learning based optimization and an inter-distance algorithm," *Journal of Petroleum Science and Engineering*, vol. 203, Article ID 108652, 2021.

[18] A. Khosravi, M. Malekan, J. J. G. Pabon, X. Zhao, and M. E. H. Assad, "Design parameter modelling of solar power

tower system using adaptive neuro-fuzzy inference system optimized with a combination of genetic algorithm and teaching learning-based optimization algorithm," *Journal of Cleaner Production*, vol. 244, Article ID 118904, 2020.

[19] P. Dong and A. Li, "Simulation of physical education teaching based on FPGA and wearable VR equipment," *Microprocessors and Microsystems*, vol. 81, Article ID 103773, 2021.

[20] N. Aelterman, M. Vansteenkiste, H. Van Keer, J. De Meyer, L. Van den Berghe, and L. Haerens, "Development and evaluation of a training on need-supportive teaching in physical education: qualitative and quantitative findings," *Teaching and Teacher Education*, vol. 29, pp. 64–75, 2013.

[21] J. Cheng, "Evaluation of physical education teaching based on web embedded system and virtual reality," *Microprocessors and Microsystems*, vol. 83, Article ID 103980, 2021.

[22] D. J. Sturm, J. Bachner, D. Renninger, S. Haug, and Y. Demetriou, "A cluster randomized trial to evaluate need-supportive teaching in physical education on physical activity of sixth-grade girls: a mixed method study," *Psychology of Sport and Exercise*, vol. 54, Article ID 101902, 2021.

Hindawi
Scientific Programming
Volume 2021, Article ID 9936217, 13 pages
https://doi.org/10.1155/2021/9936217

Research Article

Research on the Design of Government Affairs Platform in the Context of Big Data

Qian Huang[1,2] **and Xue Wen Li**[3]

[1]*School of Information Management Wuhan University, Wuhan 430072, China*
[2]*School of Marxism North Minzu University, Yinchuan 750021, China*
[3]*North Minzu University Library, Yinchuan 750021, China*

Correspondence should be addressed to Qian Huang; huangqian110@whu.edu.cn

Received 26 March 2021; Revised 26 April 2021; Accepted 5 May 2021; Published 30 May 2021

Academic Editor: Shah Nazir

Big data is a massive and diverse form of unstructured data, which needs proper analysis and management. It is another great technological revolution after the Internet, the Internet of Things, and cloud computing. This paper firstly studies the related concepts and basic theories as the origin of research. Secondly, it analyzes in depth the problems and challenges faced by Chinese government management under the impact of big data. Again, we explore the opportunities that big data brings to government management in terms of management efficiency, administrative capacity, and public services and believe that governments should seize opportunities to make changes. Brainlike computing attempts to simulate the structure and information processing process of biological neural network. This paper firstly analyzes the development status of e-government at home and abroad, studies the service-oriented architecture (SOA) and web services technology, deeply studies the e-government and SOA theory, and discusses this based on the development status of e-government in a certain region. Then, the deep learning algorithm is used to construct the monitoring platform to monitor the government behavior in real time, and the deep learning algorithm is used to conduct in-depth mining to analyze the government's intention behavior.

1. Introduction

Government management is bound to have major reforms and innovations under the help of big data. It will not only enhance the government's social management capabilities, but also improve the government's management level. However, the application of big data in China is still in a relatively basic development stage. The government will face many difficulties in applying big data for management and decision-making. It is necessary to clearly understand these issues before they can find breakthrough points one by one. The application of data in government management lays a solid foundation. Under the impact of big data, the scientific and technological progress stimulated by the massive data on the Internet will inevitably bring about tremendous changes in the way of life, thinking, and decision-making of human beings. Under these changes, the value of correctly using big data will be social life and economic development, which have far-reaching effects. First, big data revolutionizes the way humans think [1]. Big data changes the shift of human thinking patterns from causality to related relationships. Today's data in life is growing faster than ever before, and the amount of data is growing wildly. Brain-like computing seeks to design a system like biological neural network from the structure, seeks to simulate the function of the brain from the function, and seeks to greatly surpass the biological brain from the performance, also known as neuromorphic computing. The rapid update of information in the context of big data and the rapid changes in the environment require people to respond quickly when they encounter problems, rather than to explore the reasons. Second, big data revolutionizes the way humans make decisions.

In the past era of small data, human judgments on things and perceptions often rely on intuition and experience, and

the resulting bias can be seen everywhere. The scientific predictions and decisions brought about by big data will inevitably change the way people make decisions. The significance of big data to decision makers lies in advance prediction, in-process perception, and after-the-fact feedback. People will be more dependent on data and analysis than on intuition and experience; likewise, it will change the nature of leadership and management. For the government, big data analysis has promoted the scientific decision-making of the government. Decision-making entities use big data to obtain real-time information in the decision-making process, master development trends, and make accurate predictions and scientific decisions. In resource allocation, big data also acts as the third hand. Big data has played a crucial role in driving decision-makers to change decisions and improve decision-making, and this is gradually rising to the national consciousness. Third, big data drives the modern transformation of government management. First, big data helps diversify management entities. In the era of small data, because of limited access to information, people are in an environment of information confinement, and these information resources are usually in the hands of some leaders, resulting in the public lack of understanding of information, and thus, they cannot effectively participate in management [2].

The arrival of the era of big data breaks the phenomenon of monopoly of information. People can obtain massive amounts of data information through various channels such as the Internet at all times, thus supervising the management activities of managers and making them more transparent and transparent. Big data also provides a channel for public participation. People can express their opinions through social networking sites at any time and express their needs. The public is an important factor in management activities, and big data helps achieve diversification of management entities. Second, big data facilitates the transformation of one-way management from collaborative shared management. In the past, different government departments had their own information systems, which caused many data to be uncommon and isolated from each other, forming an isolated island of information. Big data broke this barrier, realized data sharing, and maximized the effectiveness of data. It has improved the economic and social development and brought more convenience to the public. Finally, big data helps improve the scientific level of the decision-making process. In the past, management decision makers often used past experience to make decisions and judgments. Today's changes in the social environment and relationships often result in decision-making mistakes. Scientific decision-making under the foundation and guarantee of big data has become a trend of development. Decision makers are aware of the importance of data and use data for analysis and judgment to make scientific decisions [3]. The specific contributions of this paper include the following:

(i) The problems and challenges faced by Chinese government management under the impact of big data are deeply analyzed

(ii) Deep learning algorithm is used to construct a monitoring platform for real-time monitoring of government behavior

(iii) The development status of e-government at home and abroad is analyzed

(iv) Using big data to update information quickly and making quick judgments

The rest of this paper is organized as follows. Section 2 discusses basic technology of SSEGA System Architecture, followed by construction of data platform based on brainlike computing in Section 3. SBEGA application in e-government platform is discussed in Section 4. Section 5 concludes the paper with summary and future research directions.

2. Government Platform Architecture Based on SSEGA System

Service-oriented architecture (SOA) was first proposed by Gartner in a 1996 report [4]. It is a service-centric, loosely coupled, dynamically optimized and reusable extended distributed applied constructor. After more than ten years of practice and development, it has become the trend of constructing integrated distributed information systems in an open and heterogeneous network environment. The loose coupling and interoperability of IT systems are brought by SOA [4], and the resulting large-scale reuse, large-scale integration, flexibility enhancement, etc. are the establishment, integration, operation, and maintenance of software systems. In particular, the innovation and development of the Internet-based software industry have brought new impetus and opportunities [5]. With the in-depth development of IT systems, the emergence of the "information island" problem and the rise of XML language and web services technology have objectively promoted the development of SOA. SOA has gradually entered people's careers and gradually shifted from concept to application. SOA requires a series of technical support, combined with web services and XML technology, to make the architecture more flexible and open; in the aspect of security authentication, SOAP can enhance interface access control and data security; in data storage, through distributed database processing, high concurrent requests can also dynamically add storage nodes. This chapter analyzes technologies such as service-oriented architecture technology and related network services and extended markup languages.

2.1. Basic Technology of SSEGA System Architecture. Combining the characteristics of SOA and web services technologies, it is necessary to change the current situation that the information between e-government systems of B/S three-tier structure is difficult to share and interconnect, and use web services technology to encapsulate core services into open-access and remotely accessible services. It is a viable way to solve this problem. As an open service implementation specification, web services have been widely accepted by information system providers and IT technicians around the

world and have gradually become the de facto standard for service implementation. [6].

Web services solve the problem of service implementation extremely well, but the emergence of services also challenges the traditional system architecture. The traditional B/S three-tier structure cannot adapt well to service-oriented applications and must be from the system architecture. There is innovation at the level. Service-Oriented Architecture (SOA) was developed under such circumstances. Although SOA is still in the process of continuous improvement, it has been accepted and sought after by major platform providers and system providers. As described in Section 1, several of the world's largest IT vendors are launching their own SOA platforms and solutions, all of which provide the impetus for the overall development of SOA.

The problem now is that although everyone is pushing their own SOA platform and solutions, in actual project applications, SOA still mostly stays in the application of integration of legacy systems, and there is a lack of reference in the field of e-government, a common system architecture. Referring to the SOA architecture template proposed by IBM, this paper proposes SOA-Based e-Government Architecture (SBEGA), as shown in Figure 1.

2.2. SOA-Based e-Government Architecture

2.2.1. Mathematical Model of SOA Algorithm. The SOA algorithm directly simulates the intelligent behavior of human random search and is directly applied to the optimization problem solving. In the calculation process of the optimization solution, the intelligent search behavior can be understood as follows: in a search in a continuous space, generally better solutions may be in the vicinity of the better solution, and the optimal solution may be better solution. Therefore, if the searcher individual is in a relatively good position, the search scope should be set in a small space; if the searcher's location is not ideal, you should choose to search in a relatively large neighborhood. Therefore, when establishing a mathematical model for search criteria, SOA uses fuzzy logic that can effectively describe the language and uncertainty reasoning, and determines the step value of the search [7].

After a series of learning, the SOA algorithm gains relevant experience, combines the aggregation behavior of intelligent population spontaneous organization, the self-interested self-priority behavior, and the pre-action behavior with predictive guidance ability, and finally establishes the individual. This is the direction of the search.

The indeterminate reasoning behavior in the SOA algorithm refers to the use of the approximation ability of the fuzzy system to simulate the intelligent search behavior of humans, and then to establish a bridge between behavior and cognition. By learning the intelligent random search behavior of human beings, we can use the fuzzy rules to describe the human search behavior: the smaller objective function corresponds to a smaller search step, and the search step is blurred [8].

Variables can usually be represented using Gaussian membership functions:

$$\mu_A(x) = \exp\left[\frac{-(x-\mu)^2}{2\delta^2}\right], \tag{1}$$

where μ_A is Gaussian membership; x is the input variable; μ, δ are the membership function parameters.

If the value of the output variable is $[\mu - 3\delta, \mu + 3\delta]$ exceeded, and if $\mu_A(\mu + 3\delta) < 0.0111$, its membership is negligible. So, you can set the value of the minimum membership to $\mu\min = 0.0111$.

To design a fuzzification system that can be used in a variety of optimization problems in the logic of uncertainty, you need to sort the objective function values according to certain decrement rules and then convert the real-type function values into natural numbers. Regarding the integer value of the total number of populations S, the result of the transformation is used as the input value of the uncertainty logic. In order to make the ordering of the function values directly proportional to the membership degree, usually the maximum membership value is $\mu\max = 1.0$ in the best position, so that the minimum membership degree is $\mu\min = 0.0111$ in the worst position, and the rest position $\mu < 1.0$ can pass formulas (2) and (3) for indication:

$$\mu_i = \mu\max - \frac{s - I_i}{s - I}(\mu\max - \mu\min), \quad i = 1, 2, 3, \ldots, s, \tag{2}$$

$$\mu_{ij} = \mathrm{rand}(\mu_i, 1), \quad j = 1, 2, \ldots, D. \tag{3}$$

In formulas (2), (3), i is defined as the value of the objective function; that is, μ_i is the membership degree of i; the search space is set to J dimension, the membership value μ_{ij} of i in the space is I; I is the function value of the population. The descending order of size is the order value of $x(t)$ at this time; D is the largest dimension in the space. In equation (3), μ_i, 1 is a random real number in $\mathrm{rand}(\mu_i, 1)$. The condition through uncertainty reasoning, that is, equations (2) and (3), can be used to determine the membership value from the latter, and then according to the behavior part of the uncertainty reasoning.

Equation (4) gives the step size:

$$a_{ij} = \delta_{ij}\sqrt{-\ln(\mu_{ij})}. \tag{4}$$

In equation (4), the parameter value of the Gaussian membership function is defined as the search step value of the J-dimensional search space, and the value can be determined byequations (5) and (6):

$$\delta_{ij} = w \cdot \mathrm{abs}(\vec{x}_{\min} - \vec{x}_{\max}), \tag{5}$$

$$w = \frac{T_{\max} - t}{T_{\max}}. \tag{6}$$

In equations (5) and (6), the position of the maximum $\vec{x}_{\min}$ and minimum function values $\vec{x}_{\max}$ in the same subgroup are set to, and defined as, the inertia weight, and

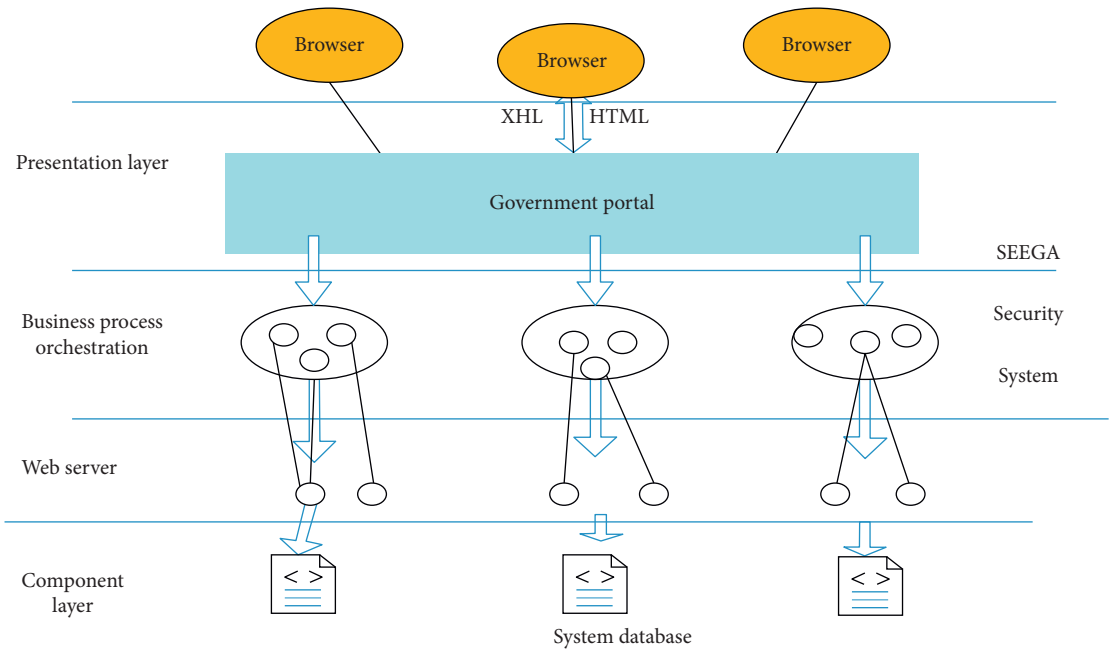


Figure 1: SOA based on the government system architecture.

their value linearly decreases with increasing the number of evolutions, and the decreasing interval is [0.1], 0.9]; and t represents the maximum number of iterations, and $T_{\max}$ represents the current number of iterations in turn; abs ( ) represents the absolute value of each input value.

2.2.2. Layers of SOA-Based e-Government Architecture. The SOA-based e-government architecture consists of the following layers:

(1) Data layer: the data layer is simply a popular database management system for storing various system parameters and official documents such as official documents used in the government system; it can also be a data set of legacy systems.

(2) Component layer: this layer consists of the enterprise components responsible for implementing the functionality. These special components are a managed and controlled collection of enterprise assets supported at the enterprise and business unit level. In most cases, this layer uses container-based technologies such as application servers that implement components, load balancing, high availability, and workload management [9].

(3) Service layer: the services selected by the business to support and disclose are at this level. They can be discovered or directly statically bound, then called, or, if possible, orchestrated into a composite service. The service disclosure layer also provides mechanisms for acquiring enterprise-wide components, business unit-specific components, and, in some cases, specific project formations and materializing their interface subsets in the form of service descriptions. Therefore, enterprise components use the functionality provided by their interfaces to provide service implementations at runtime. The interface at this level is exposed as a service description, in which they are exposed for use. They can exist independently or as a composite service.

(4) Business process orchestration layer: the composition and orchestration of services exposed in the service layer are defined in this layer. By coordinating and orchestrating, services are bundled into a single process and thus act together as a single application. These applications support special use cases and business processes [10].

(5) Presentation layer: in the framework of the presentation layer, the form of the presentation layer is mainly the government portal, and the user accesses the business process and specific functions of the government system through the government portal.

(6) Enterprise service bus (ESB): this layer enables services to be integrated and is often described as ESB. Web services introduce a set of reliable performance sets such as intelligent routing, protocol mediation, and other transformation mechanisms. The Description Language (WSDL) defines a binding that contains the address of the service. On the other hand, ESB provides a location-independent mechanism for integration [8].

(7) Security system: the security system uses the current relatively complete security technology to ensure the security of the entire government system, such as SSL and XML encryption, to ensure the security of the transport layer and the security of information exchange.

2.3. Using XML to Describe Data. According to the analysis in the previous section, in order to be able to realize the information interaction between the services in SBEGA, this information must be organized by an open, long-lived and widely accepted standard, and XML meets the above requirements. Data description specification, in SBEGA XML, will be used to describe the data that interacts between services. [11].

The Extensible Markup Language (XML) standard is a markup language for the text-based World Wide Web Consortium (W3 C) specification. Unlike HTML, which uses tags to describe appearance and data, XML strictly defines portable structured data. XML has some revolutionary features, making it the most fundamental and important technology in dynamic e-commerce. They are explained as follows:

(1) Scalable: the concept of XML is simple, but it has a wide range of applications, because anyone can define the corresponding XML tags to express the structure and type of information according to the needs of their own applications.

(2) Self-description: XML is a meta-language that can be interpreted and constrained by the definition of another XML document through an XML (previously through DTD, now XML Schema) definition.

(3) Information interoperability between heterogeneous systems: the XML format is simple to read and can represent any document material. As shown in Figure 1, as long as the system participating in the information exchange is equipped with an XML parser, it can interpret the information transmitted by other machines, thereby enabling communication of heterogeneous systems. Electronic Data Interchange (EDI) is a standard for electronic business documents, but it requires expensive software and a special network that is daunting for most companies. In contrast, XML-related software can be easily obtained, and it can use inexpensive, universal Internet transmission. [12]. The use of XML and the transmission of information and information will bring a revolutionary impact to e-commerce. XML transmission in heterogeneous systems is shown in Figure 2.

3. Construction of Data Platform Based on Brainlike Computing

3.1. Communication Mechanism Using SOAP Communication Protocol. The Simple Object Access Protocol (SOAP) is a lightweight, simple, XML-based protocol designed to interact with structured and solidified information on the Web. It can be used in conjunction with many existing Internet protocols and formats, including Hypertext Transfer Protocol (HTTP), Simple Mail Transfer Protocol (SMTP), and Multipurpose Internet Mail Extensions (MIME). It is independent of the underlying transport protocol, and the characteristics of adopting XML structured information are in line with the characteristics that SBEGA services may be heterogeneous, and interactive information is described by XML, which makes SOAP the first choice for communication protocols in SBEGA [13].

SOAP consists of four parts:

(1) SOAP encapsulation (envelop): encapsulation defines a framework that describes what is in the message, who sent it, who should accept it, and how to handle it.

(2) SOAP encoding rules are used to represent instances of the data types that the application needs to use.

(3) SOAP RPC representation is an agreement representing remote procedure calls and responses.

(4) SOAP bindings use the underlying protocol to exchange information.

Although these four parts are defined as a part of SOAP in a holistic manner, they are functionally intersecting, independent, and special. Envelope and encoding rules are defined in different XML namespaces. This makes the definition simpler. All SOAP information is an XML document. It must have a SOAP envelope and an information body, as well as an optional SOAP header. As shown, each element has a corresponding set of special rules. SOAP application system client and server communication model is shown in Figure 3.

3.2. SBEGA Service Mechanism. Web services are a new generation of Web applications, a type of software component that can be accessed through open Internet protocols. They represent a combination of component technology and Web technology. The traditional services based on CORBA, DCOM, and other proprietary protocols obviously do not conform to the openness of SBEGA architecture. The web services standard has been recognized by major platform providers and system service providers and is becoming the de facto service implementation standard. The use of web services to implement services in SBEGA is a natural choice [14].

In the web services invocation framework, clients and services communicate via standard Web protocols (HTTP or SMTP) and XML, which are completely platform-independent and implementation-oriented. Web services solve the main problems of government's government activities through the Internet, such as interoperability and integration. At the same time, web services have open and easy-to-implement standards, which enable various heterogeneous legacy applications within the government to be integrated. The architecture of web services is a layered structure, which can be divided into network layer, message delivery layer, service description layer, service discovery layer, and workflow layer. Its security, management, and quality of service run through the entire architecture of the web service [15].

The network transport layer is the basis of the Web protocol stack. To be called by the service requester, the web services must pass through the network. The standard network protocol used by web services in the Internet environment is generally HTTP, and SMTP and FTP protocols are also supported in some extension areas.

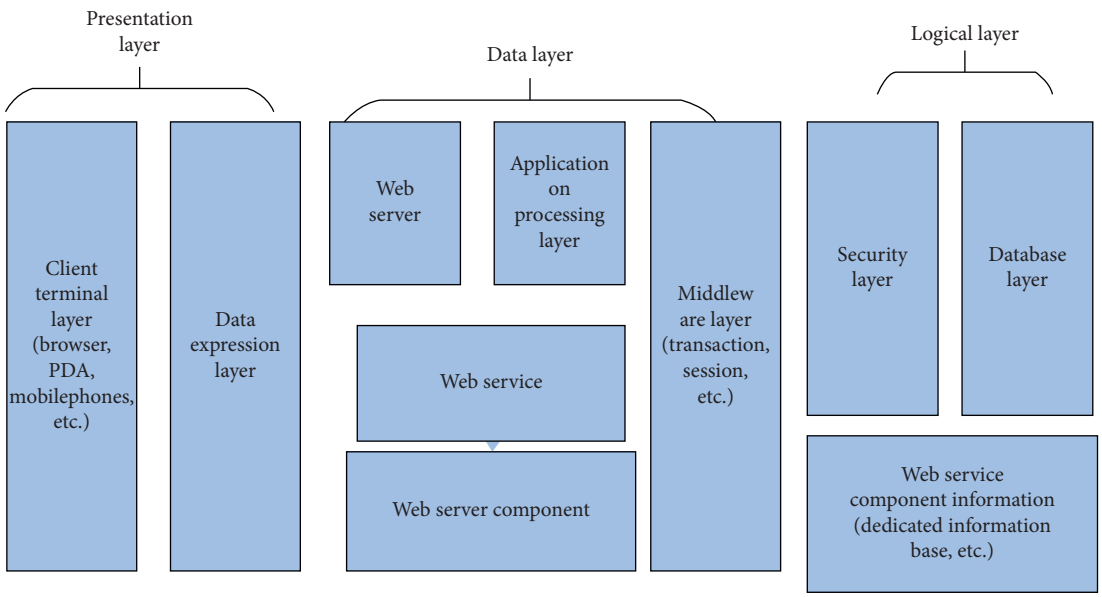


Figure 2: XML transmission in heterogeneous systems.

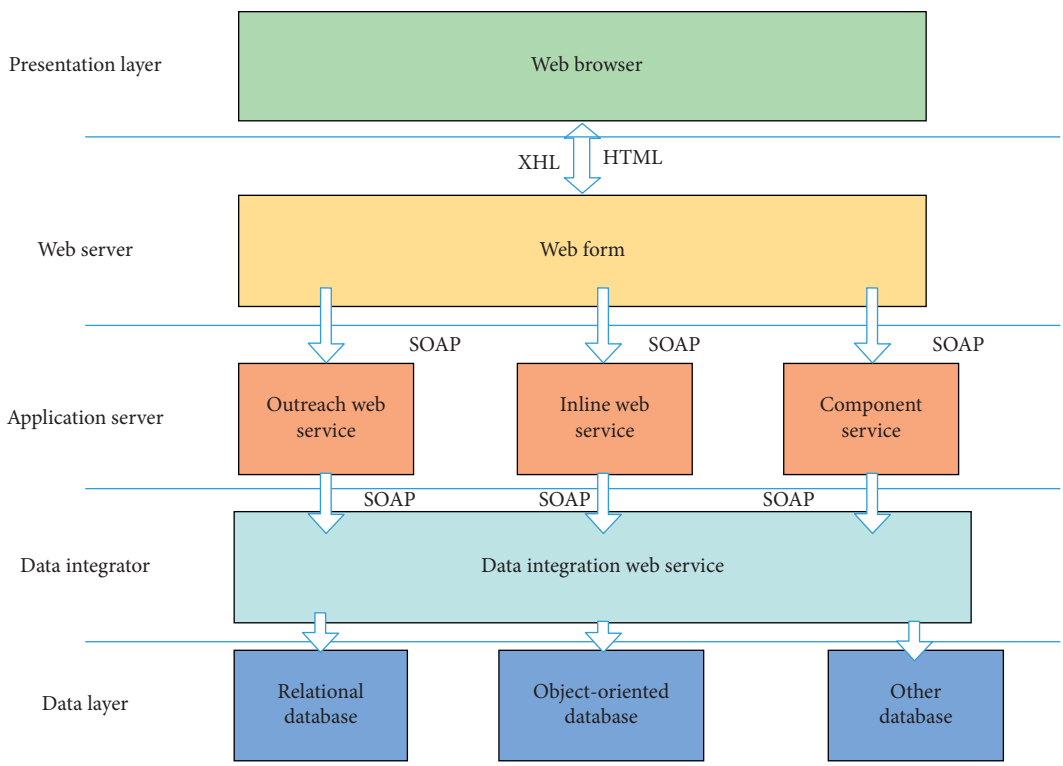


Figure 3: SOAP application system client and server communication model.

The message service layer uses the XML-based messaging protocol SOAP, which is built on top of the lower transport layer, which means that SOAP can be used alone or in conjunction with any transport protocol. All SOAP messages support the web services architecture, operations such as publishing, binding, and finding.

3.2.1. Service Description Layer. WSDL is a basic standard for defining the implementation and interface of services based on XML format. That is to say, WSDL divides the description of the service into two parts: the service implementation and the service interface.

3.2.2. Service Publishing Layer. The service provider can send the WSDL document directly to the service client, and the service provider can also publish the WSDL document to a local WSDL registry, or a public/private UDDI registry. Service clients can use these registration libraries to obtain

WSDL documents. UDDI provides a common set of SOAP APIs that implement a service registration method. The UDDI specification facilitates the creation, description, discovery, and integration of web services.

3.2.3. Service Discovery Layer. Service discovery is based on service release. If web services do not or cannot be published, then they cannot be discovered.

3.2.4. Service Workflow Layer. Web services Workflow Language (WSFL) is the standard for the service workflow layer at the top of the protocol stack. Unlike other standards in the protocol stack, WSFL is aimed at business process modeling and workflow.

3.3. SBEGA's Safety System. Through the security requirements of SBEGA, as well as the analysis of transport layer security and information interaction security, it can be seen that the transport layer security and XML security meet the different security requirements of the system, as shown in Table 1. The most important part of SBEGA's security system is to ensure the security of information exchange and transmission between different services in the system. From the table, we can see that if one of the security technologies is used alone, the security requirements of such information cannot be guaranteed. Only by combining them according to their characteristics, SBEGA's security system is established, and the functions provided by the transport layer security and XML security are fully utilized to cooperate with each other to meet the security requirements of SBEGA.

The SBEGA security system combines transport layer security and XML security to achieve the security requirements for SOAP messages passed between services in SBEGA, as shown in Figure 3. The requester uses SSL or HTTP basic authentication combined with SSL to complete the sender's authentication and uses the session key determined by the SSL negotiation to send the encrypted message to the receiver, thus establishing security between the sender and the receiver. XML digital signatures and XML encryption work at the transport layer security to provide a higher level of data confidentiality and data integrity. The nonrepudiation of the message can be provided by combining sender authentication with SSL [14]. SBEGA's security system is shown in Figure 4.

4. SBEGA Application in e-Government Platform

Over time, governments often have multiple heterogeneous independent systems based on different platforms and architectures, operating in different functional departments. With the increasing number of interactive businesses of citizens, businesses, and governments in the future, government departments need to continuously expand their systems, which in the long run will bring great inconvenience to the maintenance of the system in the future. Moreover, the subsystems of various departments are not connected, and there is almost no way for them to cooperate with each other. Therefore, information islands are often formed within the government system. Therefore, to get rid of this situation and ensure decentralized information resources under the premise of ensuring information security, the unity, integration, and management have become a key to improving the application level of e-government. At the same time, with the changes in policies, the departmental integration and adjustment of government agencies are frequent, and the business processes of the government system need to be constantly changing in order to cope with such policy changes. Process reorganization is an extremely frequent matter and requires collaborative work between multiple departments [15]. Application systems between multiple departments must agree on message formats, communication protocols, data models, etc., using standardized docking methods, so as to effectively realize the interaction of dynamic e-government modes. This requires a unified, standard-based technology, SOA architecture can well meet this need, and the system can be divided into service components according to different functions, through the UDDI registry, the public portal and the backstage of each department. The service is bound and called. Therefore, this chapter will design an e-government approval system model based on the e-government platform based on SOA architecture to realize a "one-stop," flexible and scalable e-government approval system.

4.1. e-Government Approval System Architecture. From the various advantages of SOA mentioned above, from the perspective of business and IT, service-oriented architecture is an architecture that can well adapt and solve the current e-government malpractices. The use of this structure to build an e-government approval system can solve the "information island" in the e-government field. It is the interconnection and interoperability of subsystems of various departments, integrating existing partner systems, and the rapid changes of corresponding business services to meet the needs of the public. The architecture of the e-government approval system model for service-oriented architecture is shown in Figure 5 [16].

According to the network division of each part of the model, we describe the various parts of the model as follows.

4.1.1. External Network. The model in the external network mainly runs on the Internet, providing services to the public and government office workers through Internet access and Web technologies. The one-stop office platform is mainly for internal government personnel, providing the entrances of various departments to their respective approval interfaces. It can manage the approval process and realize online approval, joint approval, handling supervision, license printing, statistical analysis, data query, and information feedback. Wait for the energy, to achieve online "one-stop" service [17]. On the public network, you can carry out project consultation, application form download, online submission application, approval result release, online information exchange, identity authentication, online complaints, and other functions [18].

Table 1: Analysis of transport layer security and XML security.

Security requirements	SSL	XML encryption	XML digital signature	Firewall
Confidentiality	Point-to-point confidentiality	End-to-end machine confidential		
Authentication		Send/receive	Verified	
Integrity	Using MAC	Using MAC and digital sign	ID authentication	Name authentication
Authorization				Authorize
Nonrepudiation			Signature	

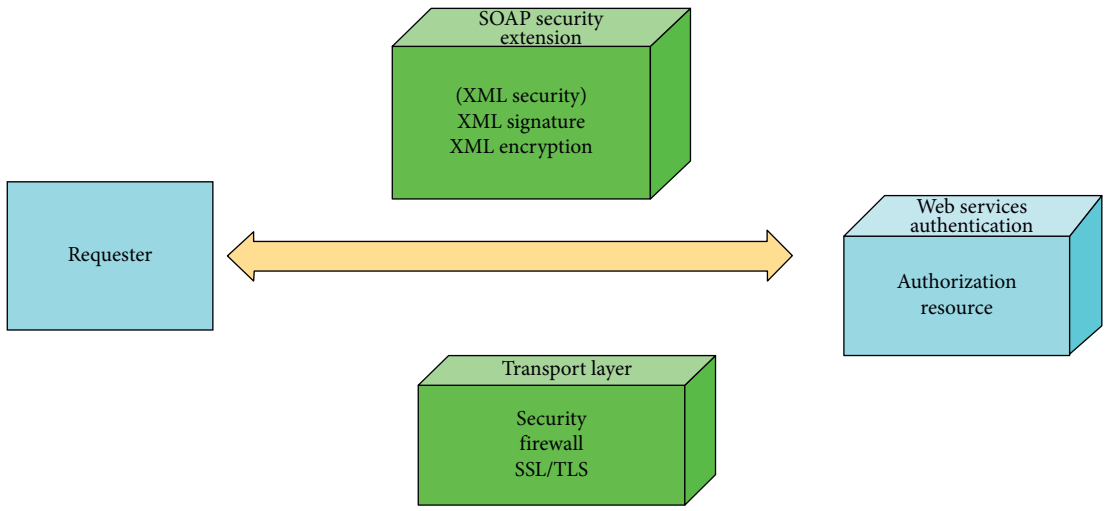


Figure 4: SBEGA's security system.

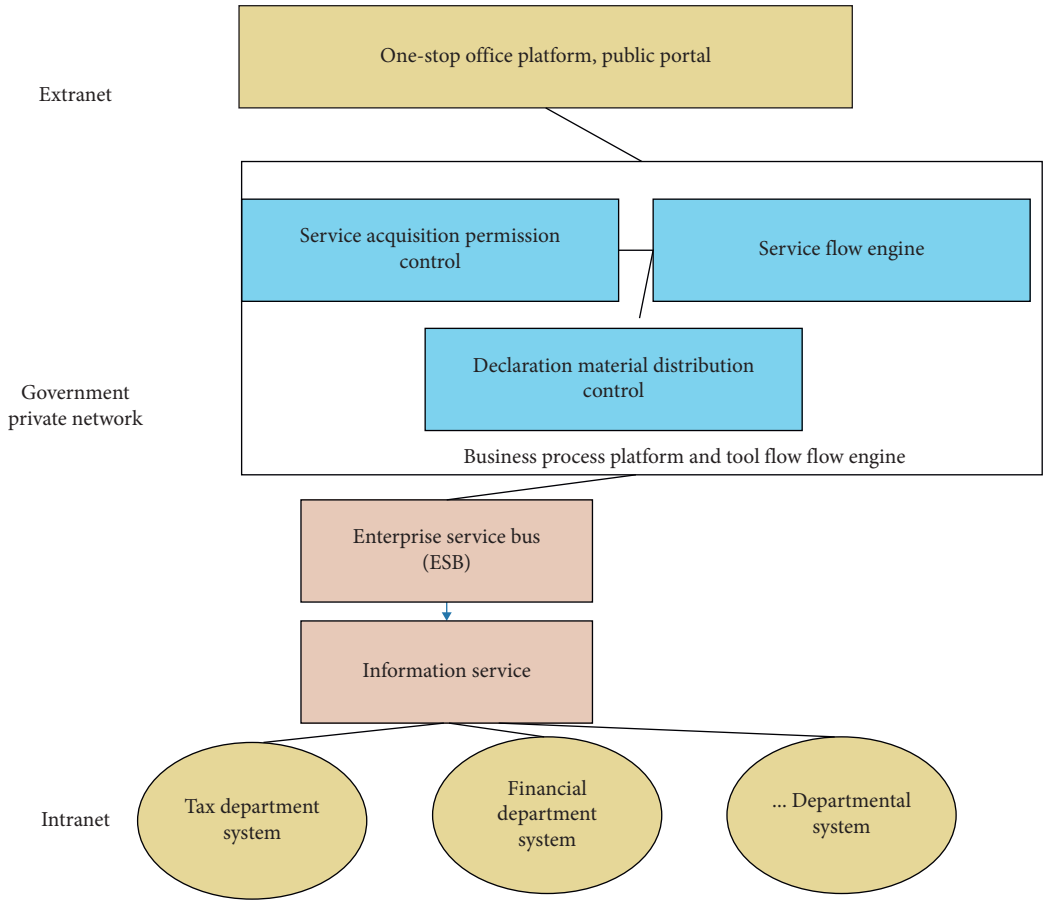


Figure 5: Architecture of the e-government approval system model for service-oriented architecture.

4.1.2. Government Special Network. The enterprise applying for approval will register the identity through the public portal. This registration process will call to check whether the enterprise is legal. The service checks whether the enterprise has legal status. After the registration is successful, it enters the approval resource library, and the enterprise user submits all the application items. After the data, the service flow engine automatically distributes it to the systems of different approval departments for review based on the nature of the data. After the approval personnel of various government departments enter the unified office platform, the system automatically pushes the approval part of each department to use the event service to the individual's desk [19]. The approval results of each part are automatically stored in the approval resource database, and the system monitors the time limit for approval. If the timeout is to be exceeded, the calendar reminder service or the SMS notification service in the tool set is called to automatically remind the approver [14].

Each department registers its service description information in the government approval UDDI and obtains services through the search for services. Some services have certain confidentiality. Therefore, not all new systems can call the registered services arbitrarily. The service gets access to the new control component to get access to it.

4.1.3. Intranet. The subsystems of each department run on the intranet, package the shared functions of each subsystem into an information service, provide the required information externally, or build new web services according to requirements, and process the shared services. Orchestrate, build new applications, recreate new services when the approval requirements change, and reengineer the original process layout, thereby improving the software reuse rate and shortening the software delivery cycle.

4.2. System Design. From the analysis of the previous section, the functional requirements of the market detection and control office can be abstracted into statistical analysis of enterprise data reporting and reported data. In order to realize a flexible, loose and scalable e-government system, the SOA-based e-government system architecture (SBEGA) proposed in this paper is used to implement the case. The architecture of the case e-government system is shown in Figure 5. In the architecture shown, the main functional modules of the case system are as follows [15].

4.2.1. User Management Service. The service provides external system user management through the web services interface. It has two components: enterprise user management and department internal user management. Their respective functions are Enterprise User Management Components.

4.2.2. Data Reporting Service. The data reporting service provides data services to enterprises and also supports the exchange of data between the future case system and other government departments' e-government systems. In the future, if the e-government system of other government departments needs to transmit the enterprise report data to the case system, the data transmission can be performed by defining the SOAP message format that satisfies the service interface requirements and accessing the service.

The service consists of data verification, conflict checking, data storage processing, and other components. These components perform data validity verification on the data transmitted by the service interface, check whether it conflicts with historical data, and perform data storage processing.

4.2.3. Data Statistics Service. The data statistics service is a service provided by the internal management personnel to collect statistics on the data reported by the enterprise. Like the data reporting service, the data statistics service also supports the exchange of data with the e-government systems of other government departments in the future. In the future, if the e-government system of other government departments needs to obtain the statistical results reported by the enterprise from the case system, it only needs to access the standard interface provided by the service and parse the data returned by the interface according to the rules [20].

4.3. Service Implementation. The service is the core of SBEGA. This section starts from this core and takes the user management service in the case system as an example to introduce the implementation of the service in the e-government system built by SBEGA. At the end, the remote access and function replacement of the service are carried out.

4.3.1. Service Realization. The case system uses the open source organization Axis 2 developed by Apache as the SOAP server running by web services. Axis 2 is newly written based on the new architecture and does not use the common code of Axis 1. Compared with Axis 1, it is more modular, more flexible, and more efficient in architecture. Using Axis 2 as the running platform of web services effectively reduces the difficulty of web services development and deployment. The process of developing web service based on Axis 2 is as follows [12].

(a) Write Service Core Business Implementation. In SBEGA, the service is the interface for the core business process to be developed externally. Take the user management service as an example. The core business is to add, update, and query enterprise users and internal users. To complete these services, many components and classes may be involved. However, users do not want to open all interfaces as services. In order to maintain the future maintainability of the system,

we design a special service class, in which all operations are open to the service, as shown in the code snippet above.

```
Package com.seagull.services;
Import’’’’‘;
Public class UserManage{
 Public OMElement adduser (OMElement userdata) {
Return resp;
Public OMElement updateuser (OMElement userdata) {
Return resp;
Public OMElement getuser (OMElement searchdata) {
Return resp;
```

(b) Write a Deployment Descriptor. After the core business is implemented, it is published as a service to the Axis 2 server, and a deployment description must be written for these implementations. The deployment descriptor is a standard XML file, and the above user-managed deployment descriptor document service.xml is as follows.

```
<service name = “UserManage”>
  <description>
    This service is to manage users in Seagull.
  </description>
  <parameter name = “ServiceClass“
      Locked = “false” > com.
seagull. Service s.UserManage</parameter>
  <operation name = “adduser”>
    <messageReceiver
  Class = “org.apache.axis2.receivers.Raw
XMLINOutMessageReceiver”/>
  </operation>
  <operation name = “updateuser”>
    <messageReceiver
  Class = “org.apache.axis2.receivers.Raw
XMLINOutMessageReceiver”/>
  </operation>
  <operation name = “getuser”>
    <messageReceiver
  Class = “org.apache.axis2.receivers.Raw
XMLINOutMessageReceiver”/>
  </operation>
</service>
```

4.3.2. Package Service. According to Apache’s specifications, the Axis 2 service is packaged as Axis Archive (.aar). This is a JAR file (created using the } ar or zip program), and the deployment descriptor document services.xml file is packaged in the 1VIETA-INF directory of this JAR archive.

4.3.3. Deployment Services. Deploying services in Axis 2 is as simple as copying the .as: file into the axis2/WEB-INF/services directory in the axis container’s axis 2 web application. Another way to deploy a service is to use the Upload Service tool in the Axis 2 Management Console. Access the Axis 2 management console via http://://localhost: <port}/axis2 and go to the administration page via the Administration link. Enter your username and password to log in. In the Tools section, select the Upload Service link, select the service to be deployed, package the .aar file, and then click Upload to complete the service deployment. This method is used to open source services on remote Axis 2 servers.

4.4. Government Affairs Platform Simulation Experiment. Public value evaluation and prediction: based on the above discovery model, customer value evaluation and customer prediction are conducted. First, the average model value of each type of customer is calculated as the classification basis of the four types of customer. The predicted customer and average model values are calculated in a fuzzy way to achieve the prediction purpose. In the government affairs system, the relevant information of customers is extracted at will and converted into data according to quantitative standards, as shown in Table 2. The distribution of customer information is shown in Figure 6.

Take the customer classification model of $\lambda = 0.40$ and find the average index of each model, as shown in Table 3. The average index distribution of customer value is shown in Figure 7.

Comparing the current value and potential value of customers, it can be seen that the potential value of bank customers finally converting into the value of return value is an uncertain amount. Therefore, the selection of customer value should be more focused on the current value, so this paper selects the value after the current value is added as the basis for clustering customer value. After addition, Class A customers are high-value customers, Class D customers are potential value customers, Class C customers are general value customers, and Class B customers are low-value customers.

Table 2: Customer information.

Classification object	Overdraft	Feedback personnel	Invalid feedback	Overall customer status	Personal response fact	Customer recommendation number	Number of replies	Random testing	The number of responses	Required and unpurchased business volume
Forecast customers	2015	25630	200	48	50	6	56000	75633	5000	7

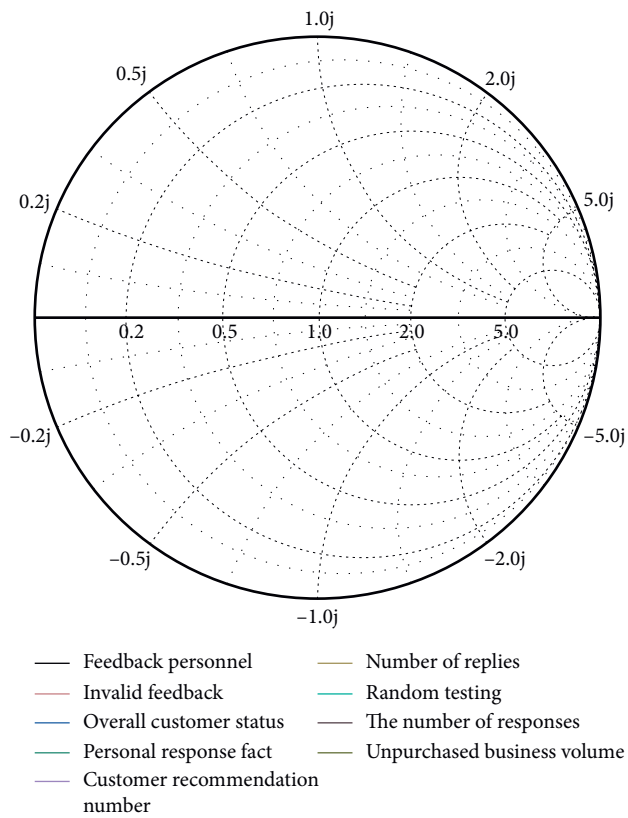


Figure 6: Customer information distribution map.

Table 3: Average indicator of customer value.

Classification object	Overdraft	Feedback personnel	Invalid feedback	Overall customer status	Personal response fact	Customer recommendation number	Number of replies	Random testing	The number of responses	Required and unpurchased business volume
A class customer	3001	16211	18	83	50	9	150000	200000	30000	2
B class customer	43567	0	9673	47	20	4	60000	3000	60000	1
C class customer	411	67327.50	27980	56.5	20	22.5	26732	2716.5	100000	3
D class customer	3267.65	41786.42	2788.58	54	41.15	4.42	117368.65	15739.35	63076.92	2.85

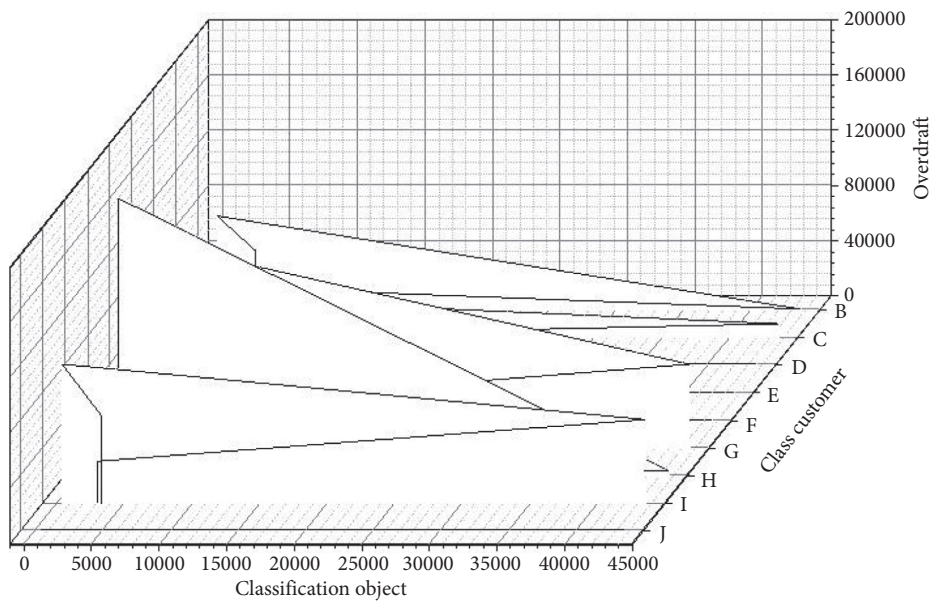


Figure 7: Average indicator distribution of customer value.

5. Conclusion

Big data is considered a diverse form of unstructured data, which need proper analysis and management, and it is excessive technological revolution after the Internet, the Internet of Things, and cloud computing. This paper analyzes the current popular SOA and web services technologies, illustrates the feasibility of their use in e-government systems, and proposes SOA-Based e-Government Architecture (SBEGA). SBEGA's basic mechanism, like other information systems, e-government networks, and information, is also subject to security threats. SBEGA uses an open architecture, which is subject to more security threats. This paper analyzes SBEGA's security objectives. The security system of SBEGA was put forward. Finally, SBEGA was successfully applied to the actual case e-government system, focusing on the system architecture design and representative service realization of the case, and testing the scalability of the case. Through the practice of the case system, SBEGA can meet the requirements of the current e-government system's openness and interconnection and provide a safe and flexible system architecture for building a dynamic and scalable e-government system, for the future construction of similar e-government systems, with reference and promotion value.

Data Availability

The datasets used and/or analyzed during the current study are available from the corresponding author upon reasonable request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

[1] L. Ying, "Design and development of public affairs government management system based on Internet," *Advanced Materials Research*, vol. 9, no. 12, pp. 1552–1555, 2014.

[2] L. Xuesi, "Based on the experience of E-government affairs how to design electronic port in China," *Applied Mechanics and Materials*, vol. 6, no. 3, pp. 1539–1542, 2014.

[3] W. Xingbiao and Y. Yiming, "Research on standardized management in government affairs microblog," *Enterprise Science and Technology and Development*, vol. 8, no. 8, pp. 859–733, 2013.

[4] J. Li, Z. Zheng, Q. Tian, G. Zhang, F. Zheng, and Y. Pan, "Research on tridiagonal matrix solver design based on a combination of processors," *Computers and Electrical Engineering*, vol. 62, pp. 1–16, 2017.

[5] X. Li, Z. Lv, Z. Zheng, C. Zhong, I. H. Hijazi, and S. Cheng, "Assessment of lively street network based on geographic information system and space syntax," *Multimedia Tools and Applications*, vol. 76, no. 17, pp. 17801–17819, 2017.

[6] S.-X. Pan, W.-J. Sun, and Z. Zheng, "Video segmentation algorithm based on superpixel link weight model," *Multimedia Tools and Applications*, vol. 76, no. 19, pp. 19741–19760, 2017.

[7] M. Chahal, S. Harit, K. K. Mishra, A. K. Sangaiah, and Z. Zheng, "A Survey on software-defined networking in vehicular ad hoc networks: challenges, applications and use cases," *Sustainable Cities and Society*, vol. 35, no. 11, pp. 830–840, 2017.

[8] T. Yang, X. Long, A. K. Sangaiah, Z. Zheng, and C. Tong, "Deep detection network for real-life traffic sign in vehicular networks," *Computer Networks*, vol. 136, no. 8, pp. 95–104, 2018.

[9] Y. Sun, C. Li, G. Li et al., "Gesture recognition based on Kinect and sEMG signal fusion," *Mobile Networks and Applications*, vol. 23, no. 4, pp. 797–805, 2018.

[10] C. Hamer, "Reinforcing the NPT," *Australian Journal of International Affairs*, vol. 65, no. 5, pp. 578–589, 2011.

[11] C. Li-Jun, C. T. Co, and L. Y. Branch, "The design exploration of the government affairs cloud platform based on trusted computing at province class," *Software Guide*, 2018.

[12] L. I. Cheng-Lin and L. Liang, "Research and implementation of government affairs cooperation cloud platform based on electronic handwriting signature and mobile security," *Communications Technology*, 2018.

[13] J. Yiwei, "Information sharing platform in electronic government affairs," *Computer and Digital Engineering*, vol. 116, no. 1, pp. 143–167, 2009.

[14] C. Bingbing, "MOOCS curriculum model: contribution and dilemma," *Foreign language Audio-Visual Teaching*, no. 3, pp. 39-40, 2014.

[15] B. Yingcai, "A probe into the reform of college english teaching in art colleges and universities under the background of mousing course," *Journal of Kaifeng Institute of Education*, no. 9, pp. 41-42, 2015.

[16] Y. U. Miao, W. Yan-Zhang, and L. Ji-Shan, "Design on E-government affairs system platform supporting governmental affair process reconfigure," *Computer Integrated Manufacturing Systems*, vol. 10, no. 3, pp. 352–358, 2004.

[17] S. Shivani, S. Tiwari, K. K. Mishra, Z. Zheng, and A. K. Sangaiah, "Providing security and privacy to huge and vulnerable songs repository using visual cryptography," *Multimedia Tools and Applications*, vol. 77, no. 9, pp. 11101–11120, 2017.

[18] N. Kumar, S. Tiwari, Z. Zheng, K. Krishn, and A. Kumar Sangaiah, "An efficient and provably secure time-limited key management scheme for outsourced data," in *Concurrency and Computation: Practice and Experience*, Wiley, New York, NY, USA, 2018.

[19] M. Carman, M. Carman, N. Bonalumi et al., "ENA health care reform platform," *Journal of Emergency Nursing*, vol. 34, no. 5, pp. 392-393, 2008.

[20] R. C. Joseph and N. A. Johnson, "Big data and transformational government," *It Professional*, vol. 15, no. 6, pp. 43–48, 2013.

Hindawi
Scientific Programming
Volume 2021, Article ID 7252896, 11 pages
https://doi.org/10.1155/2021/7252896

Research Article

A Sports Training Video Classification Model Based on Deep Learning

Yunjun Xu

School of Physical Education, Shaoyang University, Shaoyang 422000, China

Correspondence should be addressed to Yunjun Xu; yunjun.xu@adamson.edu.ph

Received 9 April 2021; Revised 9 May 2021; Accepted 22 May 2021; Published 30 May 2021

Academic Editor: Shah Nazir

A sports training video classification model based on deep learning is studied for targeting low classification accuracy caused by the randomness of objective movement in sports training video. The camera calibration technology is used to restore the position of the target in the real three-dimensional space. After the camera calibration in the video, the sports training video is preprocessed. The input video segment is divided into equal length segments to obtain the subvideo segment. The motion vector field, brightness feature, color feature, and texture feature of the subvideo segment are extracted, and the extracted features are input into the AlexNet convolutional neural network. ReLU is used as the activation function in this convolutional neural network. Local response normalization is used to suppress and enhance the output of neurons to highlight the performance of useful information, so that the output classification results are more accurate. Event matching method is used to match the convolutional neural network output to complete the sports training video classification. The experimental results of the proposed study show that the model can effectively solve the problems of target moving randomness. The classification accuracy of sports training video is more than 99%, and the classification speed is faster which is shown from the results of the experiments.

1. Introduction

With the rapid development of multimedia technology, sports get unprecedented attention and development. The mainstream research work of sports training video includes field and ground wire detection, player detection, recognition and tracking, camera calibration, event detection, and video abstract extraction. The classification of sports training video based on semantic information refers to the use of machine vision technology to automatically identify the types of sports training on the field and give the recognition results by using a certain way of expression [1]. Due to the extensive influence of sports, the introduction of machine vision technology and machine learning technology in sports training video classification has great potential commercial application value.

At present, there are few researches on sports training video classification. Zhu et al. used Gaussian mixture model to achieve player detection. The multitarget tracking method based on support vector regression particle filter was used to extract the trajectory of players and football, and the interactive space-time information between players and football trajectory was used to achieve tactical behavior expression and recognition in football game. Niu et al. achieved camera calibration by detecting and tracking the ground wire in the video image and finally achieved tactical behavior expression and recognition by using the space-time trajectory information of the interaction between players and football in real space. Matej Perse et al. proposed a two-stage framework to realize the tactical behavior recognition in basketball games. In the first stage, players' trajectory is segmented according to the Gaussian mixture model under the generalized context information in basketball games. In the second stage, players' trajectory is semantically expressed according to the key information, and the tactical behavior recognition is realized by using the template matching method. Chen et al. designed an automatic recognition system, which realized camera calibration by field line detection, and realized attack and defense pattern recognition in basketball game by using player trajectory description in

the field. Masui et al. used background subtraction to detect players and then represented the spatial distribution of players in different areas of the field by using symbol system, to realize football tactical behavior recognition. This idea was a nontracking tactical behavior recognition method. The existing tactical behavior recognition mostly used the target trajectory as the underlying visual feature, which faced many problems. Firstly, due to the mutual occlusion between targets, the randomness of target movement, and the complexity of the environment background, there are still many problems in the accuracy and persistence of target tracking; secondly, because the sports training video is mainly based on long-distance view, the identification of players and balls is poor under complex lighting conditions.

Deep learning forms more abstract high-level features by combining low-level features to discover distributed features of data. The multilayer network structure of deep model can make the network learn the organization form of features by itself [2], and get the final semantic features through multiple abstractions. In 2006, Hinton et al. proposed the first feasible depth model. Since then, deep learning has become a new research field of machine learning, known as a revolutionary new technology in the field of artificial intelligence. Deep learning constructs multilayer network model and combines low-level features to form high-level semantic features with abstract representation, so as to simulate the way of thinking of human brain for perception and recognition. At present, deep learning has been widely used in speech, image, and other data recognition, detection and other fields, and has achieved remarkable results. Following are the main contributions of the study:

(i) To study the sports training video classification model based on deep learning

(ii) Establish the sports training video classification model by using convolution neural network of deep learning method

(iii) To verify the effectiveness of the proposed approach through experiments

2. Materials and Methods

2.1. Camera Calibration. Camera calibration technology is used to restore the position of the target in the real three-dimensional space. On this basis, the radial distortion and tangential distortion in the nonlinear model are fully considered, the Rodrigues rotation equation is used to reduce the number of optimization parameters, and the steepest descent method and LM optimization method are used to solve the accurate parameters, respectively.

Because the actual lens in the video is not ideal perspective imaging, with varying degrees of distortion, this kind of distortion can be divided into radial distortion and tangential distortion [3]. In order to describe the imaging model accurately, two parameters are used to describe the lens radial distortion and tangential distortion. The relationship between ideal coordinates and distortion parameters is as follows:

$$\begin{cases} x_d = x_u + \delta_x = x_u + k_1 r^2 x_u + k_2 r^4 x_u + k_3 r^6 x_u + 2k_4 x_u y_u + k_5\left(2x_u^2 + r^2\right), \\ y_d = y_u + \delta_y = y_u + k_1 r^2 y_u + k_2 r^4 y_u + k_3 r^6 y_u + k_4\left(2y_u^2 + r^2\right) + 2k_5 x_u y_u. \end{cases} \tag{1}$$

In (1), (x_u, y_u) is the normalized image coordinate calculated by the pinhole camera model; (x_d, y_d) is the image coordinate actually containing distortion; δ_x and δ_y are the nonlinear distortion values; $r^2 = x_u^2 + y_u^2$; k_1, k_2, k_3, k_4, and k_5 are the nonlinear distortion parameters, where k_1, k_2, and k_3 are the radial distortion coefficients, which will cause the radial movement of real image points on the image plane; k_4 and k_5 are the tangential distortion coefficients.

Given the initial parameters, to solve the precise camera parameters is essentially to solve the unconstrained multi-dimensional extremum problem. Because there is a deviation between the theoretical value of pixel coordinates and the measured value after the target feature points are projected to the image plane [4–6], the optimal estimation of camera parameters needs to meet the minimum deviation. According to the nonlinear optimization theory, the objective function is expressed as follows:

$$\min F(x) = \min \sum_{i=1}^{n} \sum_{j=1}^{p} \left| m_{ij} - \hat{m}_{ij}\left(A, k_1, k_2, k_3, k_4, k_5, R_i, t_i, M_j\right) \right|^2. \tag{2}$$

In (2), n is the number of target images captured by the camera under different viewing angles; p is the number of target feature points; m_{ij} is the observed value of the coordinate of the j-th feature point of the i-th target image; $\hat{m}_{ij}$ is the theoretical value of the projection point coordinate of the target feature point under the nonlinear model; M_j is the spatial coordinate of the j-th feature point on the target.

In the process of capturing the target from different angles, the internal parameters of the camera are regarded as constant, and the external parameters are different from each shooting angle. The number of optimized parameters increases significantly with the increase of the target image [7–9]. Rodrigues rotation equation provides a method of using vector to represent rotation. If the 3×3 rotation matrix with 9 elements is represented by 3 elements of a vector $r = \begin{bmatrix} r_x & r_y & r_z \end{bmatrix}$, the external parameters of each image are reduced to 6, which greatly reduces the amount of calculation in the optimization process.

The relationship between rotation matrix and rotation vector is as follows:

$$R = \cos\theta \cdot I + \sin\theta \cdot [z]_x + (1 - \cos\theta) \cdot rr^T. \quad (3)$$

The steepest descent method searches along the negative gradient direction of the objective function until it reaches the lowest point of the objective function. For unimodal function, it can quickly get the extreme point. This method uses the principle that the function value along the negative gradient direction of the initial point decreases continuously to search. For the initial point X_0 of function F, there are sequences X_0, X_1, and X_2, which satisfies the relationship as follows:

$$X_{n+1} = X_n - \lambda_n \nabla F(X_n), \quad n \geq 0. \quad (4)$$

The corresponding function values have the following relations:

$$F(X_0) \geq F(X_1) \geq F(X_2) \geq \cdots F(X_n). \quad (5)$$

Because the objective function has the form of minimal sum of squares and the coordinates of feature points on the target image are nonlinear functions of parameters to be estimated, it belongs to nonlinear least squares optimization problem. LM method can avoid the case that $A_k^T A_k$ is ill-conditioned matrix in least squares. In LM algorithm, the descent direction is given by the following equation:

$$d_k = -\left(A_k^T A_k + \alpha_k I\right)^{-1} A_k^T f_k. \quad (6)$$

Through the above process to restore the position of the target in the real three-dimensional space, the accuracy of sports training video classification is improved.

2.2. Video Preprocessing. Before classifying the sports training videos, it needs to firstly preprocess the sports training videos. Shooting video on sports training site is usually divided into distance video, medium distance video, and close distance video [10]. The proportion of sports training remote shooting is relatively large; remote shooting can effectively obtain the whole field information. $V = \{v_1, v_2, \ldots, v_i, \ldots, v_N\}$ is used to represent video input, where V represents the video segment corresponding to a specific sports event, v_i represents the video image of frame i, and $i = 1, 2, 3, \ldots, N$, N indicates the number of frames converted into video frame image of the input video segment.

In order to classify sports training videos more accurately, the input video segments are segmented according to equal length [11], and several subvideo segments are obtained. The expression is as follows:

$$\begin{aligned} V &= \{V_1, V_2, V_3, \ldots, v_M\}, \\ v_j &= \{v_{j1}, v_{j2}, \ldots, v_{jq}, \ldots, v_{jm}\}. \end{aligned} \quad (7)$$

In the above equation, $jm = pm$, $j \neq p$, $j, p = 1, 2, \ldots, M$, $q = 1, 2, \ldots, m.v_j$ represents the j-th subvideo segment after video segmentation, v_{jq} represents the q-th frame image in the j-th subvideo segment, and M represents the number of subvideo segments. After the above processing, the input and segmentation of the sports training video are completed, and the time span of the segmented video field has a certain impact on the classification results.

2.3. Feature Extraction

2.3.1. Extraction of Motion Vector Field.

(1) Let the size of the sports training video be $M \times N \times T$, $M \times N$ denote the resolution, and T denote the length of the video sequence. The video is divided into $K \times L$ blocks; each block size is $h \times v$, where $h = M/K$ and C denotes the number of blocks in each block.

(2) A rectangular coordinate system is established and the motion vector is mapped to this coordinate system [12]. The mapping diagram of the motion vector field of the rectangular coordinate system is shown in Figure 1.

In Figure 1, $MV(i, j)$ is the block with position (i, j), $\theta \in [0, 2\pi)$ is the direction of the motion vector C. If C_x is the component of the motion vector of the C-th block in the horizontal (x) direction, C_y is the component of the motion vector of the C-th block in the vertical (y) direction, and ρ is the motion intensity of the block C; then,

$$\begin{cases} \rho = \sqrt{c_x^2 + c_y^2}, \\ \sin(\theta) = \dfrac{c_y}{\rho}, \\ \tan(\theta) = \dfrac{c_y}{c_x}. \end{cases} \quad (8)$$

(3) The coordinate system of continuous video frames is arranged in chronological order [13], and it is divided into Q equal angle sectors along the positive x direction, p is quantized to R intervals, and then the histograms of p and θ are made, respectively, so it can obtain

$$\begin{cases} \mathrm{Hist}_q = \dfrac{1}{T} \displaystyle\sum_{t=1}^{T} \sum_{i=1}^{C} q_i^t, \quad q \in [1, Q], \\ \mathrm{Hist}_r = \dfrac{1}{q} \displaystyle\sum_{t=1}^{T} \sum_{i=1}^{C} r_i^t, \quad r \in [1, R]. \end{cases} \quad (9)$$

In (9), q_i^t represents the number of motion vectors in quadrant q in frame t, and r_i^t represents the number of p quantized to r in frame t.

(4) The expectation and variance of the motion vector in the x and y directions are used to evaluate the motion in the block, namely,

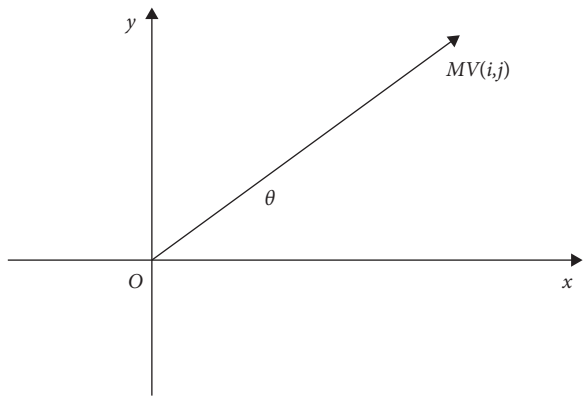


Figure 1: Diagram of motion vector field mapping.

$$\begin{cases} \mu_x = \dfrac{1}{T}\displaystyle\sum_{t=1}^{T}\sum_{i=1}^{C} C_{x,i}^{t}, \\ \mu_y = \dfrac{1}{T}\displaystyle\sum_{t=1}^{T}\sum_{i=1}^{C} C_{y,i}^{t}, \\ \sigma_x^2 = \dfrac{1}{T}\displaystyle\sum_{t=1}^{T}\sum_{i=1}^{C} \left(C_{x,i}^{t} - \mu_x\right)^2, \\ \sigma_y^2 = \dfrac{1}{T}\displaystyle\sum_{t=1}^{T}\sum_{i=1}^{C} \left(C_{y,i}^{t} - \mu_y\right)^2. \end{cases} \tag{10}$$

In (11), $C_{x,i}^{t}$ and $C_{y,i}^{t}$ represent the components of the motion vector of the i-th macroblock in the x and y directions in a frame, and μ_x, μ_y, σ_x^2, and σ_y^2 represent the expectation and variance of the motion vector of the macroblock in the x and y directions, respectively.

2.3.2. Extraction of Luminance Feature. Assuming that the frame resolution is $M \times N$, each frame is divided into $k \times k$ blocks, and the size of each block is $h \times v$, where $h = M/K$, $v = N/K$, x_i represents the brightness value of the i-th pixel in the block, and the average brightness value of each block is $\overline{X(l)}$, $l \in [1, k \times k]$, namely,

$$\overline{X(l)} = \sum_{i=1}^{h\times v} \frac{x_i}{(h \times v)}. \tag{11}$$

If y is used to represent the encoding value of the block luminance comparison, the encoding value of the luminance comparison result between the m-th block and the n-th block in the frame can be expressed by (12), where $l \le m \le k \times k$ and $2 \le n \le k \times k - 1$.

$$u = \begin{cases} 1, & \text{if } \overline{X(m)} > \overline{X(n)}, \\ 0, & \text{other}. \end{cases} \tag{12}$$

Through (12), the frames can be compared according to the average brightness of blocks and encoded with "1" and "0".

2.3.3. Color Feature Extraction. Assuming that the frame size is $M \times N$, the frame is converted into HSV model and divided into $k \times k$ blocks; each block size is $h \times v$, where $h = M/K$, $v = N/K$. $x_{i,m,n}$ represents the pixel value of the m component of the i-th pixel in the n-th block of the video, where $n \in [1, k \times k]$, $i \in [1, h \times v]$, and $m \in [H, S, V]$; then, the color characteristics of the sports training video are as follows:

$$\begin{aligned} \mu_{m,n} &= \frac{1}{h \times v}\sum_{i=1}^{h\times v} x_{i,m,n}, \\ \sigma_{m,n} &= \sqrt{\frac{1}{h \times v}\sum_{i=1}^{h\times v}\left(x_{i,m,n} - \mu_{m,n}\right)}, \\ S_{m,n} &= \sqrt[3]{\frac{1}{h \times v}\sum_{i=1}^{h\times v}\left(x_{i,m,n} - \mu_{m,n}\right)^3}. \end{aligned} \tag{13}$$

In the above equation, $\mu_{m,n}$, $x_{i,m,n}$, and $S_{m,n}$ respectively represent the mean value, variance, and third-order moment of m component in the n-th block.

2.3.4. Texture Feature Extraction. Let i have L gray levels in sports training video. G denotes a gray level cooccurrence matrix, and its element p_{ij} is the times of pixel pairs with gray level i and gray level j in i. p_{ij} is calculated as follows:

$$p_{ij} = N\{(x, y) | f(x, y) = i, f(x + \Delta x, y + \Delta y) = j\}, \tag{14}$$

where $f(x, y)$ is the gray level of the pixel (x, y), and Δx and Δy reflect the distance d and direction θ between the two points.

The most commonly used texture feature is used as the classification feature of sports video. The definition is as follows:

$$f_1 = \sum_{i=1}^{L}\sum_{j=1}^{L} \frac{p_{ij}}{1 + |i - j|}. \tag{15}$$

2.4. Sports Training Video Classification Model Based on Convolutional Neural Network

2.4.1. Neuron Layer Structure of Convolutional Neural Network. A convolutional neural network usually consists of multiple convolution layers, down sampling layers, and normalization layers. Finally, the two-dimensional feature map is connected into a vector and input to the final classifier through the fully connected layer to get the probability output.

(1) Convolution Layer. In a convolution layer, the features of the upper layer are convoluted by a learnable convolution kernel, and then the output features can be obtained through an activation function [14]. Each output may be combined to convolute the values of multiple inputs:

$$x_j^1 = f\left(\sum_{i\in M_j} x_i^{l-1} \cdot k_{ij}^{l-1} + b_j^1\right). \tag{16}$$

In the above equation, M_j represents the set of input features connected by a convolution kernel. M_j determines the connection between convolution kernel and input layer. The output feature map is obtained by convolution kernel of input feature map. Assuming that each convolution kernel extracts a pattern, each output feature map corresponds to a feature and each convolution kernel is equivalent to a feature map. This is because the convolution layer uses weight sharing technology; that is, each neuron uses the same convolution check input to do convolution and each neuron is only connected with some input neurons, which reduces the number of convolution layer parameters. Function f is the activation function of neurons, which is usually a nonlinear function.

The input of convolution layer is multiple two-dimensional planes, and each convolution core is connected with all input channels [15]. Convolution is performed in a three-dimensional space to obtain the position response output. Finally, the convolution checks the convolution of the whole input space to obtain a feature map. Usually, multiple convolution kernels are set in each convolution layer, and each convolution kernel extracts different features, so that each feature map represents the feature plane extracted by the corresponding convolution kernel.

(2) Down Sampling Layer. The purpose of the down sampling layer is to improve the robustness of the network to the small deformation of the input samples, so as to enhance the generalization performance of the network. y_{ijk} is used to represent the output of a neuron in the down sampling layer. The down sampling layer can be expressed as

$$y_{ijk} = \sum_{pq} w_{pq} x_{i,j+p,k+q}, \tag{17}$$

where w_{pq} is the normalized weighted window, which can make down sampling of every input feature map without crossing different feature maps. The number of output feature maps in the down sampling layer is the same as the number of input feature maps, which reduces the resolution of each feature map.

(3) Normalization Layer. The normalization layer is very important for improving the performance of neural network. In convolution neural network model, the normalization layer includes the normalization of the feature vector of the same feature map and the feature map located in different feature maps, which strengthens the feature map with higher response value, and drives different convolution kernels to learn different patterns [16, 17]. The subtraction and normalization operation at a given location are actually the value of the location minus the weighted value of each pixel in the neighborhood. The weight can be determined by a Gaussian weighted window. Division normalization is a common normalization algorithm, which can intensify the difference of response value and improve the effect of high characteristic of response value.

Local response normalization is a common normalization algorithm in convolutional networks. The response value can be expressed as

$$b_{x,y}^i = \frac{a_{x,y}^i}{\left(K + \alpha \sum_{j=\max(0,i-(n/2))}^{\min(N-1,i+(n/2))} \left(a_{(x,y)}^i\right)^2\right)^\beta}, \tag{18}$$

where $a_{x,y}^i$ represents the value of the i-th input feature map at the coordinate (x, y); N represents the number of input feature maps; n represents the normalization on the adjacent n maps.

The local response normalization layer contains three adjustable parameters, namely, the number of feature maps n and parameters α and β. All normalization layers adopt the same parameter setting, such that $n = 5$, $\alpha = 0.0005$, $\beta = 0.5$.

(4) Fully Connected Layer. The fully connected layer is usually at the top of the neural network, which forms a traditional multilayer perceptual network together with the decision-making layer to classify the features extracted from the convolution layer. The overfitting of convolutional neural network is mainly caused by more parameters in the fully connected layer. Dropout technology is usually added to the fully connected layer, and some neurons are randomly selected to participate in the training to prevent the network from overfitting.

A multilayer convolutional neural network is composed of the above five neuron layers, which perform different functions, respectively, and must be combined according to certain rules to achieve better results. Among the five neuron layers, only the convolution layer and the fully connected layer contain trainable parameters, and the convolution layer can retain the input spatial position information, which is required by the down sampling layer. The convolution layer is usually used alternately with the down sampling layer, so that different convolution layers can extract different scale features [18]. The fully connected layer will destroy the position information of feature planes and the difference between each feature plane. The fully connected layer is usually used as a part of the final multilayer perceptual classifier, which integrates the convolution layer and the down sampling layer to extract features and send them to the decision layer for classification.

2.4.2. Structure of Improved Convolutional Neural Network. The AlexNet convolutional neural network of deep learning is used to classify sports training videos. The AlexNet convolutional neural network consists of 23 layers, including five convolution layers and three fully connected layers.

(1) Use the New Activation Function ReLU. Generally, the activation function of artificial neuron is hyperbolic tangent function $f(x) = \tanh(x)$ or sigmoid function $f(x) = (1 + e^{-x})^{-1}$. In the experiment, it is found that when sigmoid or hyperbolic tangent function is used to calculate the error gradient by backpropagation, the derivation involves division, which leads to a large amount of calculation; once the number of layers of traditional neural network increases, the gradient fading problem occurs. The root cause is that when sigmoid or hyperbolic tangent function is used to calculate the error gradient by backpropagation, the change of function value slows down, and its derivative is close to zero, which makes other hidden layers far away from the output layer prone to gradient fading [19]; in addition, it is also a disadvantage of sigmoid function to add weight penalty factor to get sparsity and output nonzero mean value.

The advantages of ReLU function $f(x) = \max(0, x)$ are as follows: first, the calculation speed and convergence speed are faster; second, ReLU will make the output 0 when $x < 0$, resulting in network sparsity, reducing the interdependence of parameters, and alleviating the over fitting problem; third, its derivation is piecewise linear in both forward and backward propagation, avoiding the disappearance of gradient.

(2) Local Response Normalization (LRN). In neurobiology, there is a concept called "lateral inhibition", which refers to the ability of excited neurons to inhibit their adjacent neurons. That is to highlight the maximum peak in the local sensing area and increase the ability of biological perception.

It is in the neural network that the LRN layer realizes "lateral inhibition". Let a^i_{xy} be the activation value of neurons at position (x, y) of the i-th kernel function and b^i_{xy} be the activation value after normalization, and the total number of kernel functions is N; then, the mathematical model of LRN is expressed as follows:

$$b^i_{xy} = \frac{a^i_{xy}}{\left(k + \alpha \sum_{j=\max(0,i-(1/2))}^{\min(N-1,i+(n/2))} \left(a^i_{xy}\right)^2\right)^{\beta}}, \tag{19}$$

where the sum operation is normalized at the adjacent position of n around (x, y), and the super parameters k, n, α, and β need to be determined by the verification set. It is very effective to add LRN layer after using ReLU function as the activation function. The ReLU function has unlimited activation ability when $x > 0$, which needs LRN normalization. It is expected that the LRN layer can detect the features with high frequency and amplify them by suppressing the peripheral neurons; the LRN layer will suppress the uniform response in any given local neighborhood; that is, if all the values are large, then the normalization will suppress all the values uniformly. The purpose of LRN layer is to make useful information more prominent by inhibiting and enhancing neuron output.

2.5. Event Matching. Based on the output of convolutional neural network, the events of sports training test video sequence and reference video sequence are matched by event matching method. Given L_1 observation symbols of video class, a multistate traversed convolutional neural network model is trained by using features extracted from sports training video frames, to obtain the event sequence (event probability and corresponding state transition) in the corresponding reference video. The reference event sequence is used to create a dictionary for a given sports training event [20]. For the event with a specific state transition (k, l) in the reference event, the probability distribution of the event is approximated by a Gaussian density function $N(\mu_{kl}, \sigma_{kl})$, where μ_{kl} and σ_{kl} represent the mean value and variance of the density function, respectively. It is given by the following equation:

$$\begin{aligned} \mu_{kl} &= \frac{1}{L_1} \sum_{t=1}^{L_1} e^p_t(k, l), \\ \sigma_{kl} &= \sqrt{\frac{1}{L_1} \sum_{t=1}^{L_1} \left(e^p_t(k \cdot l) - \mu_{kl}\right)^2}. \end{aligned} \tag{20}$$

Each state transition is assigned a mean value and variance to represent the probability $e^p_t(k \cdot l)$ of the event occurring in the category. For the sports training video clips that do not appear in the training stage, a reference convolution neural network model is used to obtain the events. Let $e^p_t(k \cdot l)$ denote the event probability of state transition (k, l) at time t when the test sequence in the observation symbol provides a reference model. Let L_2 denote the number of observation symbols in the test sequence. The similarity between the test video clip and the reference model is expressed by the following equation:

$$s = \frac{1}{L_2} \sum_{t=1}^{L_2} \frac{1}{\sqrt{2\pi}\sigma_{kl}} \exp\left[-\frac{\left(\hat{e}^t_p(k, l) - \mu_{kl}\right)^2}{2\sigma^2_{kl}}\right]. \tag{21}$$

The similarity value s between video clips and all kinds of sports training is compared, and they are classified into the category with the highest similarity value.

3. Results and Discussion

In order to verify the feasibility and effectiveness of the sports training video classification model, eight data sets which are often used in the classification research in the network are selected as the test objects. The data sets include eight types of sports training videos, such as basketball, volleyball, and football. The detailed contents of the videos in each data set are shown in Table 1.

Table 1 shows that the experimental data set contains many types of sports training videos. Different sizes and types of sports training videos are used to test the classification performance of different models of sports training

Table 1: Details of the experimental data sets.

Data set name	Video content	Video frames/N	Duration/s	Size/MB
DataSetA	Badminton, basketball, table tennis	1582	66	4.52
DataSetB	Tennis, volleyball	1351	56	3.64
DataSetC	football, running	1254	52	2.85
DataSetD	Snooker, tennis	3269	136	7.48
DataSetE	Basketball, tennis	5642	235	9.52
DataSetF	Table tennis, volleyball	1478	62	3.48
DataSetG	Basketball, football	3151	131	8.64
DataSetH	Running, basketball	1856	77	4.85

videos. Support vector machine model and HMM model are selected as comparison models.

Three models are used to classify the sports training videos of 8 data sets, and the classification results are shown in Table 2.

The experimental results in Table 2 show that the classification of sports training videos can be realized by using the proposed model. The classification results of sports training videos by using the proposed model are similar to those of actual sports training videos, which indicates that this model has high classification performance of sports training videos.

In the result of sports training video classification of the proposed model, two images are randomly intercepted in basketball training video, as shown in Figure 2.

As can be seen from the experimental results in Figure 2, using the proposed model to classify basketball training videos can accurately classify videos according to the extracted features of sports training videos, and randomly intercepted pictures are all accurate for basketball training, which verifies that the proposed model has high classification effectiveness of sports training videos.

The classification accuracy, recall rate, and precision rate are selected as the important indexes to evaluate the classification performance of the proposed model. n_c is used to represent the number of correct recognition results, n_m is used to represent the number of wrong recognition results, and n_f is used to represent the number of failed recognition results. In order to effectively reduce the error caused by a single experiment, the average value of five experiments is selected, and 2 : 1 ratio is set to randomly divide training samples and test samples. The evaluation index equation is as follows:

$$\begin{aligned} \text{recall ratio} &= \frac{n_c}{(n_c + n_m)}, \\ \text{precision ratio} &= \frac{n_c}{(n_c + n_f)}. \end{aligned} \tag{22}$$

Statistics of the accuracy comparison results of different data sets and different types of sports training video classification are shown in Figure 3.

As can be seen from the experimental results in Figure 3, under different data sets and different types of sports training, the classification accuracy of sports training video classified by the proposed model is higher than 99%, and the classification accuracy of sports training video classified by this model is significantly higher than that of the other two models, which effectively verifies that this model has higher classification accuracy of sports training video.

Statistics of the recall rate comparison results of different data sets and different types of sports training video classification are shown in Figure 4.

As can be seen from the experimental results in Figure 4, under different data sets and different types of sports training, the recall rate of sports training videos classified by the proposed model is higher than 98.5%, and the recall rate of sports training videos classified by this model is significantly higher than that of the other two models, which verifies that this model has higher classification accuracy of sports training videos.

Statistics of the precision rate comparison results of different data sets and different types of sports training video classification are shown in Figure 5.

As can be seen from the experimental results in Figure 5, under different data sets and different types of sports training, the precision rate of sports training video classification using the proposed model is higher than 98%, and the precision rate of sports training video classification using the proposed model is significantly higher than that of the other two models, which verifies the high accuracy of sports training video classification using the proposed model.

The analysis of the above experimental results shows that the classification accuracy, recall rate, and precision rate of different data sets and different types of sports training videos are the best. Basketball and football sports training videos have strong continuity and change more frequently, so more quantitative features are needed to better obtain the change features in videos. Basketball and volleyball videos usually have close range images, while baseball and tennis videos are mostly shot from a long-distance perspective, so feature extraction is difficult. Football is also shot from a long-distance perspective; it has continuous movement in the field and can be well collected by increasing the number of states. In football and basketball videos, it most uses a single camera to track players or regions of interest; unlike other sports training, switching between multiple cameras frequently is conducive to event detection. The model can effectively improve the randomness of target movement and improve the classification accuracy by extracting video features.

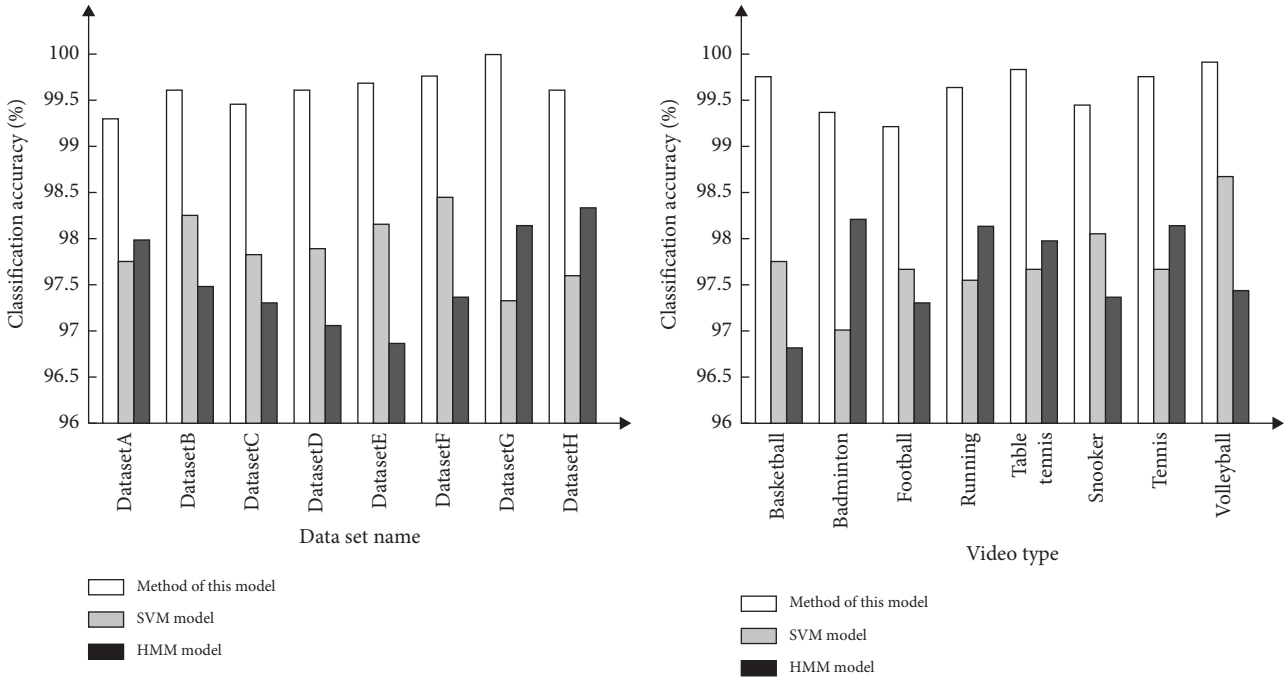
The training time and test time of sports training videos classified by three models with different data sets are counted. The comparison results are shown in Table 3.

Table 2: Classification results of sports training videos.

Types	Actual number of frames/N	Method of this model/N	SVM model/N	HMM model/N
Basketball	1481	1473	1352	1376
Badminton	2384	2415	2384	2584
Football	3436	3418	3364	3468
Running	2564	2542	2498	2348
Table tennis	1765	1759	1743	1842
Snooker	3652	3627	3584	3452
Tennis	2755	2711	2684	2679
Volleyball	1546	1638	1724	1711
Total	19583	19583	19333	19460

(a) (b)

Figure 2: Basketball training classification results.



(a) (b)

Figure 3: Comparison of classification accuracy. (a) Different data sets. (b) Different types of sports training videos.

The experimental results in Table 3 show that the classification speed of sports training video using the proposed model is the fastest, and the accurate classification results of sports training video can be obtained by using shorter training time and test time of the proposed model, which verifies that this model has higher classification efficiency of sports training video.

The above experimental results show that the proposed model can accurately classify all kinds of sports training videos, which shows that this model has good

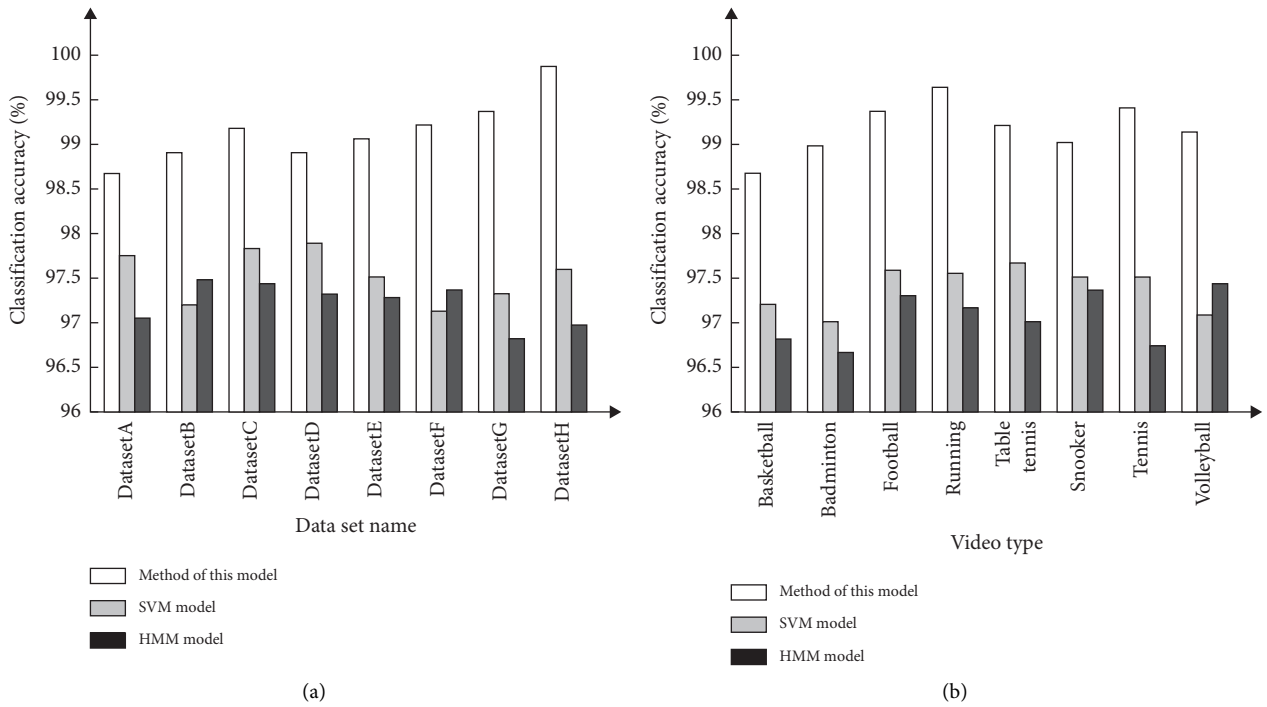


Figure 4: Comparison of classification recall rates. (a) Different data sets. (b) Different types of sports training videos.

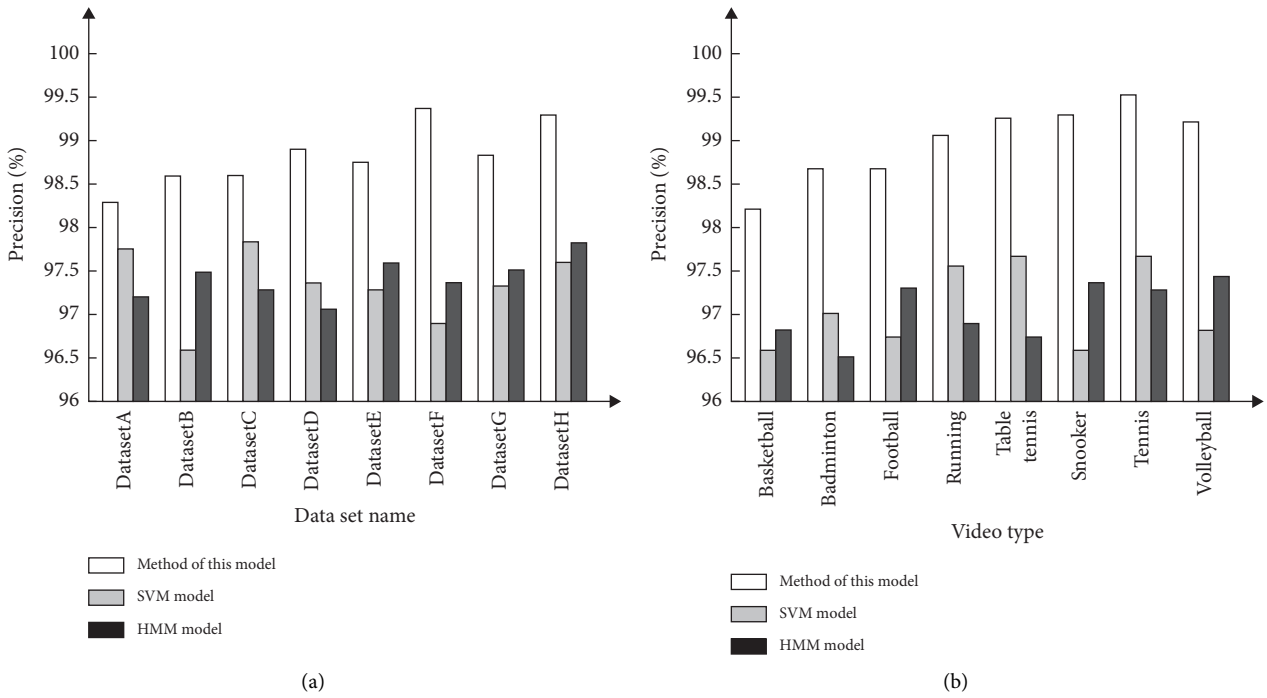


Figure 5: Comparison of classification accuracy. (a) Different data sets. (b) Different types of sports training videos.

classification performance. The main reason is that this model uses deep learning model to establish classification model, which can effectively improve the classification accuracy of sports training videos. For close range videos with similar categories, it still has high classification accuracy. This model has high accuracy and comprehensive performance in classifying all kinds of sports training videos.

Table 3: Comparison of classification speed.

Data set name	Method of this model/ms		SVM model/ms		HMM model/ms	
	Training time	Testing time	Training time	Testing time	Training time	Testing time
DataSetA	152	215	864	856	1165	658
DataSetB	352	234	915	485	1248	594
DataSetC	425	152	1052	715	1352	605
DataSetD	356	236	1135	359	1426	736
DataSetE	412	245	1658	645	1489	852
DataSetF	385	236	2354	597	2654	945
DataSetG	642	215	1842	612	2854	439
DataSetH	531	284	2358	784	1696	896

4. Conclusion

At present and with the passage of time, the amount of sports training video data in the Internet is growing rapidly. In order to effectively manage and retrieve sports training video, accurate classification of sports training video is very important for consideration. Aiming at the shortcomings of existing approaches of sports training video classification, this paper establishes sports training video classification model based on deep learning method. Convolution neural network with deep learning is used for the classification purpose in the proposed research. After classification, event matching operation is performed, and video classification is realized according to similarity. The experimental results show that the proposed model can effectively determine all kinds of sports training videos and accurately detect the occurrence of events through convolution neural network, so as to achieve high-precision classification of sports training videos. Compared with other models, the proposed model has the advantages of simple implementation, fast processing speed, high classification accuracy, high generalization ability, and adaptability.

Data Availability

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

Conflicts of Interest

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

References

[1] C. Y. Fang, K. B. Jia, and P. Y. Liu, "Identification of taxi violation behavior based on surveillance video," *Computer Simulation*, vol. 37, no. 05, pp. 331–336, 2020.

[2] Y. Kumar, M. Sheoran, G. Jajoo, and S. K. Yadav, "Automatic modulation classification based on constellation density using deep learning," *IEEE Communications Letters*, vol. 99, pp. 1–11, 2020.

[3] Q. Li, S. Zheng, Y. Huang, and D. Liu, "Automatic classification of nvst short-exposure data based on deep learning," *Publications of the Astronomical Society of the Pacific*, vol. 133, no. 1020, Article ID 024505, 2021.

[4] A. S. Garea, D. B. Heras, and F. Argüello, "Caffe cnn-based classification of hyperspectral images on gpu," *The Journal of Supercomputing*, vol. 75, no. 3, pp. 1065–1077, 2019.

[5] M. Oleynik, A. Kugic, Z. Kasáč, and M. Kreuzthaler, "Evaluating shallow and deep learning strategies for the 2018 n2c2 shared task on clinical text classification," *Journal of the American Medical Informatics Association*, vol. 26, no. 11, pp. 1247–1254, 2019.

[6] S. Park, W. K. Chung, and K. Kim, "Training-free bayesian self-adaptive classification for semg pattern recognition including motion transition," *IEEE Transactions on Biomedical Engineering*, vol. 67, no. 6, pp. 1775–1786, 2020.

[7] M. Chen, J. Cao, B. Zhang, H. Zhu, and X. Cai, "Lbp-05-classification and mutation prediction based on liver cancer hisopathological images using deep learning," *Journal of Hepatology*, vol. 70, no. 1, pp. e142–e143, 2019.

[8] T. Zebin and S. Rezvy, "Covid-19 detection and disease progression visualization: deep learning on chest x-rays for classification and coarse localization," *Applied Intelligence*, vol. 51, no. 2, pp. 1010–1021, 2021.

[9] C. Ajmi, J. Zapata, J. J. Martínez-Álvarez, G. Doménech, and R. Ruiz, "Using deep learning for defect classification on a small weld x-ray image dataset," *Journal of Nondestructive Evaluation*, vol. 39, no. 3, pp. 68–72, 2020.

[10] G. Murtaza, L. Shuib, A. W. Abdul Wahab et al., "Deep learning-based breast cancer classification through medical imaging modalities: state of the art and research challenges," *Artificial Intelligence Review*, vol. 53, no. 3, pp. 1655–1720, 2020.

[11] X. Zhang, J. Li, Z. Cai, L. Zhang, Z. Chen, and C. Liu, "Overfitting suppression training strategies for deep learning-based atrial fibrillation detection," *Medical & Biological Engineering & Computing*, vol. 59, no. 1, pp. 165–173, 2021.

[12] Y. Wang, J. Wang, W. Zhang, J. Yang, and G. Gui, "Deep learning-based cooperative automatic modulation classification method for mimo systems," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 4, pp. 4575–4579, 2020.

[13] F. Wang, H. Huang, and J. Liu, "Variational based mixed noise removal with CNN deep learning regularization," *IEEE Transactions on Image Processing: A Publication of the IEEE Signal Processing Society*, vol. 1, no. 99, pp. 1–9, 2019.

[14] D. A. Duev, A. Mahabal, F. J. Masci et al., "Real-bogus classification for the zwicky transient facility using deep learning," *Monthly Notices of the Royal Astronomical Society*, vol. 489, no. 3, pp. 3582–3590, 2019.

[15] R. K. Samala, H.-P. Chan, L. Hadjiiski, M. A. Helvie, C. D. Richter, and K. H. Cha, "Breast cancer diagnosis in digital breast tomosynthesis: effects of training sample size on multi-stage transfer learning using deep neural nets," *IEEE*

Transactions on Medical Imaging, vol. 38, no. 3, pp. 686–696, 2019.

[16] M. Shah, A. Roomans Ledo, and J. Rittscher, "Automated classification of normal and stargardt disease optical coherence tomography images using deep learning," *Acta Ophthalmologica*, vol. 98, no. 6, pp. e715–e721, 2020.

[17] Y. Liang, Z. Hu, and K. Li, "Power consumption model based on feature selection and deep learning in cloud computing scenarios," *IET Communications*, vol. 14, no. 10, pp. 1610–1618, 2020.

[18] N. Maffulli and F. Oliva, "Coper classification early after acl rupture changes with progressive neuromuscular and strength training and is associated with 2-year success: letter to the editor," *The American Journal of Sports Medicine*, vol. 47, no. 11, pp. NP64–NP65, 2019.

[19] L. R. Dugas, C. R. Labella, N. Alawad, J. Pasulka, and N. Jayanthi, "Benefits and challenges of serial sports training risk assessment and counselling in kids: the t.r.a.c.k. randomised intervention study," *British Journal of Sports Medicine*, vol. 53, no. 4, pp. 243–251, 2019.

[20] J. W. Orchard, W. Meeuwisse, W. Derman, M. Hgglund, and R. Bahr, "Sport medicine diagnostic coding system (smdsc) and the orchard sports injury and illness classification system (osiics): revised 2020 consensus versions," *British Journal of Sports Medicine*, vol. 54, no. 7, Article ID 101921, 2020.

Hindawi
Scientific Programming
Volume 2021, Article ID 5510039, 6 pages
https://doi.org/10.1155/2021/5510039

Research Article

Study on the Intentional Choice Mechanism of Course Selection Based on Swarm Intelligence Algorithm

Dawei Zhang and Daling Wang

School of Computer Science and Engineering, Northeastern University, Shenyang 110819, China

Correspondence should be addressed to Daling Wang; zhangdw@stumail.neu.edu.cn

Received 26 February 2021; Accepted 22 May 2021; Published 30 May 2021

Academic Editor: Shah Nazir

With the passage of time and recent advances in science and information technology, the development in the area of course selection based on some defined criteria has made the choice of the mechanism easy and effective. In this paper, the approach is based on an innovative perspective, and swarm intelligence was introduced for the intentional choice mechanism of course selection. The study has considered the course selection in English as an example. Swarm intelligence algorithm and integrative course selection were combined with the recommendation algorithm and the intent of course selection in English course to discuss the relevant decision mechanism. Firstly, the comprehensive selection intentional recommendation algorithm and PSO algorithm were introduced, and the algorithm was initialized. Secondly, the operation process was described in detail, and the application process was analyzed. Then, it was introduced into the English course elective process. Finally, the experimental results of the study came with the conclusion through the test that the PSO algorithm has a higher degree of accuracy and can better judge individual behaviors, which contributes to the establishment of choice-choice mechanism. The effectiveness of the study was demonstrated through experiments.

1. Introduction

With the continuous deepening of economic globalization, the society has become increasingly demanding on college students' English proficiency. It not only requires students to have a certain level of English basic knowledge but also requires students must have strong English comprehensive ability and cross-cultural communication skills [1]. This puts forward new requirements for college English teaching reform. At present, college English is influenced by traditional utilitarianism. Students' motivation to learn English has a great deviation. And, English has become a "tool" for employment [2]. In our country, higher education adopts compulsory and elective methods in English education. In a certain sense, the quality of elective courses can better reflect the perfection of the credit system because the credit system is based on the elective system [3]. However, with the reform of the curriculum in our country, the college entrance examination has canceled English subjects, which is still facing a crisis of lack of elective intention in higher education. The original compulsory high school English has also become an elective course. There is a lack of normative and curriculum diversity in the opening of English courses, and students lack guidance for elective courses. Meanwhile, elective course teaching and students' incorrect understanding of the relationship between elective courses and college entrance exam need to be improved [4]. Therefore, studying the intention of elective courses in English courses can help students to strengthen their enthusiasm for learning English courses. In-depth understanding is inadequate concerning the students' influence on the choice for taking a particular course. Imperfect understanding of modality choice has significant implications to institutions and students.

The swarm intelligence evolutionary algorithm, a kind of optimal algorithm, has attracted more attention of researchers. Both artificial life and EA are correlated intensively in evolutionary strategy, especially in the domain of genetic algorithms [5]. There are mainly two kinds of algorithms in the field of swarm optimal theory. One is ant swarm algorithm and the other is particle swarm optimization. The second one originated from the simulation of simple social systems, and it was originally the process of

foraging for flocks. However, later, it was found that it is a good optimization tool [6]. The first one (PSO) was proposed by scholars in recent years [7]. The PSO algorithm and the simulated annealing algorithm are similar as much as possible and both of them are evolutionary algorithms. PSO algorithm utilizes a random solution to start the iteration and then circles the algorithm to obtain the optimal solution. It uses the fitness function to assess the effectiveness of the solution, there are no operations such as "crossover" and "mutation," and it is easier than the rules of GA. The important thing is that it gains the global optimal by following the optimal value of the current search. For its advantages of easy implementation, high precision, and fast convergence, the algorithm has attracted more attention of the academic community, and it has demonstrated its superiority in solving practical problems [8]. PSO can be calculated parallelly. With the continuous progress of the innovation and reform process of English major teaching, the traditional single teaching model has long been unsuitable for today's students. Those boring and old-fashioned presentation methods make students feel that they can only passively accept and are far from interactive fun [9]. Therefore, it is important to improve interactivity and understand the individual conditions in teaching reform and innovation.

The contribution of the proposed research is devise swarm intelligence for the intentional choice mechanism of course selection. For the course selection, English was considered as an example of the study. Swarm intelligence algorithm and integrative course selection were integrated with the recommendation algorithm and the intent of course selection in English course for discussing the applicable mechanism of decision. The effectiveness of the study was demonstrated through experiments.

2. Methodology

2.1. Particle Swarm Algorithm Content Interpretation. The location of the Locator in the space is as shown in the dark rectangular position in the figure. And, their coordinates are $L0\,(l_{0x}, l_{0y}, l_{0z})$, $L1\,(l_{1x}, l_{1y}, l_{1z})$, $L2\,(l_{2x}, l_{2y}, l_{2z})$, and $L3\,(l_{3x}, l_{3y}, l_{3z})$. During the experiment, the four Locators emitted ultrasonic waves to the Tag. When the ultrasonic wave detected the Tag, it was reflected back to the Locator by the Tag. The propagation time of the ultrasonic wave was monitored and the TOA method was used to calculate the distance from each Locator to the Tag. Based on this initial condition, the following will begin to introduce the PSO algorithm [10]. Within the problem set is placed a particle group that consists of four particles: P_0, P_1, P_2, and P_3. The initial position of these particles is randomly placed. Follow the steps below. Firstly, ultrasonic sensors and communication between tags were used to prove the existence of ultrasonic transmission cycle. Secondly, the TOA method was used to find the distance between the Locator and the Tag [11]. The operation is as follows. The distance from Locator $L0$ to Tag is denoted as $M_{0,0}$. Distance from Locator $L1$ to Tag is denoted as$M_{1,0}$. Distance from Locator $L2$ to Tag is denoted as $M_{2,0}$. The distance from Locator $L3$ to Tag is denoted as $M_{3,0}$. Then, under the initial conditions, the coordinates of the four Locator and the four particles are also known conditions. Therefore, the distance between Locator and the particles can be calculated separately, such as the distance between Locator $L0$ and Particle P_0 [12], as follows:

$$D_{0,0} = \sqrt{(l_{0x} - p_{0x})^2 + (l_{0y} - p_{0y})^2 + (l_{0z} - p_{0z})^2}. \quad (1)$$

The distance between Locator $L1$ and Particle P_0 can be calculated as follows:

$$D_{1,0} = \sqrt{(l_{1x} - p_{0x})^2 + (l_{1y} - p_{0y})^2 + (l_{1z} - p_{0z})^2}. \quad (2)$$

The distance between Locator $L2$ and Particle P_0 can be calculated as follows:

$$D_{2,0} = \sqrt{(l_{2x} - p_{0x})^2 + (l_{2y} - p_{0y})^2 + (l_{2z} - p_{0z})^2}. \quad (3)$$

The distance between Locator $L3$ and Particle P_0 can be calculated as follows:

$$D_{3,0} = \sqrt{(l_{3x} - p_{0x})^2 + (l_{3y} - p_{0y})^2 + (l_{3z} - p_{0z})^2}. \quad (4)$$

According to the above formula, the distance between the Locator and the other three particles can be expressed as follows: $D_{0,1}$, $D_{1,1}$, $D_{2,1}$, and $D_{3,1}$ represent the distances between $L0$ and P_1, $L1$ and P_1, $L2$ and P_1, and $L3$ and P_1, respectively, $D_{0,2}$, $D_{1,2}$, $D_{2,2}$, and $D_{3,2}$ represent the distances between L0 and P_2, $L1$ and P_2, $L2$ and P_2, and $L3$ and P_2, respectively, and $D_{0,3}$, $D_{1,3}$, $D_{2,3}$, and $D_{3,3}$ represent the distances between $L0$ and P_3, $L1$ and P_3, $L2$ and P_3, and $L3$ and P_3, respectively. Then, the distance between the Tag and the Locator and the particle and the Locator was found. Combining these two conditions with the following formula, it is possible to find the particles with the closest Tag among the four particles [13]:

$$\begin{aligned}
f_0 &= (D_{0,0} - M_{0,0})^2 + (D_{1,0} - M_{1,0})^2 + (D_{2,0} - M_{2,0})^2 + (D_{3,0} - M_{3,0})^2,\\
f_1 &= (D_{0,1} - M_{0,0})^2 + (D_{1,1} - M_{1,0})^2 + (D_{2,1} - M_{2,0})^2 + (D_{3,1} - M_{3,0})^2,\\
f_2 &= (D_{0,2} - M_{0,0})^2 + (D_{1,2} - M_{1,0})^2 + (D_{2,2} - M_{2,0})^2 + (D_{3,2} - M_{3,0})^2,\\
f_3 &= (D_{0,3} - M_{0,0})^2 + (D_{1,3} - M_{1,0})^2 + (D_{2,3} - M_{2,0})^2 + (D_{3,3} - M_{3,0})^2.
\end{aligned} \quad (5)$$

Among them, $D_{0,0}$, $D_{0,1}$, and $D_{0,2}$ represent the distance between particles and Locator. $M_{0,0}$, $M_{1,0}$, $M_{2,0}$, and $M_{3,0}$ represent the distance from the Tag to the Locator. f_0, f_1, f_2, and f_3 are called distance degrees, whose size indicates the distance between the particle and the Tag. It can be seen from the formula that the smaller the distance, the smaller the distance between the particle and the tag. So [14], among the four particles, the particle with the smallest distance is calculated to be closest to the Tag. For example, if the value of f_1 is the smallest, then it can be determined that particle P_1 is closest to Tag. After this aspect, faced with the first key point in the PSO algorithm, it can be effectively processed clearly defined in the particle swarm that which one is the closest particle separated from the tag [15].

2.2. Spark Cluster Iteration Calculation. After narrating the particle swarm algorithm, in this paper, the distributed platform-Spark Cluster Particle Swarm Optimization was used. With the advantages of speed, ease of use, and sophisticated analysis, Apache Spark, as a computation platform, can be used to handle big data. It started up originally in 2009 and opened in 2010 greatly. Frankly speaking, Spark widely extends the MapRed and Ce models to hold with various kinds of calculations. Speed always plays an important role on large data sets when processing the interactive queries and stream data. Spark calculates in memory specifically [16]. In addition, Spark can perform complex computation on disk. In generally, Spark was proposed to handle various computation situations, such as batch processing applications, pass generation algorithms, interactive queries, and stream [17]. The versatility of Spark not only enables simple and convenient processing in different application scenarios but also reduces the administrative burden. Spark provides many version interfaces such as Python, Java, Scala, and SQL and provides a rich set of default tool libraries. Spark was also combined with other tools slightly. The driver program running on the master node controls the critical flow of the program [18]. The driver program defines the operations such as map, reduce, and filter. Figure 1 is a working principle diagram [19].

Then, we will elaborate on the specific implementation of the algorithm in the program. From the analysis above, firstly, it can be seen that, in the Spark application, the data file was read from a document teaching platform (such as HDFS). Secondly, the elastic distribution data set (RDD) was set up. Thirdly, the driver program was used to parallelize RDDs and assign them to various points. If the RDD is frequently reused in the application, it can perform well in cache. When RDD is filed, it is possible to perform parallel operations on the RDD. The new algorithm operation used in Spark cluster environment implementation is shown in Table 1 [20].

As you can see from the above table, Spark provides rich functions. The Spark ecosystem consists of a general execution module, a structured data module, a stream analysis module, a machine learning module, and a graph calculation module [21]. The first one is the execution system of the platform and is the core of functions. SparkCore supports cache capabilities, a common execution model, and application programming interfaces for Java, Scala, and Python, which allows Spark to efficiently calculate and stand for a wide range of applications. Spark processes structured data by QL. It provides DataFrames, program abstraction, and acts as a distributed SQL query engine. SparkSQL enables native Hive queries in Hadoop clusters to be up to 100 times faster than existing configurations and datasets. At the same time, it has strong integration capabilities with other modules in the Spark ecosystem. The stream analysis module allows for strong interactivity and analysis applications for stream data mining and historical data, while continuing Spark's ease of use and fault tolerance. Spark Streaming easily integrates with all types of common data. Machine learning module (MLLIB) is a scalable machine learning library that provides both high quality and efficiency algorithms. The MLlib library can be used as part of Spark applications in languages such as Java, Scala, and Python.

3. Results and Discussion

3.1. Particle Swarm Algorithm Validity Test. After calculating the three ideas of the ant colony algorithm to select courses for English courses, the three curriculum selection intentions were expanded to Spark cluster environment to improve the efficiency and expansibility of the elective courses. These experiments can be used to identify the performance of our proposal in this paper. In this paper, all experiments were done in a lab Spark cluster environment. The data sources are based on the original foursquare check-in data set, and the data sets with corresponding numbers are gained by replication. The experiment of verifying the efficiency and scalability of large data volume is performed on this case [22].

By repeating these experiments to modify these parameters of the influence factors of the following factors in the linear combination and probability fusion, for the linear method, the result of the selection intention preference recommendation is the highest under the values of 0.4 and 0.5. For the second one, the order is in accordance with social factors, time factors, and geographical factors. Besides, the first rough and then fine-grained result were made in order, and the effect of the selection intention preference is highest.

In the probabilistic fusion recommendation method, when the values of λ and δ are 0.2 and 0.4, respectively, the result of the selection intention preference recommendation is the highest. Among the three influencing factors, which influence the choice of course intention, the degree of influence of geographical factors and time factors is greater than that of social factors, and the degree of influence of geographical factors is stronger than that of time. Then, we compare the results of our proposal and others to validate the advantages, and it can be shown in Figure 2.

The experimental results show that the introduction of social factors, geographical factors, and time factors can hardly recommend the intention of elective courses, which verifies the conclusion of the existing research results. The three integrated methods proposed in this paper further

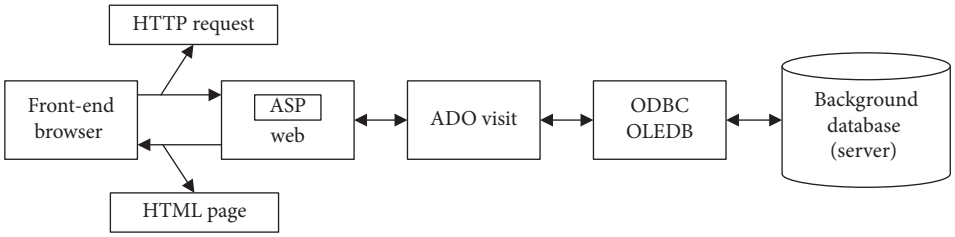


Figure 1: Schematic diagram of the Web database based on ASP access.

Table 1: RDD processing detail.

RDD operation	Description
map	Enables each element in RDD to be processed by a user-defined function; as a result, each function in RDD is the result of user-defined functions on each element in the original RDD
groupByKey	The aggregation operation is performed according to the key values of each element in RDD; when the RDD is composed of elements of (K, V) type, the aggregation operation results in each element in RDD is (K, Sequence (V)) type
collect	All elements in RDD are sent to the driver program, when the user needs to collect the results at the Master node

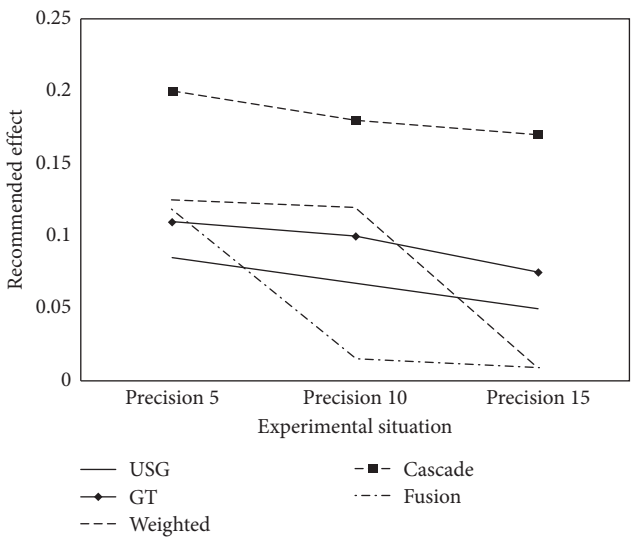


Figure 2: Results of tests: precision.

improve the effectiveness of the results of recommendation of choice intention. Specifically, we found that the linear-weighted recommendation effect was better than GT, indicating that social factors could hardly choose the outcome of recommendation intention. However, the recommendation effect of GT is better than that of T and G, indicating that geographical factors and time factors can also enhance the recommendation effect. Then, all push methods are analyzed. The results show that the linear weighting method has the highest F1_measure and can provide the most vivid comprehensive recommendation effect. In addition, the RECAU of EL combination method is relatively low, while the accuracy is the highest among all methods, and it is suitable for applications requiring high accuracy. The recommended results of the probabilistic fusion method are second only to the linear weighting method. And, all results are better than others.

3.2. Verifying the Preference of Particle Swarm Optimization. In this experiment, the efficiencies of the particle swarm algorithm for testing stand-alone environments and Spark cluster environments were compared. The extended four-square check-in data set was used to gradually add the data volume from lG to 32G, and then, the execution time was observed in the two environments, respectively. The comparison results of the three comprehensive recommended methods show the result of linear weighting. It shows that the particle swarm algorithm in the Spark is the best one. With the size of the data enlarged, the difference is even more significant. The result is shown in Figure 3 and Table 2.

Then, the scalability of the recommended method was we verified. Within these, keep the available memory of each executor as the default value of 1G. Change the number of executors in the cluster and change the number of executors in the Spark cluster environment by changing the number of cores __tota1_executor_cores available to all available executors in the cluster. That is, when the number of available cores and memory changes, the execution time of our proposal changes. Cache sizes are 1G, 2G, and 4G enlarged four square check-in data sets. Finally, the expansibility experimental results of three kinds of particle swarm optimization algorithms are as before. These show that, with more executors, the more the increase of the number of available cores and memory is, the more it decreases linearly. In addition, the stability of the algorithm was tested in two experiments. The test results are as follows. The algorithm has good convergence and stability. The convergence of the algorithm in 2 experiments is shown in Figure 4.

The above experiments testify the performance of the particle swarm algorithm. The experimental results of verifying the validity of the integrated method show that the linear one is the best one. The second one performs well on accuracy. The third ones are the second following linear-weighted method, and it performs well on sparseness problems. The experimental results of verification efficiency show that the particle swarm algorithm has higher efficiency

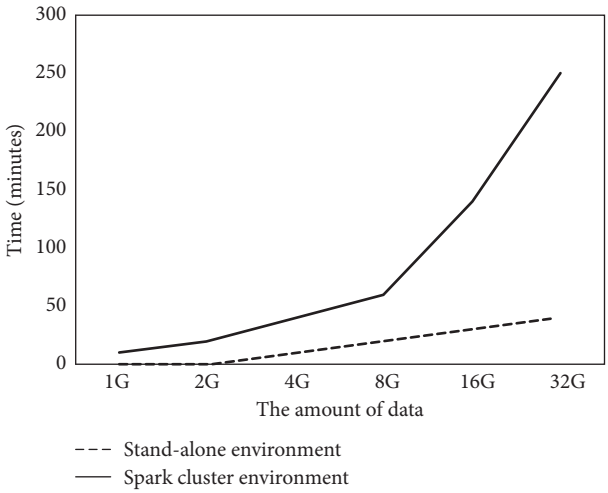


Figure 3: Different experimental cases: linear weighting.

Table 2: Different experimental cases: linear weighting.

(G)	Stand-alone environment	Spark cluster environment
1	0	5
2	0	10
4	5	30
8	15	60
16	20	140
32	30	250

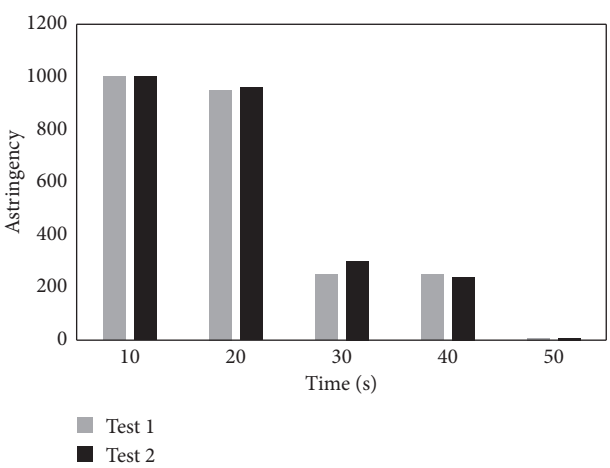


Figure 4: The convergence of the algorithm in 2 experiments.

in the Spark cluster environment than the crash environment. With the increase of the data set's scale, the efficiency advantage becomes even significantly. It shows that, as the number of available cores and memory in the cluster increases linearly, the time decreases linearly within a certain range.

4. Conclusion

At present, the development of computer and Internet technology is rapid and growing with the passage of time. With this, the combination of modern education and computer is getting closer. An effective and efficient way is needed to efficiently and precisely consider the intentional choice mechanism of course selection. For achieving the aim of the proposed study, a recommendation model and algorithm for the selection of course intention in English courses were constructed. The research results recommended by the current selection of course intentions confirm that various reinforcement factors can indeed improve the recommendation quality, but it does not consider the recommended methods for the three influencing factors. In our proposal, based on the recommendation of interest points of the existing research results, three kinds of influencing factors were integrated by three methods to improve the results of elective intentional ones. On the one hand, the problem of scalability of collaborative filtering was solved. However, efficient and easily scalable solutions were supported by point of interest recommendations. The good points and disadvantages of the three integrated methods were discussed and verified, and conclusions were conducted. Besides, the performances of the method of selection of point-of-entry intentions were verified, and it confirms that the linear one is the best choice. The experimental results perform well and show the effectiveness of the proposed study.

Data Availability

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

Conflicts of Interest

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

References

[1] M. González-Howard and K. L. Mcneill, "Learning in a community of practice: factors impacting English-learning students' engagement in scientific argumentation," *Journal of Research in Science Teaching*, vol. 53, no. 4, pp. 182–189, 2016.

[2] J. Han and H. Yin, "College English curriculum reform in mainland China: contexts, contents and changes," *Asian Education Studies*, vol. 1, no. 1, p. 1, 2016.

[3] S. Liu, "Film clips as classroom input in an elective audio-visual college English course in mainland China," *International Journal of English Language Education*, vol. 4, no. 2, p. 166, 2016.

[4] L. I. Changyu, "College English teaching reform-elective course," *Journal of Heilongjiang College of Education*, vol. 4, no. 1, pp. 16-17, 2016.

[5] A. A. A. Ari, B. O. Yenke, N. Labraoui, I. Damakoa, and A. Gueroui, "A power efficient cluster-based routing algorithm for wireless sensor networks: honeybees swarm intelligence based approach," *Journal of Network and Computer Applications*, vol. 69, no. 1, pp. 77–97, 2016.

[6] D. L. I. Dh, G. Villarubia, J. F. De Paz et al., "Multi-Sensor information fusion for optimizing electric bicycle routes using a swarm intelligence algorithm," *Sensors*, vol. 17, no. 11, p. 2501, 2017.

[7] F. Elfouly, R. Ramadan, M. Mahmoud, and M. Dessouky, "Swarm intelligence based reliable and energy balance routing algorithm for wireless sensor network," *Facta Universitatis-Series: Electronics and Energetics*, vol. 29, no. 3, pp. 339–355, 2016.

[8] D. Zouache, A. Moussaoui, and F. B. Abdelaziz, "A cooperative swarm intelligence algorithm for multi-objective discrete optimization with application to the knapsack problem," *European Journal of Operational Research*, vol. 1, no. 10, p. 264, 2017.

[9] J. L. Zhu, W. Li, H. Li et al., "A novel swarm intelligence algorithm for the evacuation routing optimization problem," *International Arab Journal of Information Technology*, vol. 14, no. 6, pp. 880–889, 2017.

[10] R. Shinde, R. Kulkarni, M. Patwardhan et al., "Conceptual schema extraction using POS annotations and weighted edit distance algorithm," *International Conference on Information Processing*, vol. 1, no. 7, pp. 719–724, 2016.

[11] S. E. Baranzini, "The genetics of autoimmune diseases: a networked perspective," *Current Opinion in Immunology*, vol. 21, no. 6, pp. 596–605, 2009.

[12] S. S. Ahmed, A. R. Ahameethunisa, W. Santosh, S. Chakravarthy, and S. Kumar, "Systems biological approach on neurological disorders: a novel molecular connectivity to aging and psychiatric diseases," *BMC Systems Biology*, vol. 5, no. 1, p. 6, 2011.

[13] S. Y. Chan, K. White, and J. Loscalzo, "Deciphering the molecular basis of human cardiovascular disease through network biology," *Current Opinion in Cardiology*, vol. 27, no. 3, pp. 202–209, 2012.

[14] M. Choura and A. Rebaï, "Exploring charged biased regions in the human proteome," *Gene*, vol. 515, no. 2, pp. 277–280, 2013.

[15] M. Choura and A. Rebaï, "Exploring disorder in the human charged biased proteins," *Journal of Receptors and Signal Transduction*, vol. 37, no. 4, pp. 386–390, 2017.

[16] D. Carvalho-Silva, A. Pierleoni, M. Pignatelli et al., "Open Targets Platform: new developments and updates two years on," *Nucleic Acids Research*, vol. 47, no. D1, 2018.

[17] The Uniprot Consortium, "UniProt: the universal protein knowledgebase," *Nucleic Acids Research*, vol. 45, no. Database issue, pp. D158–D169, 2017.

[18] G. Su, J. H. Morris, B. Demchak, and G. D. Bader, "Biological network exploration with Cytoscape 3," *Current Protocols in Bioinformatics*, vol. 47, no. 8, pp. 13.1–24, 2014.

[19] S. Johansson, M. Leonsson-Zachrisson, M. Knutsson et al., "Preclinical and healthy volunteer studies of potential drug-drug interactions between tenapanor and phosphate binders," *Clinical Pharmacology in Drug Development*, vol. 6, no. 5, pp. 448–456, 2017.

[20] S. Akhter, S. K. Nath, C. M. Tse, J. Williams, M. Zasloff, and M. Donowitz, "Squalamine, a novel cationic steroid, specifically inhibits the brush-border Na+/H+ exchanger isoform NHE3," *American Journal of Physiology*, vol. 276, no. 1Pt1, pp. 136–144, 1999.

[21] H. J. Dargie, C. T. Dollery, and J. Daniel, "Minoxidil in resistant hypertension," *The Lancet*, vol. 310, no. 8037, pp. 515–518, 1977.

[22] M. Li, A. Marubayashi, Y. Nakaya, K. Fukui, and S. Arase, "Minoxidil-induced hair growth is mediated by adenosine in cultured dermal papilla cells: possible involvement of sulfonylurea receptor 2B as a target of minoxidil," *The Journal of Investigative Dermatology*, vol. 117, pp. 1594–1600, 2001.

Hindawi
Scientific Programming
Volume 2021, Article ID 9982227, 10 pages
https://doi.org/10.1155/2021/9982227

Research Article

Enhancement of the Capability Maturity Model for Improving the Quality of Software Projects in Developing Countries

Li Hou,[1,2] **Qi Liu**,[1,2,3] **Kiran Saeed**,[4] **Saqib Ali Haidery**,[5] **M. Irfan Uddin**,[6] **and Hizbullah Khattak**[7]

[1]*School of Information Engineering, Huangshan University, Huangshan 245041, China*
[2]*Engineering Technology Research Center of Intelligent Microsystems of Anhui Province, Huangshan 245041, China*
[3]*School of Communication and Information Engineering, Shanghai University, Shanghai 200444, China*
[4]*Department of Computer Science, Preston University, Islamabad 44000, Pakistan*
[5]*Information and Communication Engineering, Nantong University, Jiangsu 210000, China*
[6]*Institute of Computing, Kohat University of Science and Technology, Kohat 26000, Pakistan*
[7]*Department of Information Technology, Hazara University, Mansehra, Pakistan*

Correspondence should be addressed to Li Hou; houli_shu@126.com

Received 25 March 2021; Accepted 20 May 2021; Published 29 May 2021

Academic Editor: Habib Ullah Khan

Business process modeling was suggested by academia and implemented by the industry to trigger device requirements. Maturity models of software processes are also similar in the software industry. Therefore, business software requirement engineering approaches-based process should aim to fulfil with the maturity models. Although researchers in the area of Software Engineering, Project Management, and Business Intelligent have addressed planning and execution and a lot of research still needs to be done, in this study, we focused on doing a questionnaire based on different prominent universities, major companies, and software houses. The purpose of the questionnaire was to dig out the main factors that negatively affect software houses and their products and services, especially in the medical field. A questionnaire was generated to accomplish these reasons, and interviews were conducted to extract the key explanation why software developers and software houses in developing countries are not familiar with the CMMI (Capability Maturity Model Integration) model for medical software projects. This analysis attempted to figure out why the tech enterprises of developing countries are behind foreign competitors. We aimed to determine why software developers do not meet the international requirements for software. The novelty of the research is that attempts were made to strengthen the new approach by adding agile at level 3 to keep the local software industry appropriate and reveal the requirements for medical software life cycle study, because accurate decision support is really necessary for healthcare systems.

1. Introduction

The classic period (1970 to 1980) is based on classic software engineering, including methods, procedures, requirement analysis tools, program architecture, scripting, testing, maintenance, and other fields of study. Most software programs now have low quality, such as missed schedule, operating costs, and other concerns. Researchers began a process to decrease production costs and increase the quality and efficiency of applications, known as improving software processes or named as improving software. Improvement of software processes has become a mainstream of software engineering study. As a result, CMMI was discovered by scientists and became the benchmark for software process capability.

CMMI is the mechanism that improves the system and provides the enterprise with the essential elements of successful processes. It may be used to monitor cross-projects, departments, or whole process enhancement organizations. CMMI is primarily used for product creation, capacity assessment, and optimization, based on project control, technical skills, and measurement of software development

processes. CMMI has been in place as an identification protocol since 1987 and has been the most authoritative credential verification method in the software industry [1].

Five stages, 18 process zones, 52 priorities, and more than 300 primary activities are included in CMMI. CMMI finds challenges resulting from software processes that modern software technologies do not boost productivity and competitiveness immediately. Higher-quality applications can be generated through this enhanced method. Not only does CMMI make multiple software projects to prevent time and resource overspending, but also it is often conducive to noticing and combining too multidisciplinary in tandem with software engineering. CMMI is an evaluation model that was developed in 1990 by Carnegie Mellon University's Software Engineering Institute (SEI). The introduction of CMMI faced a variety of problems that contributed to the advancement of CMMI as a tool of change. Based on the degree to which the organization adopts the chosen CMMI model, each organization is measured and given a CMMI ranking of 1 to 5 [2].

CMMI is a globally common software process management tool package that offers an accessible instruction system that incorporates CMMI templates, training, and methods of assessment. CMMI focuses on numerous organizational areas of concern, such as growth, procurement, and services. CMMI requires activities that include many aspects of the operation. Project management is a key process field of CMMI and the aim of this key process area is to include practices in which to handle tasks, procedures, and goods over the life cycle of the project [3].

CMMI model increased the efficiency of software and made it able to compete in the international industry, but developing countries cannot afford all levels of CMMI; therefore, they do not implement CMMI on their software; after analyzing the reason, we found that the 3rd level of CMMI is very costly, so we decided to make the small enhancement in the 3rd level of CMMI by putting agile methodology and rapid prototyping. The purpose of conducting this research is to tackle the following question: How can developing countries implement all levels of CMMI?

2. Literature Review

Other strategies related to agile methodologies and assessment models are in progress. Migrating agile methodology to structured procedures is one of the most important associated works [4], in which the technical perspective and the agile world both hold the same value in the development of an application. A structure for the application of agile concepts and values within a single enterprise was built for this purpose. In their knowledge, Vines Christ happens when an organization is accredited by CMM Level 2, ISO9001 using Extreme Programming (XP) @ Scrum. All agile approaches are combined [5]. For the technical procedures, XP was used and Scrum was implemented to support organizational and administrative challenges after a year. In a special situation, success was accomplished: in all quality level versions, the organization achieved a certification. The type was called XP 8 Scrum. This model is often used to make the production process further agile in Motorola Argentina's "Tech Global," an organization already attributed as CMM level 5. Both CMM and CMMI categorize five distinct process maturity stages based on Main Success Areas (KPAs) [6].

The authors have arranged the CMMI in [7] into five stages of maturity rise. The authors describe a direction of growing sophistication in the production processes of the enterprise and therefore in its ability to regularly and quite well complete projects. The SEI has incorporated all these CMMs together in a single holistic model called the Capability Maturity Model Integration (CMMI) in the past few years. This latest model eliminates many other older ones, like the CMM [8, 9]. Several templates and globally recognized standard and enhancement principles are currently available. CMMI and ISO can be noticed between them.

The CMMI process model, the SCAMPI evaluation method, and the related IDEAL improvement method have been developed by the SEI. The ISO 15504 process model, which is based on ISO/IEC 12207 [5, 10] and the First Amendment, its evaluation methodology ISO 15504 (Part IV), and related process enhancement ISO 15504 (Part (VII) [11, 12], was established by the ISO. A paper illustrating the agile concept was the beginning: the agile manifest [13, 14] provides a collection of concepts and standards that underpin the methodology. We have some of these techniques listed here. The aim of the SIMEP-SW project is to define the potential of obtaining CMMI certification through the use of agile approaches and how this might be accomplished. We took the strategic route between two realms with this intention: CMMI and agile methods.

3. Methods

3.1. Survey Report. In the research, a survey is conducted from educational institutes, universities, and software houses. During the survey, the aim was to cover the maximum educational institutes, and questionnaire was delivered through emictories and also distributed by hand. The participants' response was good and helpful for the survey. The questions can be seen in Table 1.

3.2. Proposed Solution. In the model, we have made changes at level 3 defined by splitting the level in two parts, part a and part b, given as follows:

(i) Processes are categorized at organizational level

(ii) Integrated rapid prototyping (RP) and joint requirement planning (JRP)

The changes at level 3 of CMMI and our proposed solution can be seen in Figure 1.

3.3. Methodology. We suggested prototypes in the form of survey and delivered these to 68 institutions, software houses, and universities by hand and by courier to know the effects of our proposed solution on the quality of software development. The flow chart of our methodology can be seen

Table 1: Survey report.

Questions	(a)	(b)	(c)	(d)
Do you think your organization has awareness about CMMI?	Nil	Good	Satisfactory	High
What are the reasons for applying CMMI in developing countries?	Project failure	Lack of business	Lack of satisfaction	Lack of customers
What is the impact of CMMI in software development?	Enhance quality	Reduce cost	Save time	Competing in international market
CMMI requirements?	Experts	Tools	Platform	Team
How does CMMI enhance quality of software projects?	Assist	Managed	Controlled	All of the above
Is CMMI effective for customers?	Yes	No	Partially	Fully
Does CMMI assure software quality?	25%	50%	75%	100%
Which one mostly affects the efficiency of software project management?	CMMI	CMM	None of these	
How does CMMI affect the capability of programmers?	Positively	Negatively	Both	None
How much is our proposed solution effective on the quality of software?	Effective	Partially effective	Not effective	
What are the requirements of software in medical field which can be essential for the specialists?	?	?	?	?

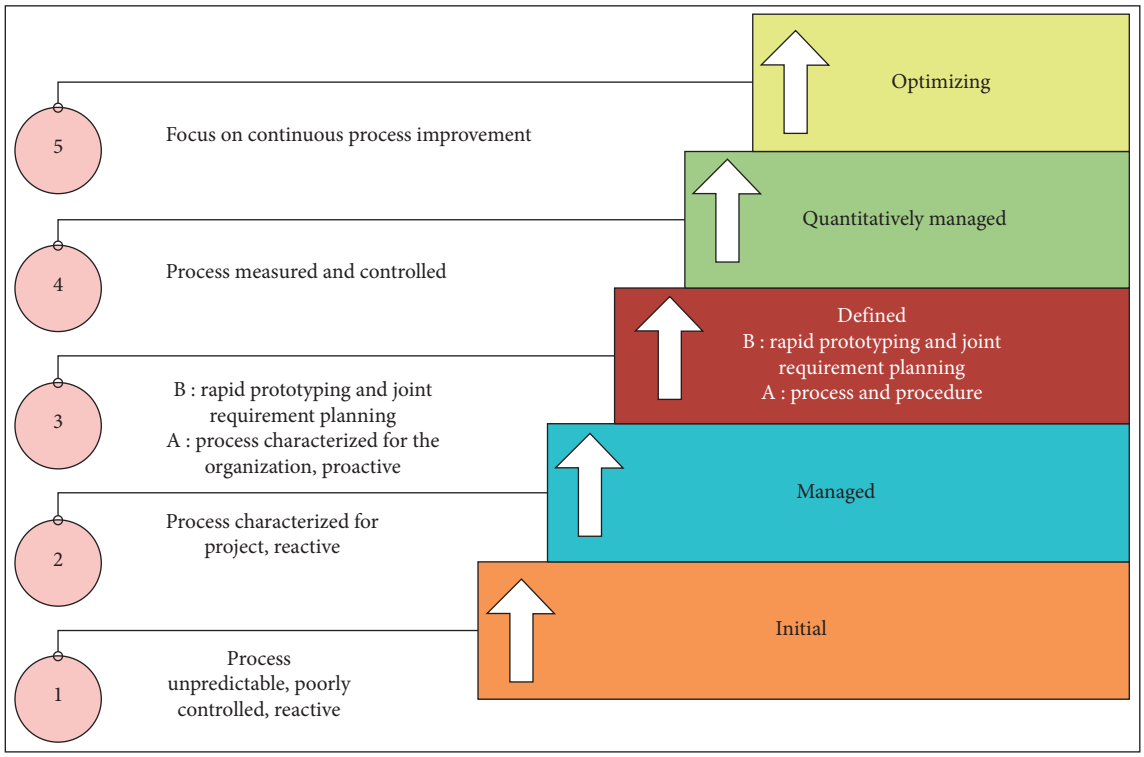


Figure 1: Proposed solution.

in Figure 2. SPSS software has been used to generate the results.

4. Results

We acquired the results through a survey asking the question, what exactly is CMMI? Most of the answers which we got from respondents demonstrated that it is a methodology used to develop and refine an organization and software development process. The results can be seen in Figure 3.

We tried to acquire the knowledge on CMMI, and 68 respondents submitted their views about it. CMMI can increase the efficiency and productivity of any piece of software, so the respondents had different knowledge and the responses are also revealed in Table 2.

There was a question about the awareness of CMMI and the answers from the respondents can be seen in Figure 4.

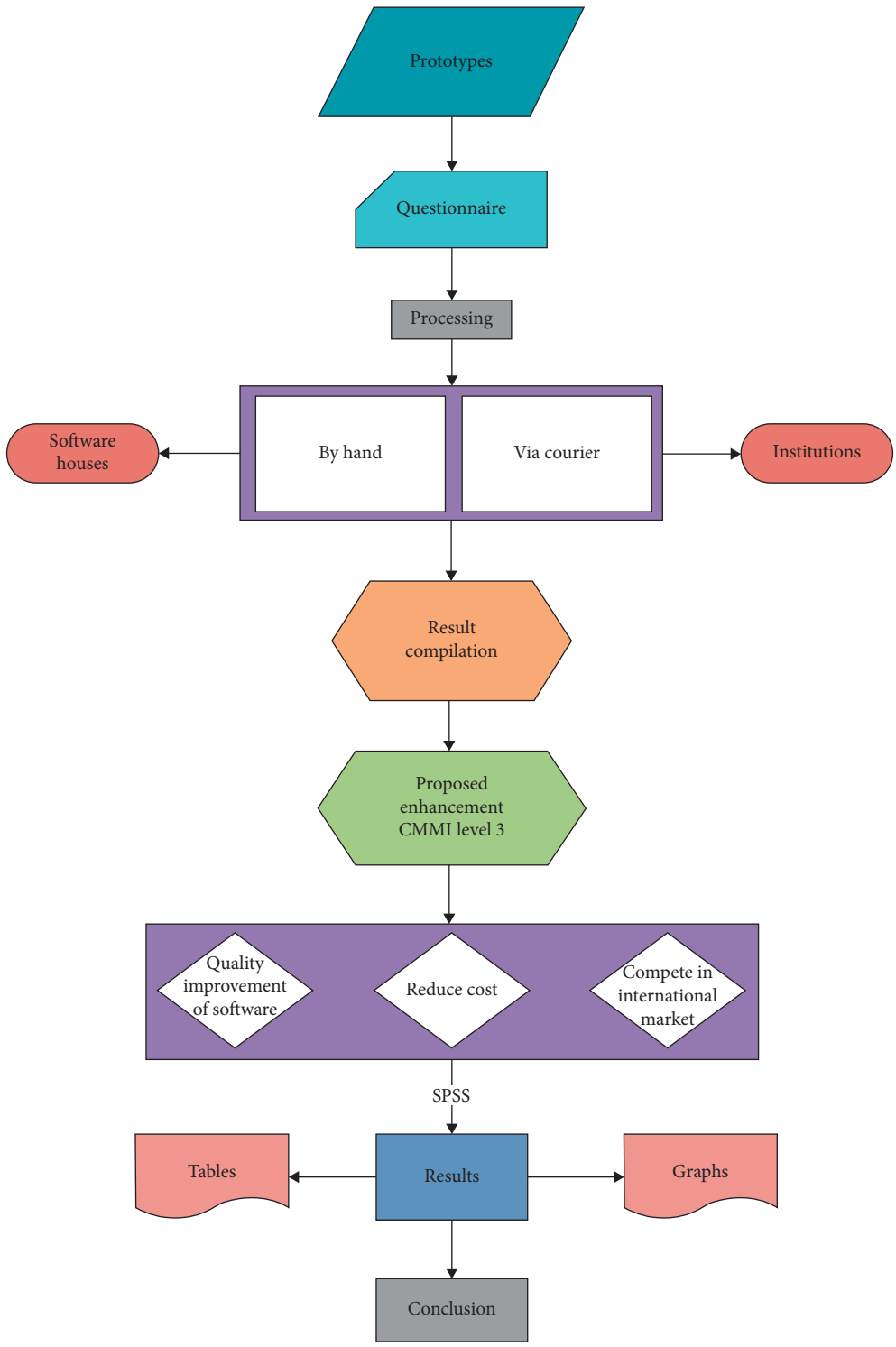


Figure 2: Methodology.

The developing countries are far behind in the production of software; that is why this industry needs to be changed for improving the quality of software and the key reasons behind this are project failure, lack of business, and lack of satisfaction. The results can be seen in Figure 5.

CMMI has great importance when it comes to the quality of software; to know the impacts of CMMI in software development, the key reasons are competing in the international market and reduce cost because in the developing countries software cost is also a big problem. The results can be seen in Figure 6.

There are many requirements for the Capability Maturity Model Integration (CMMI) in the field of software; that is why developing countries need to invest a lot if they want to

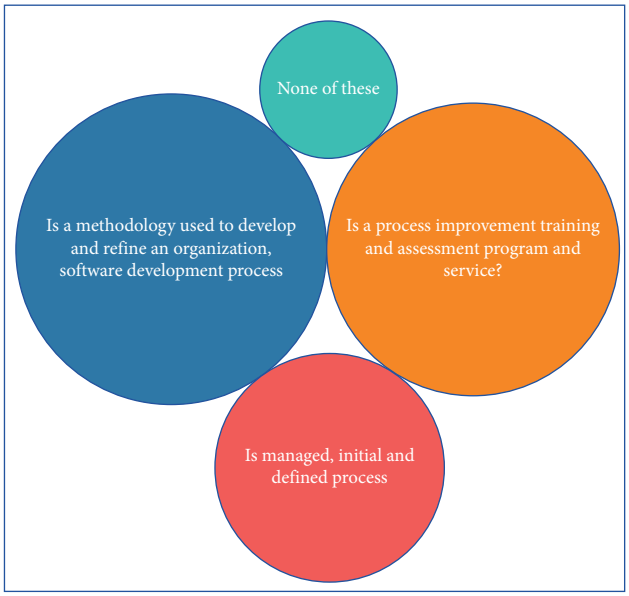


Figure 3: Views about CMMI.

Table 2: What is CMMI?

	Frequency	Percent	Valid percent	Cumulative percent
Is a process development training and assessment program and service	23	33.8	33.8	33.8
Is accomplished, initial, and defined process	14	20.6	20.6	54.4
Is a procedure used to develop and refine an organization and software development process	26	38.2	38.2	92.6
None of these	5	7.4	7.4	100.0

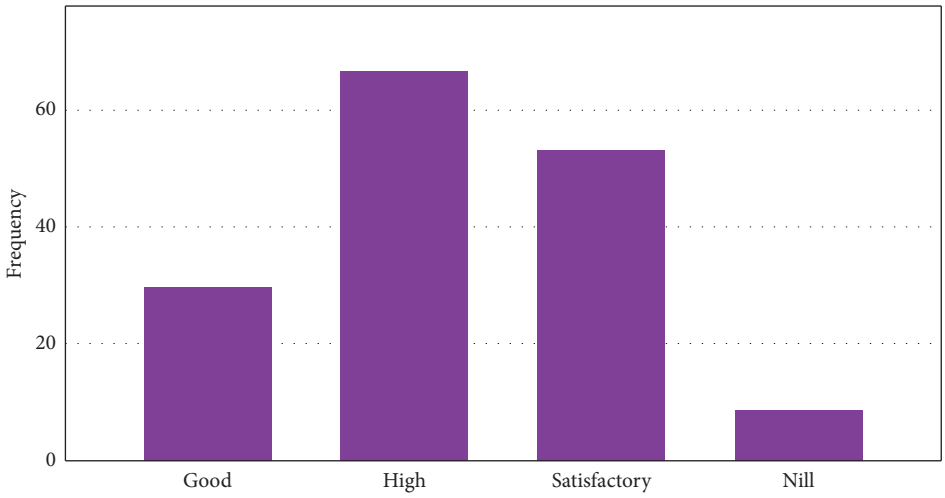


Figure 4: Awareness of CMMI.

compete on the international level software quality and some of the key reasons behind this are experts, tools, platform, and team. It can be seen in Figure 7.

CMMI enhances the quality of software projects if it is fully implemented; according to respondents of survey, CMMI can enhance assist, controlled, and managed. The results are in Figure 8.

The effectiveness of CMMI is not hidden for customers and it is being used with full trust; the respondents' answers can be seen in Figure 9.

CMMI assures the quality of software if it is deployed completely. Quality is really important for the production of software and all the customers need that the software they are going to purchase or use should have all the latest

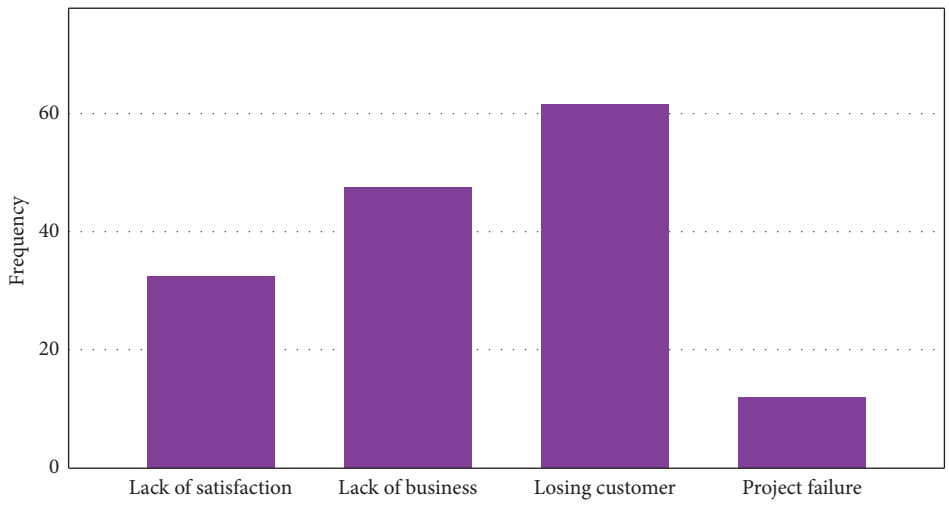


Figure 5: Reasons for applying CMMI in developing countries.

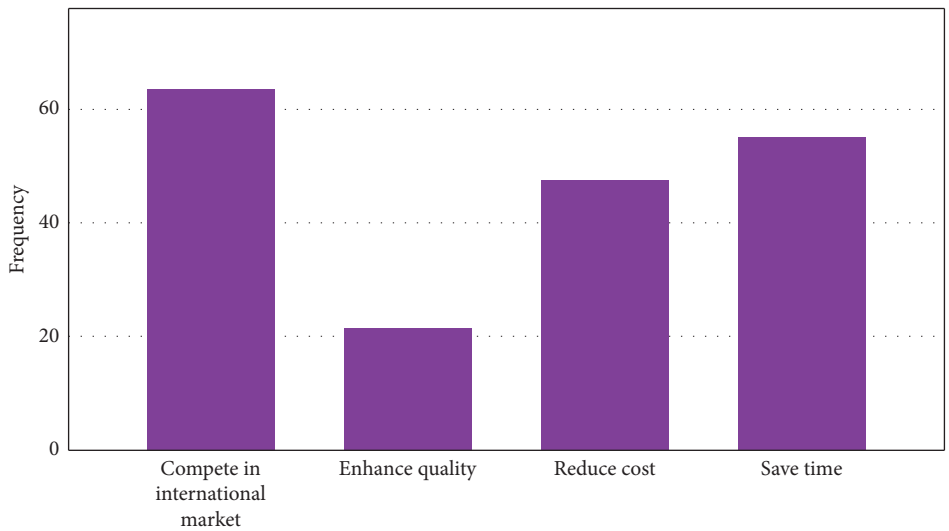


Figure 6: The impacts of CMMI.

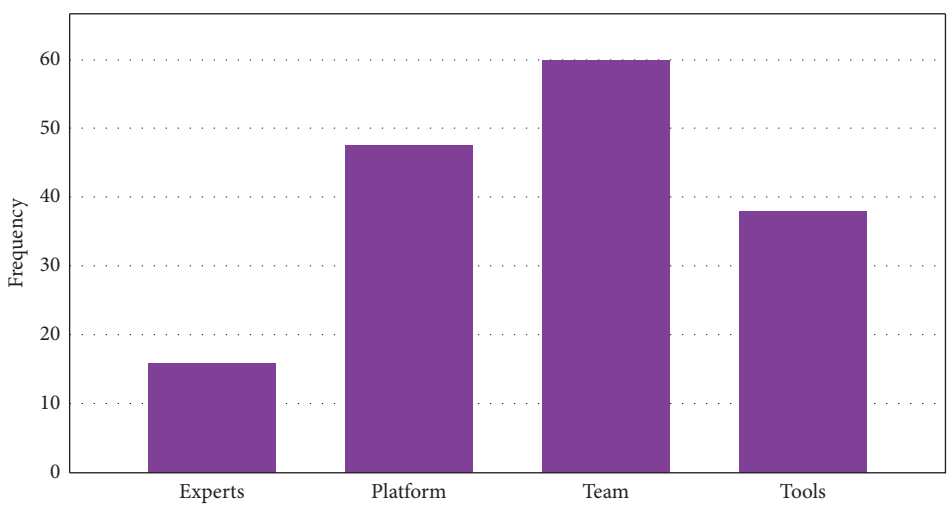


Figure 7: CMMI requirements.

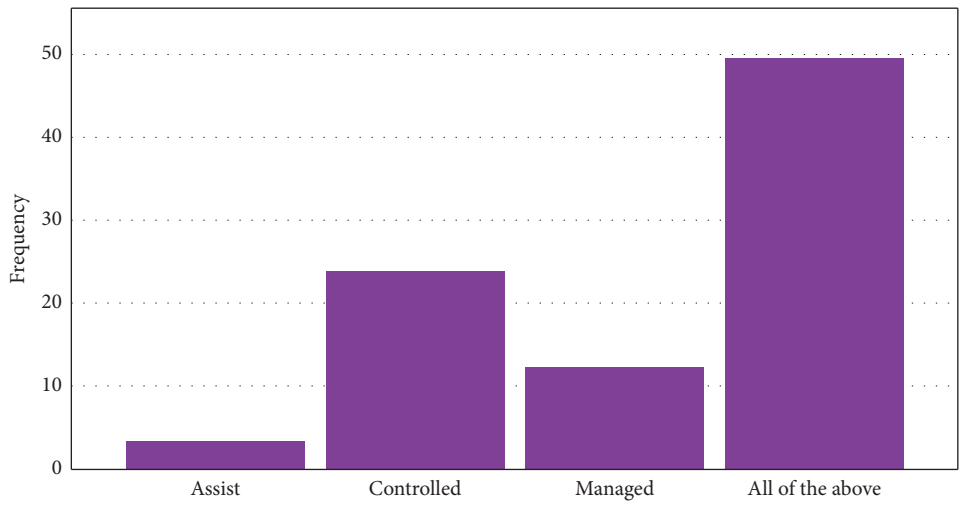


Figure 8: Enhancement of quality.

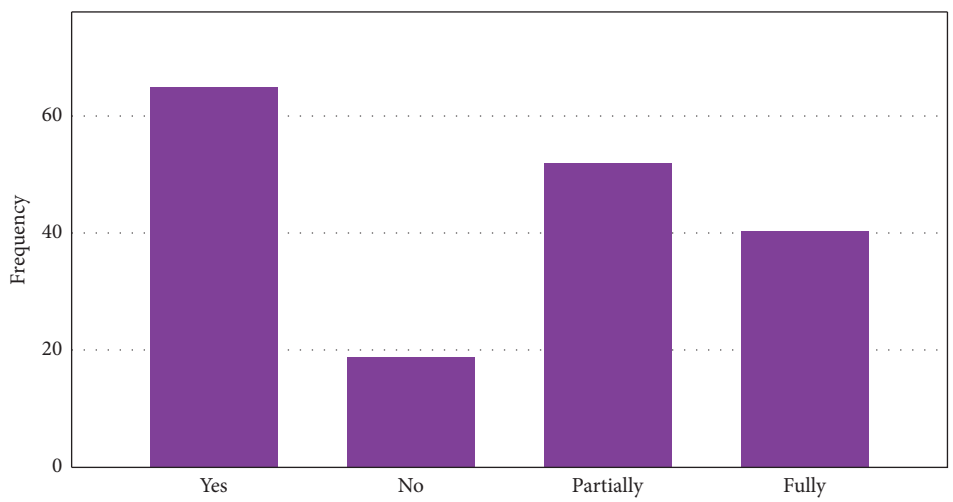


Figure 9: Effectiveness for customers.

qualities. The suggestions and prototypes which we sent to respondents have real worth because we got the answers, and the results can be seen in Figure 10.

CMMI and CMM can affect the efficiency of software project management and, according to survey, 62 respondents agree that CMMI affects the efficiency of software project management. The results are shown in Figure 11.

The capability of programmers is really important for the production of a quality product, and to know how much CMMI affects the capability of programmers, the respondents' answers were as follows, revealed in Figure 12.

CMMI has a total of 5 levels, and these are initial, managed, defined, quantitatively managed, and optimizing, and we have suggested some changes at level 3, which are process and procedure and rapid prototyping and joint requirement planning. The prototypes are being sent to software developers and organizations. They deployed the proposed solution and responded with satisfaction as can be seen in Figure 13.

The first important factor in every company is the persons that are the key drivers. There are highly educated individuals with a lot of inspiration and experience, and they know their priorities. The second engine is the instruments and resources that a company provides. In order to achieve the best performance, the new technology and sophisticated configurations are designed. Then, in order to accomplish operational targets efficiently and effectively well under time and expense, there are strategies and processes that need to be implemented. A mechanism is created by all these three main drivers. People might not even be effective every day; it can take more time and energy to understand emerging developments in technology, and policy and procedures can vary in different environments, but once an enterprise has a quality process, its efficiency can improve exponentially, and businesses that rely on processes do not have to invest spare funds into people, processes, and systems that just work smartly and boost selection. The survey is being conducted in 68 different software houses and institutions, and we got reply from them and the results can be seen in Table 3.

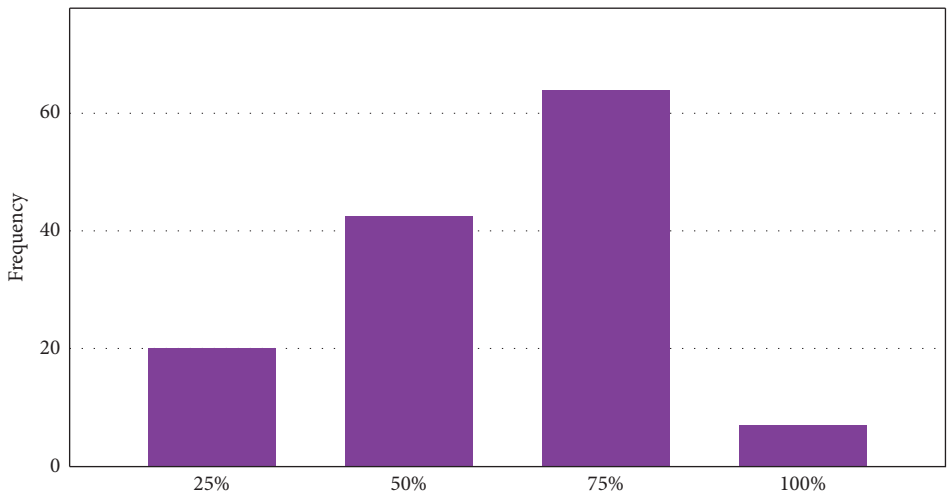


Figure 10: Software quality assurance.

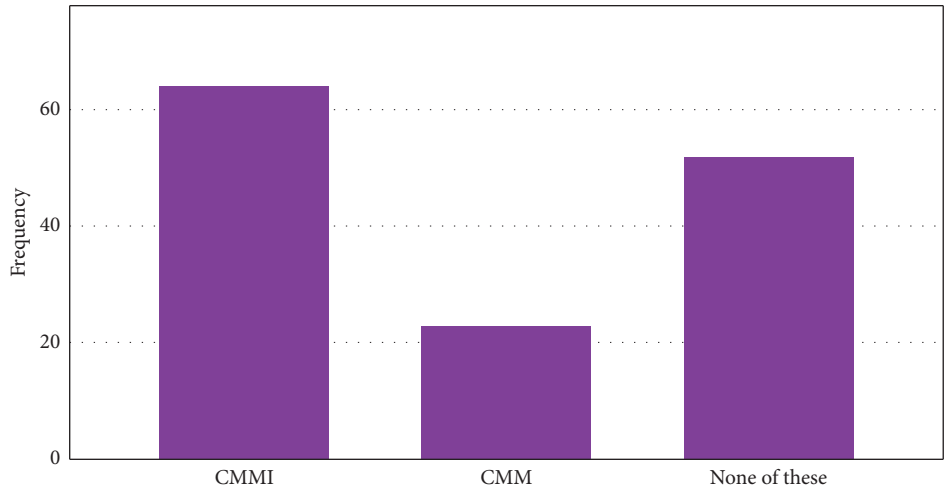


Figure 11: Efficiency of software project management.

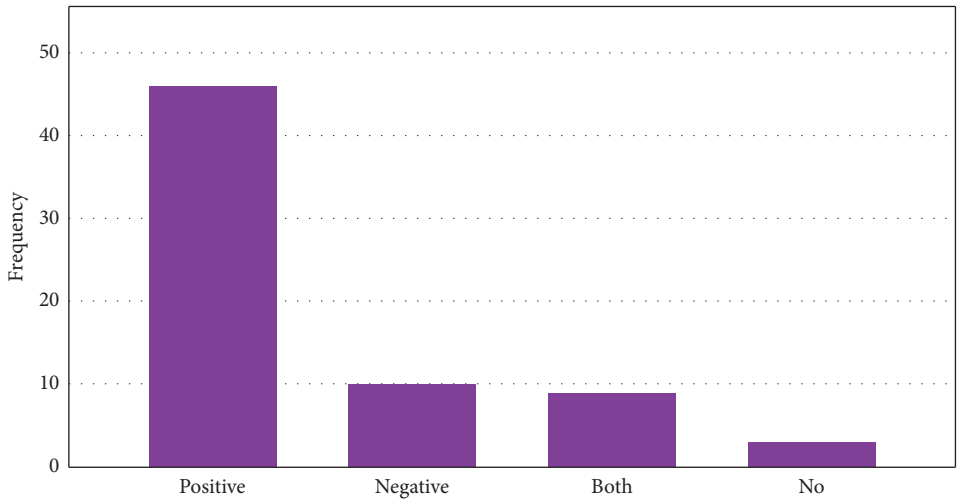


Figure 12: Effects on the capability of programmers.

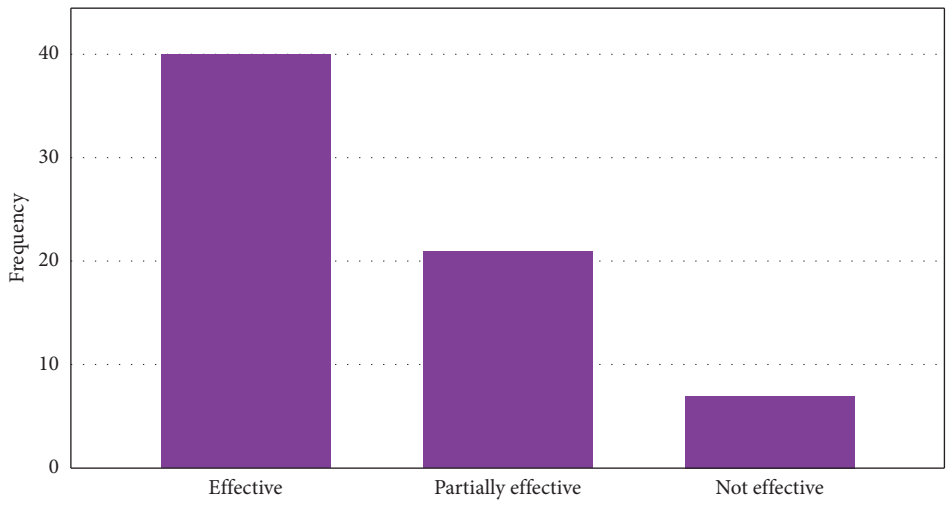


Figure 13: Effectiveness of our proposed prototype.

Table 3: Answers.

Questions	Answers
Do you think your organization has awareness about CMMI?	High
What are the reasons for applying CMMI in developing countries?	Losing customers
What is the impact of CMMI in software development?	Compete in international market
CMMI requirements?	Team
How does CMMI enhance quality of software projects?	Assist, controlled, and managed
Is CMMI effective for customers?	Yes
Does CMMI assure software quality?	75%
Which one has the biggest effect on the efficiency of software project management?	CMMI
How does CMMI affect the capability of programmers?	Positively
How much is our proposed solution effective on the quality of software?	Effective

5. Discussion

The findings are based on the primary data we collected from the survey of number of universities and tech houses that we conducted. The survey report is focused on the interpretation of data by the use of the SPSS method. The article focuses on CMMI, such that both levels relate to the process of software creation. The analysts put to the edge the factors contributing to the breakdown of the software creation phase in the survey reports. We have strong hopes that the shortcomings will be overcome by the proposed model. The findings clearly show that the tech industry cannot cope with the foreign market in today's developed countries. We find the underlying cause of failure: novice and professional workers, unpracticed software developers, and software engineers are ignorant of the CMMI model, which is why they do not use CMMI in their method of software creation.

In comparison, due to inexperience and lack of money, the second main explanation is that software companies do not have their own approach. As per our recommendation, broad methodologies should not be bought by organizations; the issue can be addressed by using the Quick Prototyping and Collaborative Requirement Preparation strategy to build their own technique using their own engineers in the tech building. This method is called agile.

The survey indicates that 38.2% of respondents replied that CMMI is a tool used to build and enhance a process of organization, as well as software creation, 33% said that CMMI is training and assessing changes in process systems and facilities, while 20.6% responded that CMMI is a controlled, original, and established process. From the study, the researchers found that 54.4% of individuals said that it was due to unawareness, while 25% said that it was due to a shortage of experts. In another survey question, 98.5% of the respondents responded that they knew about CMMI.

52.9% of the researchers found that there was a scarcity of CMMI experts in Pakistan. It was found that 38.2% of respondents believed that CMMI understanding was strong, 30.9% said that people who know about CMMI had adequate intensity, and 25% said that no one was aware of CMMI. As per the poll, 33.8% of respondents said that CMMI should be applied to limit the range loss in developed nations and

27.9% said that it should be for company growth, while 17.6% said that it will keep the client in the international environment.

The questionnaire reveals that, by supporting, handling, and monitoring, 57.4% of respondents said that CMMI increases project efficiency. Nearly 54.4% of respondents said that CMMI was efficient for consumers, while 28% said that it was slightly productive.

The results of the survey indicate that 48.5% of applicants agree that CMMI guarantees the consistency of the program. 61.8% of respondents believe that CMMI impacts the performance of software project management, according to the report. The survey indicates that almost 69.1% of respondents responded to software capabilities for CMMI effects. 45.6% of respondents said that the introduction of CMMI on projects makes them more effective, according to the study, and 39.7% said that it makes things simpler.

The research indicates that 67.6% of respondents said yes, CMMI is efficiency-oriented. It is found from the survey respondents that 100% refer to the significance of getting CMMI.

6. Conclusion

The survey revealed that the primary explanation for the loss of software projects was the absence of CMMI application in the project creation process. The issues with software production are primarily due to lack of knowledge of CMMI, professional developers of software, and financial environment. The new improved model enabled software developers to hit level 3 and also reduced the expense of buying conventional methodologies such as the Waterfall Model, the Spiral Model, and the Model for Organized Software Development. It is reported that the optimum survey participants replied that they knew about CMMI, but their responses were found to be incorrect according to the CMMI concept given while performing the survey. The lack of knowledge of CMMI was discovered as a fundamental explanation for CMMI level deployment. The research also revealed that there was a lack of country-based CMMI specialists especially for the medical related software. The need to use the highest amount of CMMI to enhance the efficiency of projects by supporting, overseeing, and monitoring is to minimize project errors in developing countries and to maintain customer input and to increase business. CMMI certainly guarantees the consistency and productivity of project management and strengthens programmers' skills, and vice versa.

Data Availability

The data used to support the study are included within the article.

Conflicts of Interest

The authors declare that there are no conflicts of interest.

Acknowledgments

This work was supported by the Anhui Natural Science Foundation (no. 1908085MF178), Anhui Excellent Young Talents Support Program Project (no. gxyqZD2019069), and Anhui Quality Engineering Projects (2019kfkc131).

References

[1] M. Choetkiertikul and T. Sunetnanta, "A risk assessment model for offshoring using CMMI quantitative approach," in *Proceedings of the 2010 Fifth International Conference on Software Engineering Advances*, IEEE, Nice, France, August 2010.

[2] I. Garcia and I. Andrea, "Using the software process improvement approach for defining a methodology for embedded systems development using the CMMI-DEV v1. 2," in *Proceedings of the 2010 10th IEEE International Conference on Computer and Information Technology*, IEEE, Bradford, UK, June 2010.

[3] J. Armstrong, R. Barbour, R. Hefner, and D. H. Kitson, "Standard CMMISM appraisal method for process improvement (SCAMPISM): improvements and integration," *Systems Engineering*, vol. 5, no. 1, pp. 19–26, 2002.

[4] B. U. Pieplow, "Project risk management handbook: a scalable approach," *Risk*, vol. 1, pp. 27–29, 2012.

[5] M. Diaz and J. King, "How CMM impacts quality, productivity, rework, and the bottom line," *Cross Talk*, vol. 15, no. 3, pp. 9–14, 2002.

[6] C. P. Team, "CMMI® for development, version 1.3, CMMI-DEV, V1. 3," Technical Report, CMU/SEI-2010-TR-033, ESC-TR-2010-033, Carnegie Mellon University, Pittsburgh, PA, USA, 2010.

[7] J. A. H. Alegrıa and M. C. Bastarrica, "Implementing CMMI using a combination of agile methods," *CLEI Electronic Journal*, vol. 9, no. 1, pp. 1–15, 2006.

[8] M. Al Yahya, R. B. Ahmad, and S. P. Lee, "Impact of CMMI based software process maturity on COCOMO II's effort estimation," *The International Arab Journal of Information Technology*, vol. 7, no. 2, pp. 129–137, 2010.

[9] J. D. Herbsleb and D. R. Goldenson, "A systematic survey of CMM experience and results," in *Proceedings of the IEEE 18th International Conference on Software Engineering*, IEEE, Berlin, Germany, March 1996.

[10] S. Gao and Q. Q. Wang, "A new security baseline reinforcement method for the power information system," *Applied Mechanics and Materials*, vol. 303–306, pp. 2407–2411, 2013.

[11] S. Zahran, *Software Process Improvement: Practical Guidelines for Business Success, Addison Wesley Longman, UK, 1997,£ 27.95, 447 pp*, Wiley Online Library, Hoboken, NJ, USA, 1999.

[12] D. Galin and M. Avrahami, "Are CMM program investments beneficial? analyzing past studies," *IEEE Software*, vol. 23, no. 6, pp. 81–87, 2006.

[13] S. Gopinath and M. Nickrad, "*How can a global Travel distribution organisation reach CMMI 5?-Solution*," M.S. thesis, KTH Royal Institute of Technology, Stockholm, Sweden, 2015.

[14] Y. L. Antonucci and R. J. Goeke, "Identification of appropriate responsibilities and positions for business process management success: seeking a valid and reliable framework," *Business Process Management Journal*, vol. 17, no. 1, pp. 127–146, 2011.

Hindawi
Scientific Programming
Volume 2021, Article ID 9993916, 8 pages
https://doi.org/10.1155/2021/9993916

Research Article

Martial Arts Training Prediction Model Based on Big Data and MEMS Sensors

Shisen Li,[1] Chao Liu ,[2] and Guoliang Yuan[2]

[1]*Department of Physical Education, North China University of Science and Technology, Tangshan 063210, Hebei, China*
[2]*College of Physical Education, Heng Shui University, Heng Shui 053000, Hebei, China*

Correspondence should be addressed to Chao Liu; liuchaotrea@163.com

Received 9 March 2021; Revised 20 April 2021; Accepted 3 May 2021; Published 27 May 2021

Academic Editor: Shaukat Ali

In martial arts teaching and sports training, the accurate capturing and analysis of martial arts athletes' posture is conducive to accurately judge sports postures, as well as correcting sports movements in a targeted manner, further improving martial arts athletes' performance and reducing physical damage. The manufacturing level of MEMS sensors continues to improve, and status perception of assembly objects is becoming more and more abundant and accurate. The shape is small and can be worn, and data can be collected continuously without obstacles. The price is relatively low, the privacy protection is strong, and the advantages are clear and prominent. A considerable number of technicians choose to use MEMS sensors as the main tool for human behavior detection data collection. Therefore, this article designs multiple MEMS inertial sensors to form a human body lower limb capture device, and its core components are composed of accelerometer, gyroscope, and magnetometer. In order to make the obtained acceleration value, angular velocity value, and magnetometer value accurately reflect the movement state of the lower limb structure, different data fusion algorithms and magnetometer ellipsoid fitting calibration algorithms are studied to realize the calculation of the posture angle of each joint point and obtain martial arts posture big data. In addition, through big data analysis, this article designs a martial arts training performance and injury risk prediction model, which can provide guidance and suggestions for martial arts teaching tasks.

1. Introduction

In martial arts teaching and sports training, the accurate capturing and analysis of martial arts athletes' posture is conducive to accurately judging sports postures, as well as correcting sports movements in a targeted manner, further improving martial arts athletes' performance and reducing physical damage. At present, human motion posture capture technology is mainly based on three methods: visual images [1–4], electromyographic signals [5–7], and wearable inertial sensors [8–10]. Among them, the human body gesture capture technology based on wearable inertial sensors has the advantages of directness, reliability, and strong applicability.

The mainstream human motion capture systems are optical and MEMS inertial sensors [11]. The optical motion capture system is currently the most authoritative motion capture method. Its capture range is wide, the response speed is fast, and the data delay is small. It can realize the real-time capture of the action. However, the data processing process of the optical work capture system includes identification, tracking, and coordinate calculation. The calculation workload is relatively large, the related equipment is expensive, the optical conditions are demanding, and the effect of motion capture in the occluded state is poor. The MEMS-type inertial sensor motion capture analysis equipment can overcome the problem of insufficient test accuracy of the optical system in the low-light environment and motion blocking state, and the cost of the equipment is low, the test accuracy is high, and the response speed is fast, which is conducive to real-time motion capture. At the same time, it can support the parallel capture of the motion states of multiple objects.

In recent years, the concept of "big data" has gradually attracted attention. Its definition originated from the explosion of massive data in the information age, and it was defined as the

technological development and innovation related to it. Martial arts [12, 13] is one of the latest sports projects to apply the concept of big data. In today's highly developed professional events, with the advancement of science and technology and the update of equipment, the data that can be collected is more and more abundant, not only for the audience. It provides the possibility of in-depth understanding of the superb skills of top-level martial arts athletes and also provides data support for martial arts athletes to improve their competitive skills and formulate training competition programs.

Therefore, this article designs multiple MEMS inertial sensors to form a human body lower limb capture device, and its core components are composed of accelerometer, gyroscope, and magnetometer. In order to make the obtained acceleration value, angular velocity value, and magnetometer value accurately reflect the movement state of the lower limb structure, different data fusion algorithms and magnetometer ellipsoid fitting calibration algorithms are studied to realize the calculation of the posture angle of each joint point and obtain martial arts posture big data. In addition, through big data analysis, this paper designs a martial arts training performance and injury risk prediction model, which can provide guidance and suggestions for martial arts teaching tasks. The following are the main contribution points of this paper:

(i) Aiming at the analysis of martial arts movement posture, this paper designs multiple MEMS inertial sensors to form a human lower limb capture device. Its core components are composed of accelerometer, gyroscope, and magnetometer. In order to make the obtained acceleration value, angular velocity value, and magnetometer value accurately reflect the movement state of the lower limb structure, different data fusion algorithms and magnetometer ellipsoid fitting and calibration algorithms are also studied, and the posture angle of each joint point is calculated.

(ii) Aiming at martial arts training performance and injury risk prediction, this paper proposes a big data analysis method and designs a support vector regression model to achieve more accurate prediction.

(iii) This paper also conducted a simulation experiment. Experimental results show that the algorithm can effectively improve the accuracy of martial arts training performance prediction and injury risk prediction.

The remainder of the paper is organized as follows Section 2 briefly describes the background of the study. Section 3 depicts the methodology of the proposed study. The experimental work is done in Section 4, and the paper is concluded in Section 5.

2. Background

There is a wearable inertial sensor product in the market as shown in Figure 1, but it has the disadvantages of large size and troublesome wearing. Some scholars fix the acceleration sensor on the human body's motion part and realize the measurement of 5 common motion postures of the human body. However, the sensitivity of the acceleration sensor used is not high enough to recognize subtle movements. The accelerometer is worn by the athlete to measure the athlete's motion posture, thereby identifying high-risk events of sports injury in the game, but the frequency sampling range of the accelerometer needs to be improved, and the capture of more rapidly changing movements is limited. In the process of human body posture capture, the accelerometer is also required to have the characteristics of small size, low power consumption, and being easy to wear. Therefore, the study of MEMS accelerometers [14, 15] with high sensitivity and high frequency response is of great significance for human body posture capture.

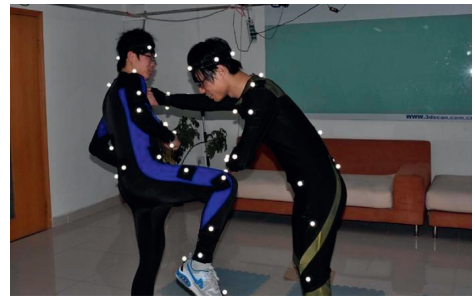


Figure 1: Martial arts motion capture system based on MEMS sensor.

2.1. MEMS Inertial Sensor. The emergence of Microelectromechanical Systems (MEMS) technology has led to a huge revolution in martial arts posture measurement technology, making martial arts posture measurement develop in the direction of low cost, low power consumption, and miniaturization. The combination of MEMS accelerometer and MEMS gyroscope is generally called an inertial measurement unit, which is used for the measurement and analysis of object posture [16]. In this paper, MEMS sensors are used to analyze the gait characteristics of martial arts athletes. In order to improve the measurement accuracy, based on the use of MEMS accelerometers and MEMS gyroscopes, magnetometers are introduced to correct the posture parameters such as acceleration and angular velocity obtained from the test to further improve the accuracy of gait parameters. Obtain and provide powerful experimental data support for gait feature analysis. The working principles, structural features, and uses of MEMS accelerometers, MEMS gyroscopes, and MEMS magnetometers are introduced below. On this basis, a sensor system suitable for motion capture in this research is selected.

2.1.1. MEMS Gyroscope. According to different components, gyroscopes can be divided into mechanical gyroscopes [16], fiber optic gyroscopes [17], laser gyroscopes [18], MEMS gyroscopes, and so forth. The mechanical gyroscope is a mechanical device that uses the gyro torque generated by the vibration of the carrier to measure the

angular movement. It has extremely high requirements on the processing technology, the structure is very complicated, and the accuracy is restricted in many aspects. The fiber optic gyroscope (FOG) uses the Sagnac effect to sense the change of direction, thereby realizing the function of a mechanical gyroscope. Its working principle is based on light interference, with high sensitivity and stable and reliable operation. The principle of the laser gyroscope is to use an optical oscillator as an optical loop, and its distinguishing features are simple structure and long working life. MEMS gyroscope is a microelectromechanical gyroscope, which refers to a technology that uses silicon as a substrate for micron-level processing. It generally uses the principle of the angular velocity of the vibrating object sensor. MEMS gyroscopes have the characteristics of low power consumption, small size, low cost, easy installation, and mass production. Therefore, they are widely used in the low-end market. Various gyroscopes are shown in Figure 2. Due to the high cost and large volume of mechanical gyroscopes, fiber optic gyroscopes, and laser gyroscopes, they are not suitable for human wear. Therefore, this article chooses MEMS gyroscope as one of the devices in the inertial measurement unit. However, most MEMS gyroscopes have the characteristics of high noise and serious drift. To make up for this shortcoming, this article uses accelerometers and magnetometers to compensate the gyroscope's attitude measurement accuracy, which is one of the research focuses of this article. In this paper, the angular velocity provided by the gyroscope is used for data fusion with acceleration and magnetic value to calculate the attitude angle; on the other hand, the angular velocity is used as a monitoring signal for zero-speed judgment and as a signal for multisensor feature layer fusion.

2.1.2. MEMS Accelerometer. With the continuous development and maturity of MEMS acceleration sensor technology, MEMS acceleration sensors with high sensitivity, small size, and low price have been applied to electronic consumer products, such as mobile phones, smart watches, and notebook computers. The MEMS accelerometer is mainly used to measure the acceleration of the carrier. The carrier speed can be obtained by integrating the acceleration once, and the carrier displacement can be obtained by integrating again. Compared with the gyroscope, the accelerometer has good static stability, while the gyroscope has very good dynamic stability, which leads to relatively unreliable data at rest. The accelerometer can only accurately measure the attitude when the linear acceleration is negligible when stationary or in motion, while the azimuth update of the gyroscope will be affected by drift, and the data of the magnetometer is easily disturbed by the surrounding magnetic field. Therefore, this article chooses accelerometer as one of the inertial measurement unit devices. On the one hand, it is used to measure the acceleration value of the key nodes of the human body and obtain the corresponding static information. On the other hand, it can also be used to correct the drift error of the gyroscope and magnetometer to realize the human body.

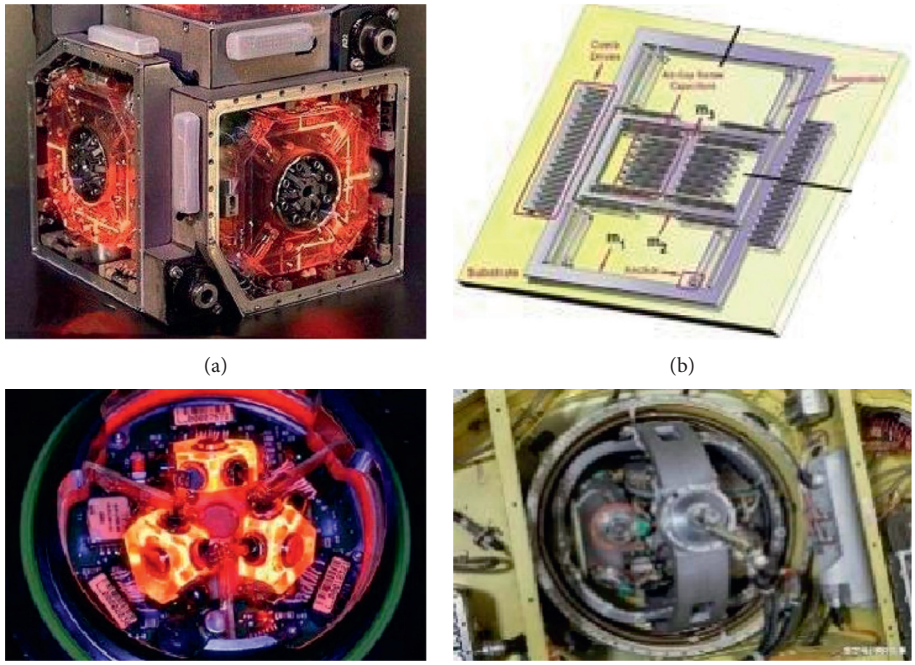
2.2. Support Vector Machine. Support vector machine (SVM) [19] was proposed in 1964. Its theoretical basis follows the principle of structural risk minimization. After that, many scholars have studied the solution method of SVM. In 1992, Boser et al. [20] researched and proposed a block algorithm for solving SVM, which decomposed the original solution problem into a series of subproblems to be optimized and reduced the difficulty of solving calculation. In 1997, Osuna [21] proposed the use of a decomposition algorithm to solve the SVM. While ensuring that the subproblem remains unchanged, the variable that needs to be optimized is called the working set, and other nonworking set variables remain unchanged. The shortcomings of the problem size in the algorithm are too large. In 2000, Mangasarian and Musicant [22] introduced the smoothing function into the SVM solution algorithm and transformed the original constraint problem into an unconstrained problem and then used Newton's iteration method to solve it. Mangasarian and Musicant [23] also proposed the Lagrangian SVM algorithm, which converges linearly and uses the implicit Lagrangian formula to get the iterative formula of LSVM, which is more convenient to implement. Although scholars and researchers have done a lot of research on SVM solving algorithms and have achieved good results, designing more efficient SVM solving algorithms is helpful to the application and development of SVM, which is still a direction worth studying.

3. Methodology

3.1. Martial Arts Motion Capture Device Based on MEMS Inertial Sensor Unit. The main chip of the sensor unit constructed in this paper adopts TDK IIM-46230, and its output parameters include the acceleration value, angular velocity value, and magnetic force value of the carrier, which satisfy the research of the attitude fitting algorithm in this paper. The sensor package structure is shown in Figure 3, and the parameter index of the inertial sensor is shown in Table 1.

In the inertial sensor unit, the main performance indicators of the gyroscope include measurement range, resolution, zero bias, and zero bias stability. The zero bias is an important parameter of the gyroscope, so only its zero bias output is tested. The accelerometer test mainly measures the modulus value of its XYZ axis, and the modulus value should be equal to the local gravitational acceleration at rest. The measurement of the magnetometer also measures the modulus of its XYZ axis. Place the inertial sensor in an environment without magnetic field interference for 1 hour, and collect its output data of angular velocity, acceleration, and magnetic force.

3.2. Composition of Martial Arts Posture Analysis System Based on MEMS Inertial Sensor. The martial arts athlete's posture analysis system is mainly composed of inertial sensors, data forwarding routers, and computer terminals. The form of its composition is called the human body sensor network (BNS). Specifically, it refers to the human-centered



(a) (b) (c) (d)

Figure 2: Four different kinds of gyroscopes. (a) Laser gyroscope. (b) MEMS gyroscope. (c) Fiber optic gyroscope. (d) Mechanical gyroscope.

integration of intelligent sensors, biomedicine, mechatronics, data acquisition preprocessing, multisensor data fusion, artificial intelligence, wireless communication, and other forms of multidisciplinary knowledge to realize the collection and fusion of various human motion data.

Calculation. The 7 inertial sensors of the lower limbs of martial arts athletes are used as the data collection layer of the entire system. These 7 sensors are, respectively, bound to the left instep, right instep, left calf, right calf, left thigh, right thigh, and crotch of the human body to collect various parts. The acceleration, angular velocity, and magnetic value of the magnetic field are shown in Figure 4(a). The transmission rate of the inertial sensor is 60 Hz, the indoor transmission distance is 40 m, the outdoor transmission distance is 110 m, and the working time is 10 h. The network forwarding layer is a data forwarding router, and the data of the inertial sensor is forwarded to the computer terminal through the router in the form of 2.4 G Wi-Fi forwarding. The router of the data forwarding layer is shown in Figure 4(b). The data of multiple inertial sensors collected by the data forwarding router is connected to the computer terminal through a USB cable. The inertial sensor data fusion, lower limb posture fitting, calculation of gait parameters, and various algorithm research simulations in this paper are all completed on the computing terminal, as shown in Figure 4(c).

The human sensor network of the martial arts athlete's gait analysis system is divided into three levels, namely, the data collection layer, the data transmission layer, and the data fusion calculation layer, as shown in Figure 5.

3.3. Support Vector Regression Prediction Model. As the martial arts training performance and injury risk prediction problem contains a large number of complex features, the data dimensionality is very high, and the support vector regression (SVR) machine model is a commonly used model for processing high-dimensional data, which is used in large-scale complex data problems There will be very good robustness and accuracy. The core of the support vector regression model is the support vector machine (SVM), which maps high-dimensional data to a linear space and classifies it by linear regression.

Support vector machine is a two-class classification model that defines the linear classifier with the largest interval in the feature space. When the kernel function is introduced, support vector machines can also solve nonlinear problems. The main idea of the support vector machine is to find an optimal hyperplane in the feature space to completely separate the positive and negative samples, as shown in Figure 6.

For a training data set on a given feature space,

$$T = \{(x_1, y_1), (x_2, y_2), \ldots, (x_N, y_N)\}, \tag{1}$$

where $x_i \in R^n$, $y_i \in \{+1, -1\}$, x_i represents the i-th feature vector, and y_i is the class label of x_i. The model can be expressed as the following convex quadratic programming problem:

$$\min_{w,b,\zeta} \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{N} \zeta_i, \tag{2}$$

$$\text{s.t.} \quad y_i(wx_i + b) \geq 1 - \zeta_i, \quad i = 1, 2, \ldots, N, \tag{3}$$

(a) (b)

Figure 3: TDK IIM-46234 inertial sensor.

Table 1: The parameters of TDK IIM-46234 inertial sensor.

Types	Gyro	Accelerometer	Magnetometer
Range	280°/s	4 g	20 Guss
Nonlinearity	0.1%	0.5%	0.1% FS
Noise density	0.009	100 ug	5 mG
xyz orthogonality	0.1	0.1	1
Sensitivity	17.5 LSB	2048 LSB	0.6uT/LSB
Start time	30 ms	25 ms	22 ms
Output rate	120 Hz	120 Hz	120 Hz
Axis	3-axis	3-axis	3-axis

$$\zeta_i \geq 0, \quad i = 1, 2, \ldots, N, \tag{4}$$

where ζ_i is a slack variable, and $C > 0$ is called a penalty parameter, which is generally determined by the actual situation. A large value of C means a large penalty for misclassification, and a small value of C means a small penalty for misclassification. Minimizing the objective function (equation (2)) contains two meanings: (1) Make $1/2\|w\|^2$ as small as possible; that is, maximize the interval. (2) Make the number of misclassified points as small as possible, and C is a parameter that reconciles the two.

The above equations represent a convex quadratic programming problem, so the solution about WB exists; and it can be proved that the solution of w is unique, but the solution of b is not unique, and the solution of b is in an interval. In order to solve the convex quadratic programming problem, equations (2)–(4) can be transformed into the corresponding dual problem:

$$\begin{aligned} \min \quad & L(w, b, a) = \frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N} a_i a_j y_i y_j \left(x_i \times x_j\right) - \sum_{i=1}^{N} a_i, \\ \text{s.t.} \quad & \sum_{i=1}^{N} a_i y_i = 0, \\ & 0 \leq a_i \leq C, \quad i = 1, 2, \ldots, N, \end{aligned} \tag{5}$$

where a is the Lagrange multiplier and $(x_i \times x_j)$ is the inner product of x_i and x_j. The optimal solutions w^* and b^* of the original problem can be expressed as

$$\begin{aligned} w^* &= \sum_{i=1}^{N} a_i^* y_i x_i, \\ b^* &= y_j - \sum_{i=1}^{N} y_i a_i^* \left(x_i \times x_j\right). \end{aligned} \tag{6}$$

Thus, the separating hyperplane can be expressed as

$$\sum_{i=1}^{N} a_i^* y_i \left(x \times x_i\right) + b^* = 0. \tag{7}$$

The classification decision function can be expressed as

$$f(x) = \operatorname{sign}\left(\sum_{i=1}^{N} a_i^* y_i \left(x \times x_i\right) + b^*\right). \tag{8}$$

In the martial arts training performance and injury risk prediction, the real value is directly output:

$$y = \sum_{i=1}^{N} a_i^* y_i \left(x \times x_i\right) + b^*. \tag{9}$$

4. Experiments

4.1. Experimental Environment. The hardware platform and software simulation environment of the experiment in this paper are shown in Table 2.

4.2. AUC Evaluation Criteria. Martial arts athletes training injury risk prediction is a typical two-classification problem. It is difficult to reasonably reflect the effect of classification models that deal with large-scale data using traditional evaluation methods such as accuracy. For a two-category problem, first set a threshold according to the prediction result. When the estimated value is higher than the threshold, it means click; otherwise, it means no click. For a

(a) (b) (c)

Figure 4: Physical map of martial arts athlete sensor network. (a) Data acquisition sensors. (b) Data forwarding muter. (c) Computer terminal.

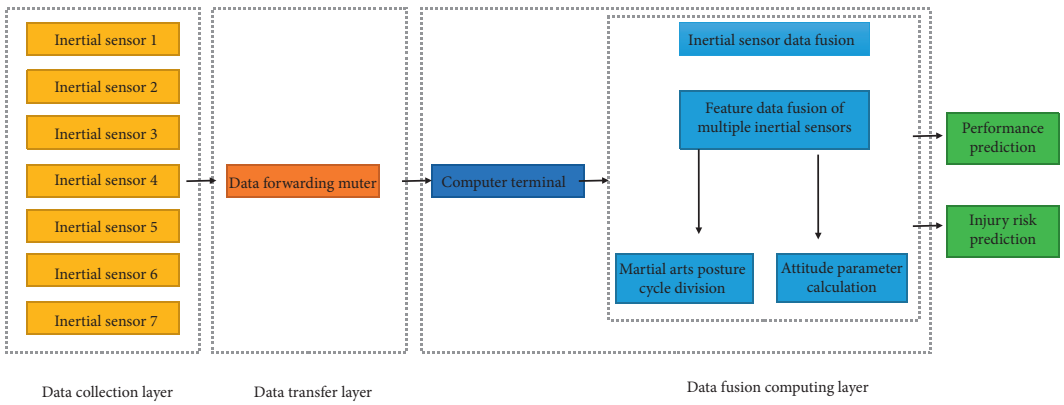


Figure 5: Martial arts athlete's training analysis and prediction system.

specific example, the predicted effect can be reflected by the four types in Table 3.

For the entire test set, the training injury risk prediction results of martial arts' athletes are usually described by the ROC curve. FPR is the abscissa of the curve, indicating that the risk-free records are classified into the prediction report, as shown in formula (10); the ordinate is TPR, which means that the risky records are correctly assigned to the forecast report, as shown in formula (11). Under normal circumstances, the TPR value is greater than the FPR value; otherwise, it means that the prediction result is reversed. The AUC value is used to represent the area under the curve, which is an important measurement index for the evaluation of the prediction results of the martial arts athletes' training injury wind. The higher the value, the better the forecasting effect. Moreover, the value of AUC is between 0.5 and 1 under normal circumstances, and the ROC curve is convex, as shown in Figure 7.

$$\mathrm{FPR} = \frac{\mathrm{FP}}{\mathrm{FP} + \mathrm{TN}}, \tag{10}$$

$$\mathrm{TPR} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FN}}. \tag{11}$$

4.3. Experimental Results and Analysis. In this paper, the training set randomly sampled from the data set is 10,000, and the test set is 3,000. Three basic models are used to predict martial arts training performance and injury risk,

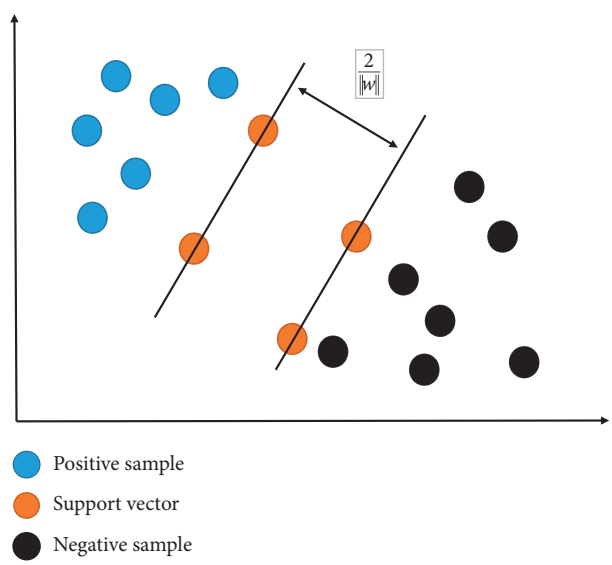


Figure 6: Support vector machine model.

and the AUC value and training time are used as evaluation indicators to explore the advantages and disadvantages of the logistic regression model.

This article compares the three basic models. The features extracted by the model in the experiment are shown in Table 4, where "f_Sparse" is the category sparse feature, "p_CTR" is the historical click rate feature, "Similar" is the similarity feature, "Pos" is the location feature, and "W2V" is

Table 2: Experimental hardware platform and software simulation environment.

CPU	Intel(R) Core(TM) i5-2450M CPU @ 2.50 GHz (2500 MHz)
RAM	4.00 GB
Operating system	Ubuntu 14.04 LTS
Development environment	PyCharm
Programming language	Python

Table 3: Confusion matrix.

TP	True positive	Predict the instances from the positive sample correctly
TN	True negetive	Predict the instances from the negative samples correctly
FP	False positive	Predict the instances from the positive sample incorrectly
FN	False negetive	Misprediction of instances from the negative samples

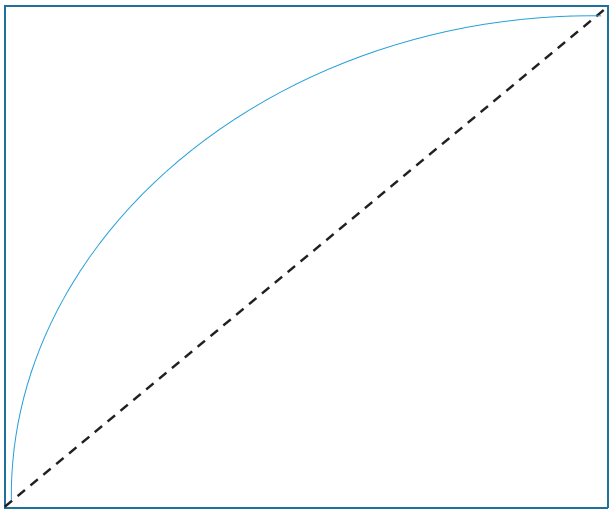


Figure 7: ROC curve.

Table 4: Confusion matrix.

Model	AUC	Time (min)
Naive Bayes	0.642	18
Logistic regression	0.689	15
Support vector regression (ours)	0.758	12

the word vector feature. For the case where the prior probability in the naive Bayes model is 0, the experiment uses plus one smoothing for processing. The experimentally predicted AUC and training time are shown in Table 4.

Experiments show that the prediction effect of the naive Bayes model is obviously inferior to those of the logistic regression model and the support vector regression model. Not only is the AUC value low, but also the training time is the longest. The reason may be a hypothetical estimation based on training data in prior probability, as well as a certain correlation between features, while naive Bayes trains the model based on the assumption of feature independence. In addition, the AUC value predicted by the logistic regression model is slightly lower than that of the support vector regression model, and the training time is longer. The support vector regression model can capture the key samples in the training set and use them as support vectors, which makes the model very robust.

5. Conclusion

The accurate capturing and analysis of martial arts athletes' posture in martial arts teaching and sports training is beneficial to accurately judge sports postures, as well as adjusting sports movements in a targeted manner with further refining martial arts athletes' performance and decreasing physical damage. To improve the prediction accuracy of the training injury risk of martial arts athletes and reduce physical damage, this article designs multiple MEMS inertial sensors to form a human body lower limb capture device, and its core components are composed of accelerometers, gyroscopes, and magnetometers. In order to make the obtained acceleration value, angular velocity value, and magnetometer value accurately reflect the movement state of the lower limb structure, different data fusion algorithms and magnetometer ellipsoid fitting calibration algorithms are studied to realize the calculation of the posture angle of each joint point and obtain martial arts posture big data. In addition, through big data analysis, this paper designs a martial arts training performance and injury risk prediction model, which can provide guidance and suggestions for martial arts teaching tasks.

Data Availability

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

Conflicts of Interest

All the authors declare no conflicts of interest.

References

[1] L. Zhang, X. Wang, X. Dong, L. Sun, W. Cai, and X. Ning, "Finger vein image enhancement based on guided tri-gaussian filters," *ASP Transactions on Pattern Recognition and Intelligent Systems*, vol. 1, no. 1, pp. 17–23, 2021.

[2] W. Cai, B. Liu, Z. Wei, M. Li, and J. Kan, "TARDB-Net: triple-attention guided residual dense and BiLSTM networks for hyperspectral image classification," *Multimedia Tools and Applications*, vol. 80, no. 7, Article ID 11291, 2021.

[3] X. Ning, Y. Wang, W. Tian, L. I. U. Liang, and W. Cai, "A biomimetic covering learning method based on principle of

homology continuity," *ASP Transactions on Pattern Recognition and Intelligent Systems*, vol. 1, no. 1, pp. 9–16, 2021.

[4] X. Ning, X. Wang, S. Xu et al., "A review of research on co-training, concurrency and computation: practice and experience," *Bio-Engineering Applications*, vol. 32, Article ID e6276, 2021.

[5] H. Su, S. E. Ovur, X. Zhou, W. Qi, G. Ferrigno, and E. De Momi, "Depth vision guided hand gesture recognition using electromyographic signals," *Advanced Robotics*, vol. 34, no. 15, pp. 985–997, 2020.

[6] D. Buongiorno, G. D. Cascarano, I. De Feudis et al., "Deep learning for processing electromyographic signals: a taxonomy-based survey," *Neurocomputing*, 2020, In press.

[7] S. Sapienza, P. M. Ros, D. A. F. Guzman et al., "On-line event-driven hand gesture recognition based on surface electromyographic signals," in *Proceedings of the 2018 IEEE International Symposium on Circuits and Systems (ISCAS)*, IEEE, Florence, Italy, May 2018.

[8] V. Camomilla, E. Bergamini, S. Fantozzi, and G. Vannozzi, "Trends supporting the in-field use of wearable inertial sensors for sport performance evaluation: a systematic review," *Sensors*, vol. 18, no. 3, p. 873, 2018.

[9] A. Ancillao, S. Tedesco, J. Barton, and B. O'Flynn, "Indirect measurement of ground reaction forces and moments by means of wearable inertial sensors: a systematic review," *Sensors*, vol. 18, no. 8, p. 2564, 2018.

[10] M. Guo, Z. Wang, N. Yang, Z. Li, and T. An, "A multisensor multiclassifier hierarchical fusion model based on entropy weight for human activity recognition using wearable inertial sensors," *IEEE Transactions on Human-Machine Systems*, vol. 49, no. 1, pp. 105–111, 2018.

[11] Z. Chao, F. Pu, Y. Yin, B. Han, and X. Chen, "Research on real-time local rainfall prediction based on MEMS sensors," *Journal of Sensors*, vol. 2018, Article ID 6184713, 9 pages, 2018.

[12] R. A. Fabio and G. E. Towey, "Cognitive and personality factors in the regular practice of martial arts," *The Journal of Sports Medicine and Physical Fitness*, vol. 58, no. 6, pp. 933–943, 2017.

[13] L. Podrigalo, S. Iermakov, V. Potop et al., "Special aspects of psycho-physiological reactions of different skillfulness athletes, practicing martial arts," *Journal of Physical Education and Sport*, vol. 17, pp. 519–526, 2017.

[14] S. Łuczak, R. Grepl, and M. Bodnicki, "Selection of MEMS accelerometers for tilt measurements," *Journal of Sensors*, vol. 2017, Article ID 9796146, 2017.

[15] M. Varanis, A. Silva, A. Mereles, and R. Pederiva, "MEMS accelerometers for mechanical vibrations analysis: a comprehensive review with applications," *Journal of the Brazilian Society of Mechanical Sciences and Engineering*, vol. 40, no. 11, pp. 1–18, 2018.

[16] L. G. Pagani, L. Guerinoni, L. Falorni, P. Fedeli, and G. Langfelder, "Investigation of gyroscopes mechanical and electronic phase drift with 2 μrad/$\sqrt{}$ hz resolution and 12 μrad/K accuracy," in *Proceedings of the 2020 IEEE International Symposium on Inertial Sensors and Systems (INERTIAL)*, March 2020.

[17] G. A. Sanders, L. K. Strandjord, J. Wu et al., "Development of compact resonator fiber optic gyroscopes," in *Proceedings of the 2017 IEEE International Symposium on Inertial Sensors and Systems (INERTIAL)*, IEEE, Kauai, HI, USA, March 2017.

[18] A. D. Di Virgilio, N. Beverini, G. Carelli, D. Ciampini, F. Fuso, and E. Maccioni, "Analysis of ring laser gyroscopes including laser dynamics," *The European Physical Journal C*, vol. 79, no. 7, pp. 1–8, 2019.

[19] W. S. Noble, "What is a support vector machine?" *Nature Biotechnology*, vol. 24, no. 12, pp. 1565–1567, 2006.

[20] B. E. Boser, I. M. Guyon, and V. N. Vapnik, "A training algorithm for optimal margin classifiers," in *Proceedings of the 5th Annual Workshop on Computational Learning Theory*, Pittsburgh, PA, USA, July 1992.

[21] E. E. Osuna, "Support vector machines: training and applications," Doctoral Dissertation, Massachusetts Institute of Technology, Cambridge, MA, USA, 1998.

[22] O. L. Mangasarian and D. R. Musicant, "Active support vector machine classification," in *Proceedings of the 13th International Conference on Neural Information Processing Systems NIPS*, MIT Press, Cambridge, MA, USA, April 2000.

[23] O. L. Mangasarian and D. R. Musicant, "Lagrangian support vector machines," *Journal of Machine Learning Research*, vol. 1, pp. 161–177, 2001.

Hindawi
Scientific Programming
Volume 2021, Article ID 1651560, 9 pages
https://doi.org/10.1155/2021/1651560

Research Article

Music Feature Extraction and Classification Algorithm Based on Deep Learning

Jingwen Zhang

Music and Dance, Xi'an Peihua University, Xi'an Province 710199, China

Correspondence should be addressed to Jingwen Zhang; zhangjingwenpeihua@163.com

Received 5 April 2021; Revised 26 April 2021; Accepted 28 April 2021; Published 26 May 2021

Academic Editor: Shah Nazir

With the rapid development of information technology and communication, digital music has grown and exploded. Regarding how to quickly and accurately retrieve the music that users want from huge bulk of music repository, music feature extraction and classification are considered as an important part of music information retrieval and have become a research hotspot in recent years. Traditional music classification approaches use a large number of artificially designed acoustic features. The design of features requires knowledge and in-depth understanding in the domain of music. The features of different classification tasks are often not universal and comprehensive. The existing approach has two shortcomings as follows: ensuring the validity and accuracy of features by manually extracting features and the traditional machine learning classification approaches not performing well on multiclassification problems and not having the ability to be trained on large-scale data. Therefore, this paper converts the audio signal of music into a sound spectrum as a unified representation, avoiding the problem of manual feature selection. According to the characteristics of the sound spectrum, the research has combined 1D convolution, gating mechanism, residual connection, and attention mechanism and proposed a music feature extraction and classification model based on convolutional neural network, which can extract more relevant sound spectrum characteristics of the music category. Finally, this paper designs comparison and ablation experiments. The experimental results show that this approach is better than traditional manual models and machine learning-based approaches.

1. Introduction

With the rapid development of multimedia and digital technologies [1–3], there are more and more digital music resources on the Internet, and consumers' music consumption habits have shifted from physical music to online music platforms. Massive music resources and a huge online music library stimulate users to generate a variety of complex music retrieval needs. For example, at a certain moment, users are eager to listen to a certain genre or a song with a certain emotion. At this time, the music label is essential to the quality of music retrieval. In addition to music retrieval, many recommendation and subscription scenarios also require music category information to provide users with more accurate content [4, 5].

Music is diverse; it is made of different elements such as melody, rhythm, and harmony combinations according to certain rules of art forms. Understanding the music of different forms often requires some background knowledge, not as a music classification standard, so almost all music media platforms use text labels as the basis of the classification of music or retrieval. Music labels are text descriptors that express musical properties in high dimensions, such as "happy" and "sad" to express emotions, and "electronic" and "blues" to express musical styles [6, 7].

Music genre classification [8–10] is an important branch of music information retrieval. Correct music classification is of great significance for improving the efficiency of music information retrieval. At present, music classification mainly includes text classification and classification based on music content. Text classification is mainly based on music metadata information, such as singer, lyrics, songwriter, age, music name, and other labeled text information. The advantages of this classification method are easy to implement,

simple to operate, and fast to retrieve, but the shortcomings are also obvious. First of all, this method relies on manually labeled music data, which requires a lot of manpower, and manual labeling is difficult to avoid incorrectly labeling music information problems. Secondly, this text method does not involve the audio data of the music itself. Audio data includes many key characteristics of music, such as pitch, timbre, and melody. These characteristics are almost impossible to label with text; and based on the classification of the content, the features of the original music data are extracted, and the extracted feature data are used to train the classifier, so as to achieve the purpose of music classification. Therefore, music classification based on content has also become a research hotspot in recent years. Based on this, the research direction of this article is also based on content-based music classification [11].

The emergence of deep learning has brought music classification technology into a new period of development. Deep learning has been widely used in image processing, speech recognition, and other fields, and its performance on many tasks surpasses traditional machine learning methods. Scholars have also begun to use deep learning technology to study related issues in the field of music information retrieval, so it is necessary to research music classification methods based on deep learning to improve the effect of music classification [12]. The following are the main innovation points of this paper:

(i) The paper aims to convert the audio signal of music into a sound spectrum as a unified representation, avoiding the problem of manual feature selection.

(ii) It aims to use 1D convolution, gating mechanism, residual connection, and attention mechanism, and it proposes a music feature extraction and classification model based on convolutional neural network, which can extract and correlate more closely related sound spectrum features.

(iii) Sufficient comparison and ablation experiments have been carried out. The experimental results have proved the effectiveness and superiority of our algorithm, surpassing several other well-known methods.

The organization of the paper is as given. Section 2 depicts the background knowledge of the proposed study. The methodology of the paper is shown in Section 3 with the details in the subsections. Experiments and results are presented in Section 4. The paper is concluded in Section 5.

2. Background

As a very important component in the field of music information retrieval, music feature extraction and classification recognition have been widely studied since the 1990s. In 1995, Benyamin Matityaho and Furst [13] proposed a method for frequency-domain analysis of music signals. First, fast Fourier transform is performed on the audio data, and then the logarithmic scale transformation is performed to use the obtained data as feature data. Training was done in a neural network [14–17] containing two hidden layers and two music genres were finally identified: classical and pop music. Tzanetakis and Cook [18] systematically proposed in 2002 the division of the characteristics of music into three feature data sets, namely, timbre texture characteristics, rhythm content characteristics, and tonal content characteristics; the authors adopted the Gaussian mixture model and *K*. The proximity method is used as a classifier. It is worth mentioning that, due to the numerous music genres, there was no relatively fixed classification standard in the academic circles before. Since the groundbreaking research results of Tzanetakis, the ten music genres contained in the GTZAN data set used by George Tzanetakis have become music information. The classification standard was generally recognized in the search field.

As George Tzanetakis' research results laid a lot of foundation for us, later scholars in the field of automatic recognition of music genres mainly focused on two aspects. On the one hand, they made corresponding improvements in the selection of music feature extraction and the dimension of feature vectors. On the other hand, they improved the choice of classification algorithm. The extraction of music features is a very critical part of music genre recognition. If the extracted features cannot represent the essential characteristics of music, then the music classification effect will undoubtedly be very bad. Scaringella et al. [19] divided the music signal characteristics into three categories: pitch, timbre, and rhythm. At present, the commonly used characteristics of music signals mainly include short-term zero-crossing rate, short-term energy, linear prediction coefficient, frequency spectrum, flux, Mel frequency inverse coefficient, spectral centroid, and spectral contrast. Since these characteristics are both in the time domain and in the frequency domain, they can reflect the musical perception characteristics of pitch, rhythm, timbre, and loudness to some extent. The process of music feature extraction is generally to first perform frame processing on the original audio signal, then perform related calculations based on the mathematical statistical significance of the features, and finally use the calculated results as the training data of the classifier in the form of vectors. Because music feature extraction is based on music signal analysis, the current audio-based music signal analysis techniques mainly include time-domain analysis methods and frequency-domain analysis methods. The so-called time-domain analysis method is to analyze and count the waveform state of the music signal from the time dimension. Frequency-domain analysis converts the music signal in the time domain into the frequency domain through Fourier transform, so many useful features in the frequency domain can be obtained, for example, Mel to general coefficient, spectral centroid, pitch frequency, subband energy, spectrogram, etc. Literature [20] cascades together the Mel-to-Pop coefficient and pitch frequency, spectral centroid, subband energy, and other perceptual characteristics to form a high-dimensional feature vector. In the music classification algorithm, traditional machine learning methods are mainly used, such as support vector machines, Gaussian mixture models, decision trees, nearest neighbors, hidden Markov, and artificial neural

networks [21–24]. In addition, there are some improvements to the above algorithms. For example, literature [25] adds a genetic algorithm to the Gaussian mixture model, which improves the accuracy of classification from the experimental results.

3. Methodology

3.1. Music Signal Features

3.1.1. Spectral Centroid. The spectral centroid is a metric used to characterize the frequency spectrum in digital signal processing. It indicates where the "centroid" of the frequency spectrum is located. It feels that it has a close relationship with the brightness of the sound. Generally speaking, the smaller the value is, the more energy is concentrated in the low frequency range. Since the spectral centroid can better reflect the brightness of the sound, it is widely used in digital audio and music signal processing. It is used as a measure of the timbre of music. Its mathematical definition is as follows:

$$C_t = \frac{\sum_{n=1}^{N} M_t[n] * n}{\sum_{n=1}^{N} M_t[n]}, \tag{1}$$

where $M_t[n]$ represents the magnitude of the Fourier transform of the t-th frame at the frequency group n.

3.1.2. Spectral Flux. The spectrum flux is generally a measure of the rate of change of the signal spectrum. It is calculated by comparing the spectrum of the current frame with the spectrum of the previous frame. More precisely, it is usually calculated as the 2-norm between two normalized spectrums. Since the spectrum is normalized, the spectrum flux calculated in this way does not depend on the phase; only the amplitudes can be compared. Spectrum flux is generally used to determine the timbre of an audio signal or to determine whether to pronounce. Its mathematical definition is as follows:

$$F_t = \sum_{n=1}^{N} \left(N_t[n] - N_{t-1}[n]\right). \tag{2}$$

3.1.3. Spectral Contrast. Spectral contrast is a feature used to classify music genres. Spectral contrast is expressed as the difference in decibels between peaks and valleys in the frequency spectrum, which can represent the relative spectral characteristics of music. It can be seen from the experimental results of the literature [26] that the spectral contrast has a good ability to discriminate music genres.

3.1.4. Mel-Scale Frequency Cepstral Coefficients. Since the cochlea has filtering characteristics (as shown in Figure 1), different frequencies can be mapped to different positions of the basilar membrane. So the cochlea is often regarded as a filter bank. Based on this feature, psychologists obtained a set of filter banks similar to the cochlear effect through psychological experiments, that is, the Mel frequency filter bank. Since the sound level perceived by the human ear is not linearly related to its frequency, researchers have proposed a new concept called Mel frequency. The Mel frequency scale is more in line with the auditory characteristics of the human ear. The relationship between Mel frequency and frequency f is as follows:

$$f_{\text{mel}} = 25951 g\left(1 + \frac{f}{700}\right), \tag{3}$$

where f_{mel} is the converted Mel frequency, f is the frequency, and the unit is Hz.

Firstly, the audio signal is divided into frames, pre-emphasized, and then windowed, and then short-time Fourier transform (STFT) is performed to obtain its frequency spectrum. Secondly, set the Mel filter bank of L channels on the Mel frequency. The L value is determined by the highest frequency of the signal, generally 12 to 16, and each Mel filter has the same interval on the Mel frequency. Let $o(l)$, $c(l)$, and $h(l)$ be the lower limit frequency, center frequency, and upper limit frequency of the l-th triangular filter, respectively; then, the relationship between the three frequencies of adjacent triangular filters is as follows:

$$c(l) = h(l-1) = o(l+1). \tag{4}$$

Pass the linear amplitude spectrum of the signal through the Mel filter to get the output of the filter:

$$Y(l) = \sum_{k=o(l)}^{h(l)} W_l(k)\left|X_m(k)\right|, \quad l = 1, 2, \ldots, L. \tag{5}$$

The frequency features of the filter are

$$W_l(k) = \begin{cases} \dfrac{k - o(l)}{c(l) - o(l)}, & o(l) \le k \le c(l), \\[2ex] \dfrac{h(l) - k}{h(l) - c(l)}, & c(l) \le k \le h(l). \end{cases} \tag{6}$$

Take the natural logarithm of the filter output value, and then transform the discrete cosine to MFCC. The expression is as follows:

$$\text{MFCC}_{\text{MFCC}}(n) = \sum_{l=1}^{L} \lg Y(l) * \cos\left[\pi(l - 0.5)\frac{n}{L}\right], \quad n = 1, 2, \ldots, L. \tag{7}$$

3.2. 1D Residual Gated Convolutional Neural Model

3.2.1. Selection of Convolution Kernel. Convolutional neural networks can well identify potential patterns in the data. By superimposing convolution kernels to perform repeated convolution operations, more abstract features can be obtained in the deep layers of the network. One-dimensional convolution is often used to deal with problems related to time series. Unlike two-dimensional convolution that

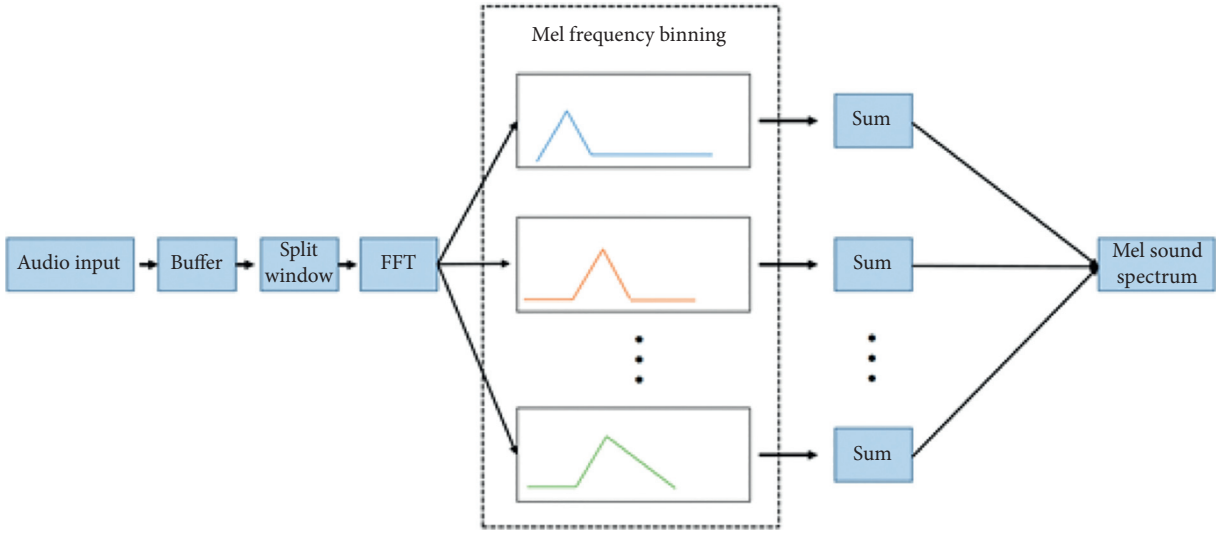


Figure 1: The calculation process of Mel sound spectrum.

attempts to convolve in multiple directions, one-dimensional convolution focuses more on capturing the translation invariance of data features in a specific direction. When dealing with time-related data, this direction is often the direction of time change. One-dimensional convolution is often used to analyze time series or sensor data and is suitable for signal data analysis within a fixed period of time, such as audio signals.

Figure 2 shows the convolution process of one-dimensional convolution and two-dimensional convolution on the sound spectrum. It can be seen that the receptive field of the one-dimensional convolution kernel covers all frequency ranges on the sound spectrum, which is only performed on the time axis. Convolution can capture the percussion components of the musical instruments appearing on the sound spectrum, and their overtones and other musical elements. Unlike the one-dimensional convolution that only convolves in the time direction, the two-dimensional convolution performs convolution in the two dimensions of time and frequency and can extract specific patterns of frequency within a certain time range, such as the rise and fall of pitch. In the field of music classification, many models use two-dimensional convolution as the basic convolution structure of convolutional networks.

The time perception of two-dimensional convolution is not as good as one-dimensional convolution, and the range of perception in the frequency range is not as broad as one-dimensional convolution, and the computational complexity of one-dimensional convolutional neural networks is smaller. In addition, two-dimensional convolution also performs convolution in the frequency dimension of the sound spectrum, which is inexplicable for sound signals. Therefore, the model in this article will use one-dimensional convolution as the basic convolution structure, which is more in line with the fact that the audio signal is expanded in time and has less correlation in the frequency range.

The essential difference between one-dimensional convolution and two-dimensional convolution lies in the translation direction, and its calculation method is not essentially different from that of two-dimensional convolution. Although the original audio signal is a time series, after it is converted into a sound spectrum, its expression is similar to a single-channel grayscale picture, so the calculation of convolution can be expressed by the following equation:

$$a_{ij} = h\left(\sum_{m=0}^{f_w-1}\sum_{n=0}^{f_h-1} w_{mn}x_{i+m,j+n} + b\right), \tag{8}$$

where a_{ij} is the width and height of the feature map, h is the activation function used by the convolution layer, f_w is the width of the convolution kernel, f_h is the height of the convolution kernel, b is the offset of the convolution, and w and x represent the weight matrix and data input of the product core, respectively. In the one-dimensional convolution operation based on the sound spectrum, f_h and the frequency range l of the sound spectrum have the following relationship:

$$l = f_h. \tag{9}$$

That is, the height of the convolution kernel in one-dimensional convolution is equal to the frequency range in the sound spectrum, and the receptive field of the convolution kernel covers the entire frequency axis, so as to capture a specific frequency pattern. Then the convolution operation can be expressed as

$$\mathrm{conv}(X, W) = \sum_{m=0}^{F_w-1}\sum_{n=0}^{L-1} w_{mn}x_{i+m,j+n}. \tag{10}$$

Assuming that the output of the convolution kernel is R and the bias matrix is B, then the convolution operation can be simply expressed as

$$R = \mathrm{conv}(X, W) + B. \tag{11}$$

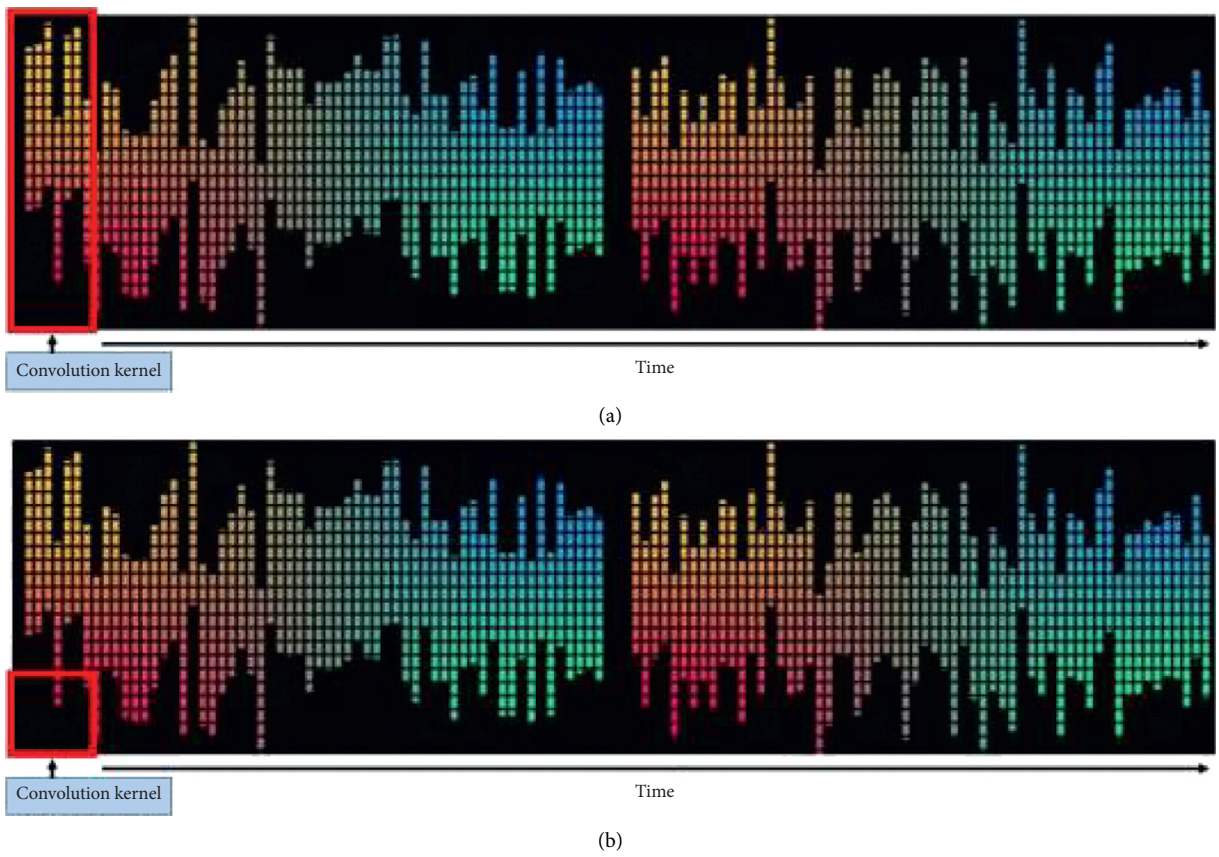


Figure 2: Convolution process comparison of 1D convolution and 2D convolution on the sound spectrum. (a) 1D conv. (b) 2D conv.

The width R_w of R can be obtained by the following formula:

$$R_w = \frac{t - f_w + 2p}{s} + 1, \tag{12}$$

where t represents the length of the sound spectrum on the time axis, that is, the width of the sound spectrum. p represents the size of the padding, and f_w represents the width of the convolution kernel. Since the one-dimensional convolution only performs translation in the time dimension of the sound spectrum, the height R_h of the output feature map R is as follows:

$$R_h = 1. \tag{13}$$

In other words, R_h has nothing to do with the frequency range l of the acoustic spectrum and the high f_h of the convolution kernel. After one-dimensional convolution, the dimension of the acoustic spectrum changes to that of the two-dimensional convolution. After one-dimensional convolution, the specification of the feature graph is also reduced.

3.2.2. Gated Linear Units. Assuming that the sound spectrum sequence to be processed is $X = [x_1, x_2, \ldots, x_n]$ and the output of the convolution kernel is Y, then the gated linear unit can be expressed as

$$Y = \mathrm{Conv1D}_1(X) \otimes \sigma\left(\mathrm{Conv1D}_2(X)\right). \tag{14}$$

The two $\mathrm{Conv1D}_1$ and Conv1D in the above formula represent two identical one-dimensional convolutions, but the weights are not shared. $\otimes$ represents the (element-wise) operation, and σ represents the Sigmoid activation function. One of the results after the two convolutions is activated by the Sigmoid function, and the other is not added with the activation function, and then the creation gate is multiplied bit by bit. Formally, it is equivalent to adding a "valve" to each output of one-dimensional convolution to control the flow. The convolution-based gating mechanism is different from the complex threshold mechanism in the LSTM network. It does not need to forget the gate, only an input gate, which also makes the network model based on the gated convolution unit perform better than LSTM in training speed.

Figure 3 shows the basic structure of the one-dimensional gated convolution unit. You can see the data flow inside the one-dimensional gated convolution unit. After the input of the convolution unit undergoes two identical convolutions, one of the 1D convolution kernels is extra. The activation operation of the Sigmoid function is performed, and then the output of another convolution kernel is multiplied bit by bit to produce the output of this layer.

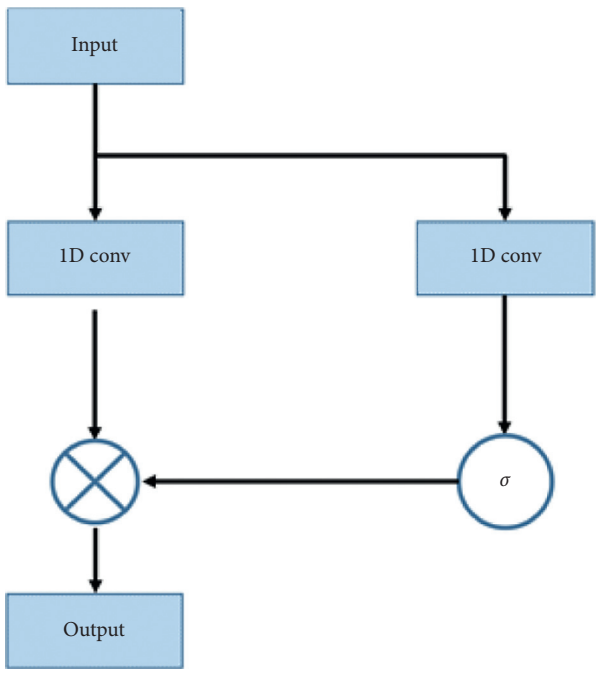


Figure 3: Schematic diagram of 1D convolutional gated unit.

3.2.3. Residual Connection. The entire convolutional neural network can be regarded as a process of information extraction. The more the layers of the network, the stronger the ability of the network to gradually extract from the underlying features to the highly abstract features. When the network layers are deepened, the model is more likely to discover high-level abstract features related to music categories. Increasing the depth of the network too much will cause gradient disappearance and explosion problems to the model. The solution to gradient disappearance and explosion is generally to add regular initialization and an intermediate regularization layer, but the network degradation problem also arises. When the network begins to degenerate, the accuracy on the training set will decrease as the number of network layers increases. This problem is essentially different from overfitting, which will show excellent results on the training set.

The basic residual module is shown in Figure 4. It can be seen that the residual structure has an additional identity mapping channel, so that when the depth of the network increases and it is not conducive to the enhancement of network performance, the network can directly skip these useless layers. Directly accept the output of the upper layer. The calculation equation of the residual structure is as follows:

$$x_{l+1} = x_l + F(x_l, W_l). \tag{15}$$

3.3. Music Feature Extraction and Classification Model. The model in this paper can be divided into GLU stacking layer, global pooling feature aggregation layer, and fully connected layer from input to output. The overall structure of the network is shown in Figure 5.

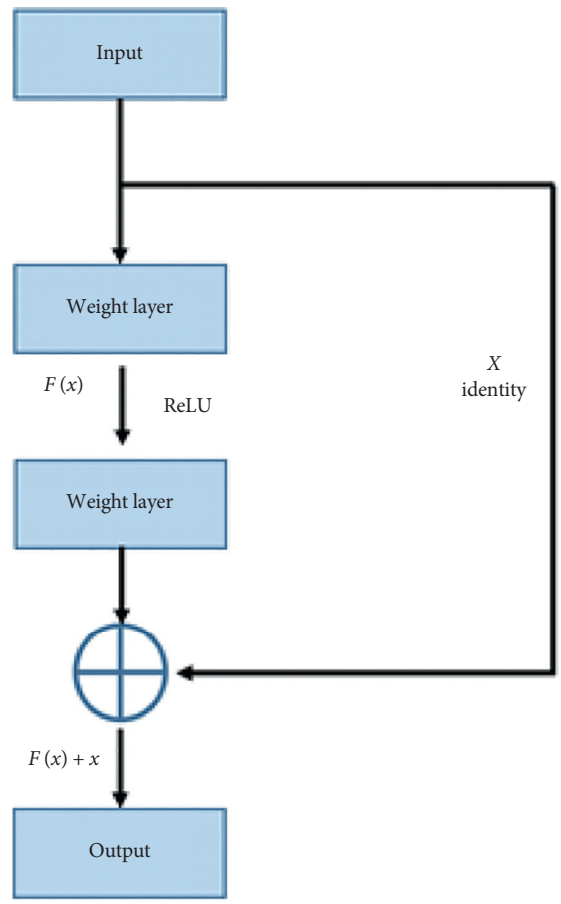


Figure 4: Schematic diagram of residual module.

To make full use of the statistical information of the pooling layer, the model in this chapter combines the global maximum pooling and the global average pooling to form a global pooling feature aggregation layer. The feature maps obtained from the GLU block stacking layer undergo global average pooling and global maximum pooling to obtain average pooling statistics and maximum pooling statistics, respectively. The results of the pooling operation here are all one-dimensional. In Figure 5, two rectangular blocks of different colors are used to represent these two one-dimensional features, and the two pooled statistical features are spliced into the next layer of fully connected network.

4. Experiments and Results

4.1. Experimental Setup. Due to the repetitive information in the multichannel of the original audio, all audio is converted to mono, and downsampling is performed at a sampling rate of 16 kHz. The Fourier transform window length used when converting the Mel sound spectrum is 512, the window jump size is 256, and the number of frequency bins is 128. The original audio sample is segmented by the segmentation

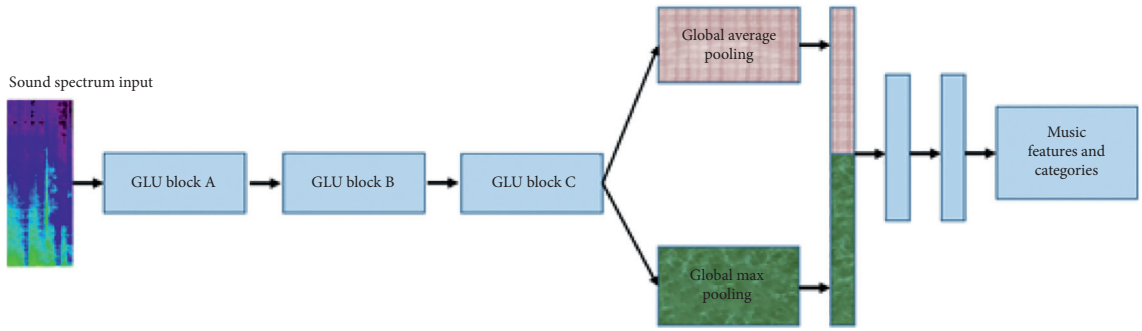


Figure 5: Schematic diagram of the overall model.

method. The slice duration is 5 seconds, and the overlap rate is 50%. The Mel sound spectrum specification of a single slice generated after processing according to the above settings is (313, 128), and each audio sample produces 11 slices of the same size.

4.2. Data Set. The experiment in this chapter uses the GTZAN data set, which is widely used to verify the performance of music classification methods and is the most popular music classification data set. The GTZAN data set has 10 music genre categories (as shown in Table 1). The number of audio samples in each genre category is 100, the sample duration is 30 seconds, and the sampling rate is 22050 Hz.

4.3. Evaluation Index. The classification accuracy rate (Acc) is selected as the evaluation index of the music classification method proposed in this chapter. The calculation method of classification accuracy is as follows:

$$\text{Acc} = \frac{N_C}{N} \times 100\%. \tag{16}$$

4.4. Experimental Results. Different models produce different recognition results by learning different deep features. In order to make a fair comparison, all experiments were implemented in the same environment, and all parameters were retained, comparing the proposed model with SVM, CNN, GLU, RCNN, and RGLU.

The above five types of networks with different structures are tested with the same experimental settings, and the results are shown in Table 2 and Figure 6. The GLU network using the gated structure has higher accuracy than the ordinary convolutional network CNN, which indicates that the stacking of multiple gated convolutions used in the model in this chapter is more conducive to the sound spectrum characteristics than the ordinary convolution learning. The gating structure makes the features passed to the next layer of the network pay more attention to the sound spectrum features that are more important for the music classification task, and the information that is not related to the music classification task is ignored by the gating mechanism. The results of the comparison experiment verify that the gating structure is based on the effectiveness of the sound spectrum

Table 1: Introduction to the GTZAN data set.

ID	Category
1	Rock
2	Reggae
3	Pop
4	Metal
5	Jazz
6	Hip hop
7	Disco
8	Country
9	Classical
10	Bruce

Table 2: Comparative experiment results on the GTZAN data set.

Method	Acc	Std
SVM	0.52	0.03
CNN	0.70	0.01
GLU	0.75	0.03
RCNN	0.80	0.01
RGLU	0.82	0.02
Ours	**0.87**	**0.01**

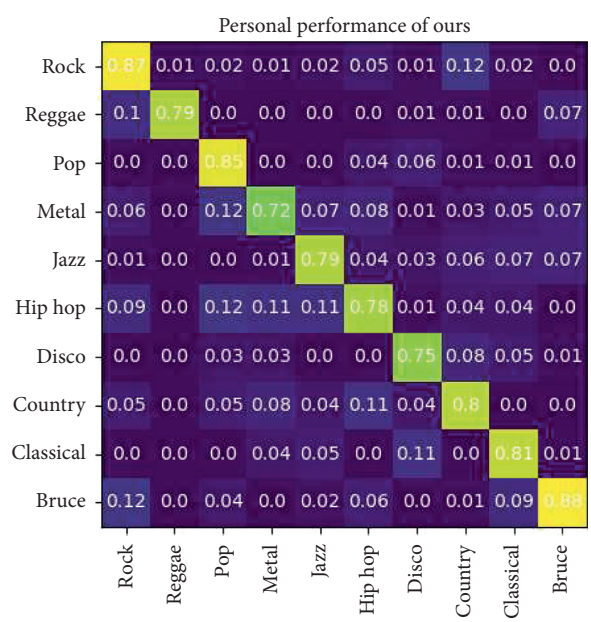


Figure 6: Confusion matrix.

Table 3: Results of ablation studies.

Method	Acc
Global average pooling	0.85
Global max pooling	0.84
Ours	**0.87**

in the task of music classification. From the perspective of information filtering, GLU can be used as another implementation of the attention mechanism. Unlike the RGLU structure that determines an attention weight for each feature map channel, GLU can adaptively determine the time during the network learning process. The attention weight in the one-dimensional convolution is expanded in time; this kind of gated structure that increases attention in the time dimension, combined with the one-dimensional convolution in the time dimension, can get better performance. Compared with CNN and GLU without residual structure, the accuracy of RCNN and RGLU with added residual structure has been improved, which shows that the use of residual connection can improve the accuracy of classification to a certain extent. It is worth noting that the accuracy of RGLU using the residual structure is improved compared to GLU, and the accuracy of RCNN is greater than that of CNN. This indicates that the combination of residual structure and gated convolution is more beneficial for the transmission of information in the network. Therefore, this experiment fully proves the effectiveness and superiority of our algorithm.

4.5. Ablation Study of Global Pooling. This section will compare the classification performance of different pooling features and their combinations in the global pooling feature aggregation layer. We used three pooling methods to conduct experiments, and the experimental results are shown in Table 3.

The aggregation of the two global pooling features can make the model obtain a higher accuracy rate. Using the global average pooling feature alone is more accurate than using the global maximum pooling feature alone, which means the overall statistical information in the spectroscopic feature map is more conducive to classification. The model in this chapter combines two types of global pooling features, which enables the fully connected layer to grasp more statistical information of the features abstracted by the convolutional layer and makes the classification performance of the model stronger.

5. Conclusion

Digital music has grown and exploded with the growing developments of information technology and communication. Music feature extraction and classification are considered as a significant portion of music information retrieval. The design of features requires knowledge and in-depth understanding in the domain of music. The features of different classification tasks are often not universal and comprehensive. Traditional music classification approaches use a large number of artificially designed acoustic features. It is difficult to effectively extract music features due to manual and traditional machine learning methods. Therefore, the contribution of this paper is to convert the audio signal of music into a sound spectrum as a unified representation, avoiding the problem of manual feature selection. According to the characteristics of the sound spectrum, combined with one-dimensional convolution, gating mechanism, residual connection, and attention mechanism, a music feature extraction and classification model based on convolutional neural network is proposed, which can extract and correlate more closely related sound spectrum features of music category. Finally, this paper designs a comparison and ablation experiment. Experimental results show that this method is superior to traditional manual models and machine learning-based methods.

Data Availability

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

Conflicts of Interest

The author declares no conflicts of interest.

References

[1] A. El Saddik, "Digital twins: the convergence of multimedia technologies," *IEEE Multimedia*, vol. 25, no. 2, pp. 87–92, 2018.

[2] I. Gorbunova and H. Hiner, "Music computer technologies and interactive systems of education in digital age school," in *Proceedings of the International Conference Communicative Strategies of Information Society (CSIS 2018)*, pp. 124–128, Atlantis Press, Almaty, Kazakhstan, February 2019.

[3] C. C. S. Liem, E. Gómez, and G. Tzanetakis, "Multimedia technologies for enriched music performance, production, and consumption," *IEEE MultiMedia*, vol. 24, no. 1, pp. 20–23, 2017.

[4] Y. An, S. Sun, and S. Wang, "Naive Bayes classifiers for music emotion classification based on lyrics," in *Proceedings of the 2017 IEEE/ACIS 16th International Conference on Computer and Information Science (ICIS)*, pp. 635–638, IEEE, Wuhan, China, 2017 May.

[5] Y. R. Pandeya and J. Lee, "Deep learning-based late fusion of multimodal information for emotion classification of music video," *Multimedia Tools and Applications*, vol. 80, no. 2, pp. 2887–2905, 2021.

[6] M. Russo, L. Kraljević, M. Stella, and M. Sikora, "Cochleogram-based approach for detecting perceived emotions in music," *Information Processing & Management*, vol. 57, no. 5, Article ID 102270, 2020.

[7] M. Chełkowska-Zacharewicz and M. Paliga, "Music emotions and associations in film music listening: the example of leitmotifs from the Lord of the Rings movies," *Roczniki Psychologiczne*, vol. 22, no. 2, pp. 151–175, 2019.

[8] S. Oramas, F. Barbieri, O. Nieto, and X. Serra, "Multimodal deep learning for music genre classification," *Transactions of the International Society for Music Information Retrieval*, vol. 1, no. 1, pp. 4–21, 2018.

[9] S. Oramas, O. Nieto, F. Barbieri, and X. Serra, "Multi-label music genre classification from audio, text, and images using deep features," 2017, https://arxiv.org/abs/1707.04916.

[10] A. Rosner and B. Kostek, "Automatic music genre classification based on musical instrument track separation," *Journal of Intelligent Information Systems*, vol. 50, no. 2, pp. 363–384, 2018.

[11] M. D. S. Anisetty, G. Shetty, S. Hiriyannaiah, S. Gaddadevara Matt, K. G. Srinivasa, and A. Kanavalli, "Content-based music classification using ensemble of classifiers," in *Proceedings of the International Conference on Intelligent Human Computer Interaction*, pp. 285–292, Springer, Cham, Germany, 2018 December.

[12] J. Nam, K. Choi, J. Lee, S. Y. Chou, and Y. H. Yang, "Deep learning for audio-based music classification and tagging: teaching computers to distinguish rock from bach," *IEEE Signal Processing Magazine*, vol. 36, no. 1, pp. 41–51, 2018.

[13] B. Matityaho and M. Furst, "Neural network based model for classification of music type," in *Proceedings of the Eighteenth Convention of Electrical and Electronics Engineers in Israel*, pp. 4–3, IEEE, Tel Aviv, Israel, 1995 March.

[14] X. Ning, K. Gong, W. Li, L. Zhang, X. Bai, and S. Tian, "Feature refinement and filter network for person re-identification," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 12, p. 1, 2020.

[15] W. Cai, Z. Wei, R. Liu, Y. Zhuang, Y. Wang, and X. Ning, "Remote sensing image recognition based on multi-attention residual fusion networks," *ASP Transactions on Pattern Recognition and Intelligent Systems*, vol. 1, no. 1, pp. 1–7, 2021.

[16] S. Li, X. Ning, L. Yu et al., "Multi-angle head pose classification when wearing the mask for face recognition under the COVID-19 coronavirus epidemic," in *Proceedings of the 2020 International Conference on High Performance Big Data and Intelligent Systems (HPBD&IS)*, pp. 1–5, IEEE, Shenzhen, China, May 2020.

[17] W. Cai, B. Liu, Z. Wei, M. Li, and J. Kan, "TARDB-net: triple-attention guided residual dense and BiLSTM networks for hyperspectral image classification," *Multimedia Tools and Applications*, vol. 80, no. 7, pp. 11291–11312, 2021.

[18] G. Tzanetakis and P. Cook, "Musical genre classification of audio signals," *IEEE Transactions on Speech and Audio Processing*, vol. 10, no. 5, pp. 293–302, 2002.

[19] N. Scaringella, G. Zoia, and D. Mlynek, "Automatic genre classification of music content: a survey," *IEEE Signal Processing Magazine*, vol. 23, no. 2, pp. 133–141, 2006.

[20] M. McKinney and J. Breebaart, "Features for audio and music classification," 2003.

[21] X. Ning, F. Nan, S. Xu, L. Yu, and L. Zhang, *Multi-View Frontal Face Image Generation: A Survey. Concurrency and Computation: Practice and Experience*, Wiley, Hoboken, NJ, USA, 2020.

[22] X. Zhang, Y. Yang, Z. Li, X. Ning, Y. Qin, and W. Cai, "An improved encoder-decoder network based on strip pool method applied to segmentation of farmland vacancy field," *Entropy*, vol. 23, no. 4, p. 435, 2021.

[23] Z. L. Yang, S. Y. Zhang, Y. T. Hu, Z. W. Hu, and Y. F. Huang, "VAE-stega: linguistic steganography based on variational auto-encoder," *IEEE Transactions on Information Forensics and Security*, vol. 16, pp. 880–895, 2020.

[24] X. Ning, X. Wang, S Xu et al., *A Review of Research on co-training. Concurrency and Computation: Practice and Experience*, Beijing Union University, Beijing, China, 2021.

[25] K. I. Molla and K. Hirose, "On the effectiveness of MFCCs and their statistical distribution properties in speaker identification," in *Proceedings of the 2004 IEEE Symposium on Virtual Environments, Human-Computer Interfaces and Measurement Systems, 2004 (VCIMS)*, pp. 136–141, IEEE, Boston, MA, USA, July 2004.

[26] D. N. Jiang, L. Lu, H. J. Zhang, J. H. Tao, and L. H. Cai, "Music type classification by spectral contrast feature," in *Proceedings of the IEEE International Conference on Multimedia and Expo*, vol. 1, pp. 113–116, IEEE, Lausanne, Switzerland, August 2002.

Hindawi
Scientific Programming
Volume 2021, Article ID 9972589, 9 pages
https://doi.org/10.1155/2021/9972589

Research Article

K-Modes Clustering Algorithm Based on Weighted Overlap Distance and Its Application in Intrusion Detection

Yawen Dai,[1] Guanghui Yuan,[2] Zhaoyuan Yang,[3] and Bin Wang[4]

[1]*Institute of Rail Transit, Tongji University, Shanghai 201804, China*
[2]*School of Economics and Management, Shanghai University of Political Science and Law, Shanghai 201701, China*
[3]*School of Finance and Business, Shanghai Normal University, Shanghai 201418, China*
[4]*School of Humanities, Shanghai University of Finance and Economics, Shanghai 200433, China*

Correspondence should be addressed to Zhaoyuan Yang; 1853871@tongji.edu.cn

Received 3 April 2021; Revised 4 May 2021; Accepted 18 May 2021; Published 25 May 2021

Academic Editor: Shah Nazir

In order to better apply the K-modes algorithm to intrusion detection, this paper overcomes the problems of the existing K-modes algorithm based on rough set theory. Firstly, for the problem of K-modes clustering in the initial class center selection, an initial class center selection algorithm Ini_Weight based on weighted density and weighted overlap distance is proposed. Secondly, based on the Ini_Weight algorithm, a new K-modes clustering algorithm WODKM based on weighted overlap distance is proposed. Thirdly, the WODKM clustering algorithm is applied to intrusion detection to obtain a new unsupervised intrusion detection model. The model detects the intrusion by dividing the clusters in the clustering result into normal clusters and abnormal clusters and analyzing the weighted average density of the object x to be detected in each cluster and the weighted overlapping distance of x and each center point. We verified the intrusion detection performance of the model on the KDD Cup 99 dataset. The experimental results of the current study show that the proposed intrusion detection model achieves efficient results and solves the problems existing in the present-day intrusion detection system to some extent.

1. Introduction

For the threat of network security, there are many corresponding network security defence technologies. For example, traditional network security defence technologies include firewall, antivirus software, digital signature, digital authentication, and data encryption. Traditional network security defence technology is only a static defence method, and it is difficult to effectively protect our network security [1]. Intrusion detection helps systems count the number of errors in the system and analyze the causes of system vulnerabilities, detecting known intrusions and alarms, auditing of new anomalous behaviors, and integrity assessment of critical data files.

In view of the problems existing in existing intrusion detection systems, data mining techniques have been widely used in the field of intrusion detection in recent years. Among them, cluster-based unsupervised intrusion detection has caused a lot of attention. At present, many clustering algorithms have been applied to intrusion detection. As an effective extension of the K-means algorithm, the K-modes algorithm can effectively deal with categorical data. It inherits the efficient features of the K-means algorithm, and the algorithm is simple, easy to implement, and widely used in many fields. Therefore, the K-modes algorithm has a very broad application prospect in the field of intrusion detection. However, the research on applying K-modes algorithm to intrusion detection is still rare [2, 3].

In order to better apply the K-modes algorithm to intrusion detection, this paper solves the problems of the existing K-modes algorithm based on rough set theory and applies the improved algorithm to intrusion detection [4]. Firstly, for the problem of K-modes clustering in the initial class center selection, an initial class center selection algorithm Ini_Weight based on weighted density and

weighted overlap distance is proposed. Secondly, based on the Ini_Weight algorithm, a new K-modes clustering algorithm WODKM based on weighted overlap distance is proposed. Thirdly, the WODKM clustering algorithm is applied to intrusion detection to obtain a new unsupervised intrusion detection model UIDM_WODKM.

The model detects the intrusion by dividing the clusters in the clustering result into normal clusters and abnormal clusters and analyzing the weighted average density of the object x to be detected in each cluster and the weighted overlapping distance of x and each center point. We verified the intrusion detection performance of the model on the KDD Cup 99 dataset.

The organization of the paper is as follows: Section 2 describes the intrusion detection system. Section 3 briefly discusses the K-modes clustering algorithm based on weighted overlap distance. The application of K-modes clustering algorithm based on weighted overlap distance in intrusion detection is given in Section 4. The paper is concluded in Section 5.

2. Intrusion Detection

Intrusion detection is a proactive network security defence strategy that can make up for the shortcomings of traditional static security policies and thus becomes a reasonable complement to traditional static defence strategies such as firewalls [5]. System administrators' security management capabilities have been extended through auditing, monitoring, intrusion identification and response, and reducing the workload of system administrators.

Through the intrusion detection system, system administrators can grasp the status of the network system in real time, including the status of programs, files, and devices, and help network system administrators to develop accurate and complete strategies [6]. In addition, the intrusion detection system dilutes the restrictions on professionals in network security, making it easy for nonprofessional personnel to manage network systems. In addition, the intrusion detection system can respond to the discovered network intrusion, illegal operation, timely and proactive response, and achieve the purpose of active defence [7].

Intrusion detection system (IDS) is a collection of network security systems or systems that detect abnormal behaviors in the system or changes in network status and can alert or take proactive responses [8]. It differs from other network security systems in that it is also an active security protection technology [9–11]. Intrusion detection in an intrusion detection system is mainly divided into the following steps:

(1) Collect message: intrusion detection first collects the corresponding system and network information. The main content includes the content of network traffic, system health, status, and behavior of user connection activities.

(2) Signal analysis: the information collected above is generally analyzed by three technical means: pattern matching, statistical analysis, and integrity analysis. The first two methods are used for real-time intrusion detection, while the integrity analysis is used for postmortem analysis.

(3) Real-time recording, alarm, and response: finally, the intrusion detection system can timely respond to detected intrusions or other network attacks. This includes logging intrusion information and alerting the network administrator.

At present, there are many kinds of intrusion detection systems. According to different classification standards, we classify the current intrusion detection system as follows.

2.1. Classify according to the Principle of Detection

(1) Anomaly detection: anomaly detection first summarizes the characteristics of normal behavior and then judges whether the user behavior is intrusive according to the user's activities or the use of resources in the system. Anomaly detection needs to establish a model to determine the normal user behavior and activities before detection and then to determine which qualified behaviors can be marked as "exceptions" and to make corresponding processing.

(2) Abuse detection: misuse detection first establishes the feature library based on the behavioral characteristics of the abnormal operation collected. When the observed user activity or system resource usage matches the records in the feature library, it is considered an intrusion. When the intrusion behavior matches the normal user activity, the system produces a false alarm; when the feature library does not store the features matching with a new attack behavior, the system produces a false alarm. Therefore, the focus of abuse detection is how to update the feature library automatically to reduce the false alarm rate.

(3) Mixed detection: hybrid detection combines the advantages of the above two methods, mainly based on the normal data flow of the system to detect intrusion behavior. And before making a decision, we not only analyze the normal behavior of the system but also observe the suspicious intrusion behavior. Therefore, the results of mixed detection are more comprehensive. It is reliable.

2.2. Classify according to Data Sources

(1) Host-based intrusion detection system: host-based intrusion detection system obtains data from the host where the system runs. It mainly protects the security of the host where the system is located. Usually, this system is installed on the host which needs to be checked, mainly on the relatively critical system files and executable files.

(2) Network-based intrusion detection system: with the development of network technology, the host-based

intrusion detection system has been difficult to meet the needs of network security, and the network-based intrusion detection system came into being. The data source of this kind of system analysis is the data packet on the network. It mainly analyzes and detects the intrusion behavior on the corresponding network segment according to the data flow, data packet, and corresponding protocol on the network.

2.3. Classify according to the Architecture of the Intrusion Detection System

(1) Centralized intrusion detection system: centralized intrusion detection systems usually have multiple auditors on different hosts, but only one central server is used for intrusion detection. In this way, the audit program sends the collected data trace to the central server, which then analyzes and detects the data. The scalability and configurability of such intrusion detection systems are relatively poor.

(2) Distributed intrusion detection system: the distributed intrusion detection system (DIDS) sends the corresponding tasks to several different host-based IDS by the central detection server, and these host-based IDS are nonhierarchical. Each of them is responsible for monitoring suspicious activities on the corresponding host. This kind of intrusion detection system has relatively high scalability and security, but its maintenance cost is high and the workload of the host is relatively heavy.

3. K-Modes Clustering Algorithm Based on Weighted Overlap Distance

The schematic of the intrusion detection system is shown in Figure 1.

3.1. K-Modes Clustering Algorithm. The K-means algorithm is a clustering algorithm commonly used in the field of data mining. However, this algorithm can only process numeric data, but cannot process subtype data [12]. In response to this problem, Huang et al. further proposed the K-modes clustering algorithm.

Clustering is the process of dividing dataset into a plurality of groups or clusters composed of similar objects [13] The similarity between objects in the same group is made as high as possible, and the similarity of objects between different groups is as low as possible. Clustering is an unsupervised machine learning method that does not have any prior knowledge of the datasets that need to be classified [14]. Datasets are automatically divided into groups or clusters based solely on similarity metrics. Try to make the similarity between samples in the same group high and the sample similarity between different groups is low. The groups in the cluster do not need to be defined in advance and are automatically divided according to the inherent similarity of the data according to the actual characteristics of the data [13]. The cluster analysis system inputs the criteria for measuring the similarity between data and the dataset that needs to be classified, and the output is the result of the well-classified class. The additional result of cluster analysis is a comprehensive description of each group, which can be used to provide a more in-depth analysis of the characteristics of the dataset [15].

The K-means clustering algorithm is a popular clustering algorithm with simple and efficient characteristics, but it can only deal with numerical data. The K-modes clustering algorithm is an extension of the K-means algorithm, which can deal with class attribute data. It inherits the characteristics of the K-means clustering algorithm which is efficient and easy to implement, so it is widely used in various fields.

Distance measure (or similarity measure) is the most basic and important part of the K-modes clustering algorithm. As mentioned earlier, the existing geometric distance measurement for numeric data is not suitable for categorical data. It is necessary to design some new distance measurement mechanisms according to the characteristics of categorical data. At present, in the K-modes clustering algorithms, most of them use simple overlap distance to measure the dissimilarity between any two objects X and Y; that is, the number of attributes with different values of X and Y is taken as the distance between X and Y. The process of clustering and segmentation is shown in Figure 2.

When calculating the overlap distance between any two objects, we mainly get the final result by comparing the values of the two objects on each attribute. If we use the same way to deal with all the attributes in the information table without distinguishing them, this approach obviously does not conform to objective reality and eventually will lead to the clustering results deviation, thus affecting the performance of the K-modes algorithm.

The K-modes algorithm uses a simple overlapping distance measure to calculate the distance between objects [4, 16]. The specific definition of the distance measure is as follows: for any two objects X and y, X and Y in the universe U, the distance $d(x, y) = \sum_{a \in A} \delta_a(x, y)$, which for any $a \in A$, $\delta_a(x, y)$ denotes a simple overlap distance between X and Y on the conditional property a, that is, if $f(x, a) = f(y, a), \delta_a(x, y) = 0$; otherwise, $\delta_a(x, y) = 1$.

The schematic of the K-means algorithm is shown in Figure 3.

Obviously, the traditional K-modes clustering algorithm has the following problems [17–20]:

(1) Select the initial center point by random selection: choosing the initial center point in this way will lead to the instability of clustering results. If we want to get relatively good clustering results, we need to repeatedly execute the K-modes clustering algorithm and select the best group of data from it. For intrusion detection, it is inappropriate to select the initial center point by using the above method because the data to be processed by the network intrusion detection system is usually very huge. Repeated K-modes

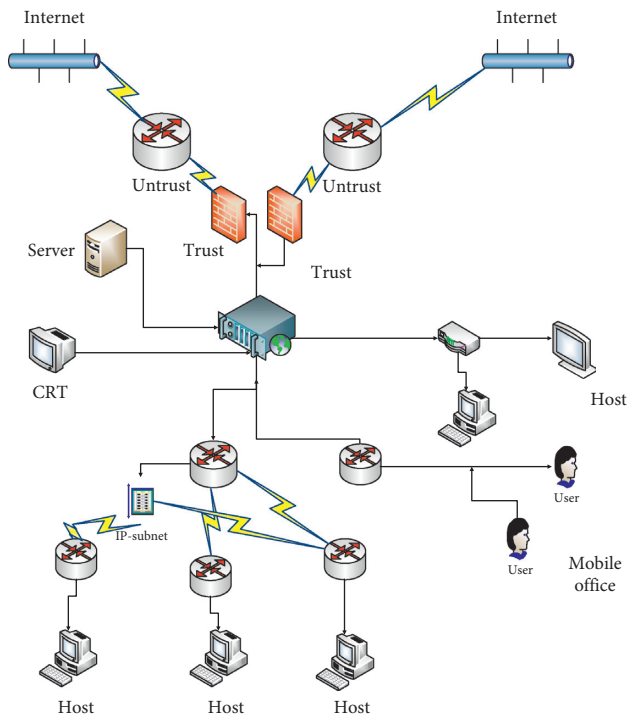


Figure 1: Intrusion detection detecting system.

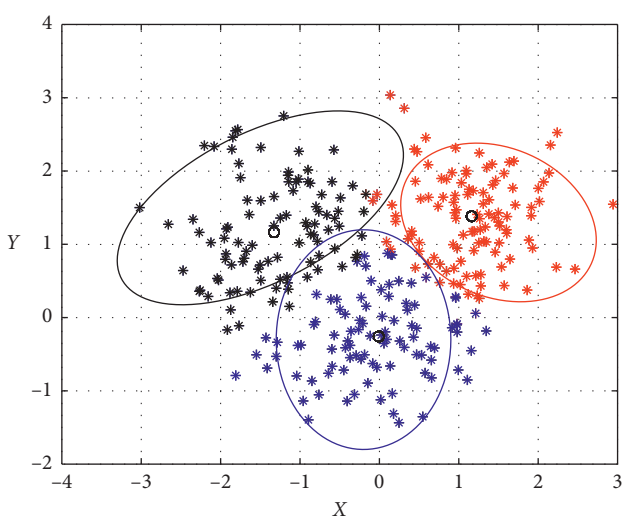


Figure 2: Process of clustering and segmentation.

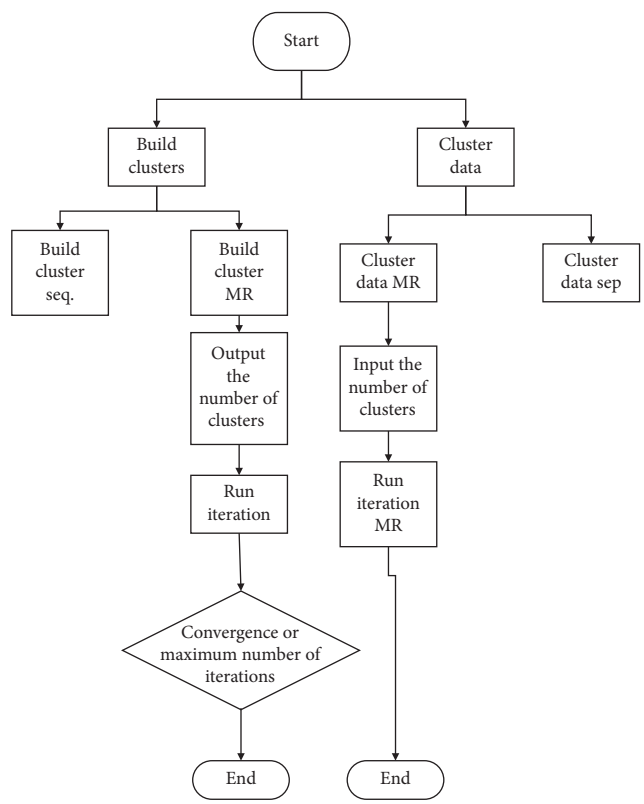


Figure 3: K-modes algorithm.

clustering algorithms on massive, high-dimensional network data are often not feasible in time.

(2) Traditional K-modes algorithms usually use simple overlap distance to measure the dissimilarity between objects X and y. Simple overlapping distance is simple and intuitive, but there are still some problems in its practical application. For example, it assumes that all attributes play the same role in calculating the distance between objects; that is, each attribute has a weight equal to 1. However, in many practical cases, the impact of different attributes on distance calculation is likely to be different. Therefore, it is necessary to assign a weight value to each attribute and through the size of the weight value to reflect the contribution of each attribute to distance calculation.

3.2. Weighted Overlap Distance. Aiming at the problems existing in the traditional K-modes clustering algorithm, this paper proposes a new K-modes clustering algorithm WODKM based on weighted overlap distance. Different from the traditional K-modes clustering algorithm, the WODKM algorithm uses the initial class center selection algorithm Ini_Weight based on weighted density and weighted overlap distance proposed in Chapter 3 to select the initial center point. In addition, the WODKM algorithm uses a weighted overlap distance metric instead of a simple overlap distance metric when calculating the distance between objects and the distance between the object and the center point (Algorithm 1).

The specific steps of the WODKM algorithm are in Algorithm 1.

In the worst case, the time complexity of steps (1)–(5) in the Ini_Weight algorithm is O $(|A|2 \times |U|)$, The time complexity of steps (6) – (10) is O $(K2 \times |A| \times |U|)$, where K is the number of clusters. Therefore, in the worst case, the time complexity of the algorithm is O $(K2 \times |A| \times |U| + |A|2 \times |U|)$, and the space complexity is O $(|A| \times |U\ |)$. Obviously, the time complexity of Ini_Weight is linearly relative to the number of objects.

To evaluate the performance of the Ini_Weight algorithm, we experimented with the following four subtype UCI datasets. The basic information of the above four datasets is shown in Table 1.

We compare the Ini_Weight algorithm with four existing initial center selection methods: (1) random method [21]; (2) Cao's algorithm [22]; (3) Wu's algorithm [23]; and (4) Khan's algorithm. The main steps of the experiment are as follows: first, different initialization algorithms are used to select the initial center; secondly, K-modes clustering is performed based on the initial center selected in the previous step (using the classical K-modes algorithm proposed by Huang for clustering); finally, compare the clustering results corresponding to different initialization algorithms. Table 2 shows the K-modes clustering results for different initialization algorithms on the four datasets Soybean, Zoo, Breast, and Mushroom.

There is

$$\mathrm{PR} = \frac{\sum_{i=1}^{K} a_i/a_i + b_i}{K},$$
$$\mathrm{RE} = \frac{\sum_{i=1}^{K} a_i/a_i + c_i}{K}, \tag{1}$$
$$\mathrm{AC} = \frac{\sum_{i=1}^{K} a_i}{|U|}.$$

From Table 2, we can see that the Ini_Weight algorithm performs better than the random method because the AC, PR, and RE of Ini_Weight outperform the random method on the three datasets of Soybean, Breast, and Mashroom. On Zoo, although Ini_Weight's PR is slightly worse than the random method, it is significantly superior to the random method in AC and RE. In addition, the clustering results obtained by random methods are different each time, and our method can get stable clustering results.

Although Cao's method and Wu's method are also superior to random methods, Ini_Weight is superior to both methods. On Soybean, Breast, and Mashroom, Ini_Weight yields AC, PR, and RE higher than or equal to these two methods. On Zoo, although Ini_Weight's PR is better than Cao, Wu is low, but it is higher than those 2 methods on AC and RE. In addition, Ini_Weight performs better than Khan's because on any data set, Ini_Weight yields higher AC, PR, and RE than Khan's.

We can also see that Ini_Weight is always higher than other methods on RE. This shows that our method can strictly control the objects in a cluster without being wrongly assigned to other clusters. In addition, our method produces poorer PR on Zoo than Cao, Wu, and random methods, but better on AC. This is because PR and AC are two different performance metrics. In the clustering results obtained by Ini_Weight, one of the clusters has zero accuracy (i.e., no objects in Zoo are allocated to this cluster correctly), so that the PR obtained by Ini_Weight is lower than that by Cao, Wu, and random methods. However, although the accuracy of the cluster is zero, the number of objects belonging to this cluster in Zoo is very small, so Ini_Weight is still higher on AC than Cao, Wu, and random methods. We also compared the accuracy of several algorithms, as shown in Figure 4.

4. Application of K-Modes Clustering Algorithm Based on Weighted Overlap Distance in Intrusion Detection

We used the network test environment shown in Figure 5 for testing.

Under normal circumstances, Host 5 uses Tcpdump to collect 10 minutes of network packets every 1 h. The data collection was performed six times in succession. The first five sets of data were merged to form the standard normal behavior set dataNormal, and the sixth set of datasets was the normal behavior set dataNatural. Then the distributed denial of service attack tool trino is installed on host Host 1, and trino uses master to control host Host 2 and host Host 3 to launch an attack on host Host 5. The network packet of 10 min is also collected to form the abnormal behavior set data Abnoraml (Algorithm 2).

The clustering process is shown in Figure 6.

During the experiment, we used four attack types. For each of the four types of attacks, random and non-returning are performed according to different proportions. Each type of attack extracts a certain amount accordingly, and the corresponding category attribute is also removed. The attack datasets of D, P, R, and U are, respectively, obtained, and the attack records of DOS, Probe, R2L, and U2R types are, respectively, stored in D, P, R, and U. The proportion of the four attack types corresponding to the extraction and the corresponding number of records are shown in Table 3.

In the experiment, we set the threshold μ to the proportion of abnormal behavior in the entire training set. In order to verify the performance of the UIDM_WODKM

Input: Information table IS = (U, A, V, f), where U = {x1, ..., xn}, A = {a1, ..., am}; the number of clusters k expected.
Output: k initial center points.
Input: Information table IS = (U, A, V, f), where U = {x1, ..., xn}, A = {a1, ..., am}; the number of clusters k expected.
Output: k initial center points.
Initialization: Let $C = \Phi$, where C is the initial set of center points that have been selected
(1) Calculate the division U/IND (A-{a}) and U/IND ({a}) respectively by counting sorting;
(1.1) Calculate the information entropy E (A-{a}) of IND (A-{a});
(1.2) Calculate the importance of the attribute a Sig (a), and thus obtain the weight of a weight (a);
(1.3) For any $x \in$ U, calculate |[x]{a}| according to the division U/IND ({a}), and
(2) Calculate WDens (x) for any $x \in$ U;
(3) Select the object y with the largest weighted average density from U as the first initial center, and $C = C\{y\}$;
(4) If $|C| < k$, go to step (5), otherwise go to step (10);
(5) Assume that C = {c1, c2, ..., cq}, repeated for any $x \in$ U -C
(5.1) Calculate the weighted overlapping distance wd (x, ci) of x and ci, where ci $\in$ C, $1 \le i \le q$;
(5.2) Calculate Pos_ Center (x);
(6) Select the object y that is the most likely to be the initial center from U-C as the new initial center.
And let $C = C\{y\}$;
(7) If $|C| < k$, go to step (5), otherwise go to step (8);
(8) Return k initial centers in C.

Algorithm 1: The initial center selects the Ini_Weight algorithm.

Table 1: UCI datasets.

Datasets	Soybean	Zoo	Breast	Mushroom
Number of categories	4	7	2	2
Number of objects	47	101	699	8124
Number of attributes	35	16	9	22
With or without missing values	No	No	Yes	Yes

Table 2: Initialization results of different algorithms.

Datasets	Clustering	Initializing methods				
		Random	Cao	Wu	Khan	Ini_Weight
Soybean	AC	0.8356	0.8812	0.8812	0.8911	0.9406
	PR	0.8186	0.8702	0.8702	0.7224	0.7676
	RE	0.6123	0.6714	0.6714	0.7716	0.8143
Zoo	AC	0.8356	0.8812	0.8812	0.8911	0.9406
	PR	0.8186	0.8702	0.8702	0.7224	0.7676
	RE	0.6123	0.6714	0.6714	0.7716	0.8143
	AC	0.8461	0.9113	0.9113	0.9127	0.9385
Breast	PR	0.8700	0.9292	0.9292	0.9318	0.9479
	RE	0.7833	0.8773	0.8773	0.8783	0.9167
	AC	0.7318	0.8754	0.8754	0.8815	0.8858
Mushroom	PR	0.7520	0.9019	0.9019	0.8975	0.9080
	RE	0.7278	0.8709	0.8709	0.8780	0.8817

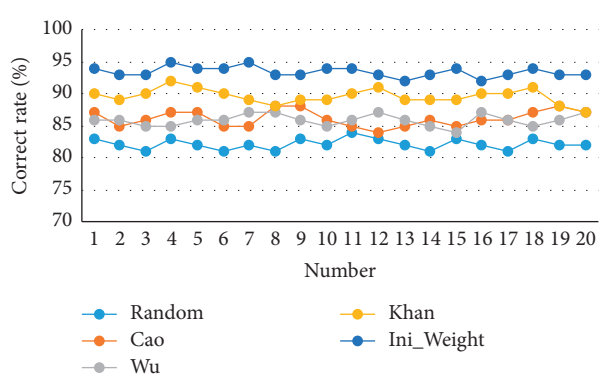


Figure 4: Comparison of the accuracy of different algorithms.

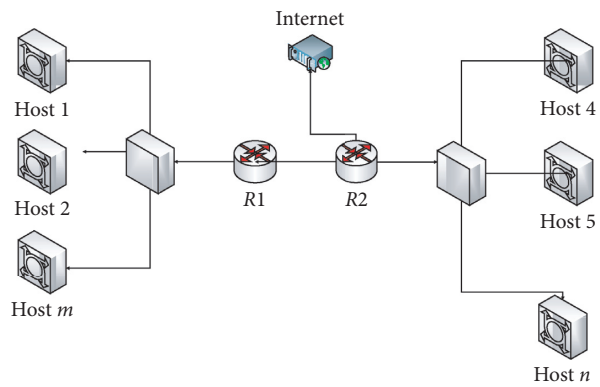


Figure 5: Test network topology.

Input: Information table IS = (U, A, V, f), where $U = \{x1,..., xn\}$, $A = \{a1,..., am\}$; the number of clusters k expected.
Output: A collection of k clusters.
Step 1: according to the previously proposed Ini_Weight algorithm, select k objects from the universe U as the initial center point: $z1$, ..., zk.
Step 2: calculate the weighted overlap distance between each object x in U and each center point zi, and divide the object x into the cluster represented by the nearest center point.
Step 3: for each current cluster c, recalculate the center point of c based on the frequency of the object's value on each attribute in c.
Step 4: repeat steps 2 and 3 until the value of the objective function does not change. After each iteration, let t ++.
Step 5: return the clustering results.

Algorithm 2: The clustering algorithm used in the experiment.

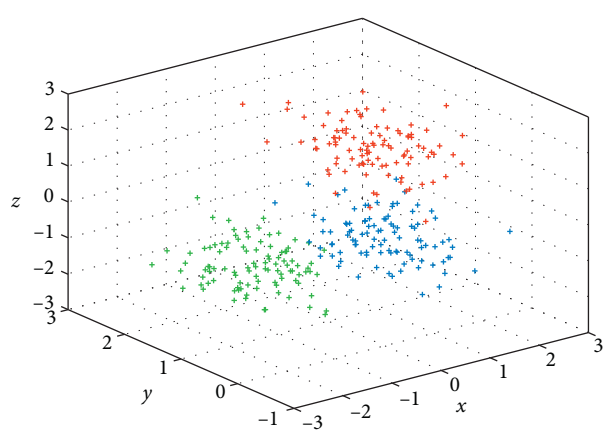


Figure 6: Class process in the experiment.

model in intrusion detection, we compared it with the FCM clustering algorithm and the clustering algorithm proposed by Wu. The FCM algorithm is a fuzzy clustering algorithm based on the objective function, which uses Euclidean distance to measure the similarity of objects. The specific experimental results are shown in Figure 7.

In Figure 5, DR (detection rate) represents the detection rate, and FPR (false positive rate) represents the false alarm rate. By comparing the detection results of the three intrusion detection algorithms listed in the table, the following conclusions can be drawn. The UIDM_WODKM model is better for Probe, U2R, and R2L attacks than the FCM and Wu algorithms. The detection effect of the DOS attack is worse than the FCM algorithm, but it is better than the Wu algorithm. In addition, the overall performance of the UIDM_WODKM model is better than the traditional FCM algorithm and Wu algorithm, especially the UIDM_WODKM model has a much lower false alarm rate than the FCM algorithm. Since Wu's algorithm does not analyze the false alarm rate, we cannot compare it.

Table 3: The percentage and number of four-attack categories.

Attack type	Extraction ratio (%)	Extraction number
DOS	0.311	1217
Probe	24.84	1020
R2L	50	563
U2R	100	52

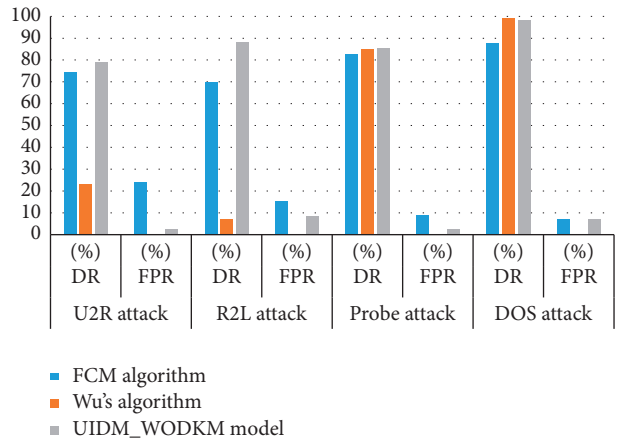


Figure 7: Comparison of results of different algorithms.

5. Conclusion

In this paper, a new distance metric and density metric are proposed based on the concepts of attribute importance and rough entropy in rough set theory, the weighted overlap distance and the weighted average density are used, and an initial class center selection algorithm Ini_Weight based on the weighted density and the weighted overlap distance is proposed [21]. The Ini_Weight algorithm takes full account of the different importance of each attribute so that important attributes are given a larger weight, and unimportant attributes are given a smaller weight. The Ini_Weight algorithm distinguishes the influence of different attributes on the clustering result by the weight of the weight, thus solving the problem of the traditional K-modes algorithm in the initial center point selection. We have carried out experiments on multiple subtype UCI datasets. The experimental results show that the performance of the Ini_Weight algorithm is better than the existing initial center point selection method, which effectively improves the accuracy of the initial center point selection.[24]

Data Availability

All data are available from the corresponding author upon request.

Conflicts of Interest

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

References

[1] H. Sandberg, S. Amin, and K. H. Johansson, "Cyberphysical security in networked control systems: an introduction to the issue," *IEEE Control Systems*, vol. 35, no. 1, pp. 20–23, 2015.

[2] P. Maji, S. K. Pal, and A. Skowron, "Preface: pattern recognition and mining," *Natural Computing*, vol. 15, no. 3, pp. 355–357, 2016.

[3] D. G. Ferrari and L. N. De Castro, "Clustering algorithm selection by meta-learning systems: a new distance-based problem characterization and ranking combination methods," *Information Sciences*, vol. 301, pp. 181–194, 2015.

[4] R. J. Kuo, Y. Potti, and F. E. Zulvia, "Application of meta-heuristic based fuzzy K-modes algorithm to supplier clustering," *Computers & Industrial Engineering*, vol. 120, pp. 298–307, 2018.

[5] R. Zuech, T. M. Khoshgoftaar, and R. Wald, "Intrusion detection and big heterogeneous data: a survey," *Journal of Big Data*, vol. 2, no. 1, p. 3, 2015.

[6] A. S. Eesa, Z. Orman, and A. M. A. Brifcani, "A novel feature-selection approach based on the cuttlefish optimization algorithm for intrusion detection systems," *Expert Systems with Applications*, vol. 42, no. 5, pp. 2670–2679, 2015.

[7] W.-C. Lin, S.-W. Ke, and C.-F. Tsai, "CANN: an intrusion detection system based on combining cluster centers and nearest neighbors," *Knowledge-based Systems*, vol. 78, pp. 13–21, 2015.

[8] R. A. R. Ashfaq, X.-Z. Wang, J. Z. Huang, H. Abbas, and Y.-L. He, "Fuzziness based semi-supervised learning approach for intrusion detection system," *Information Sciences*, vol. 378, pp. 484–497, 2017.

[9] J. Jabez and B. Muthukumar, "Intrusion Detection System (IDS): anomaly detection using outlier detection approach," *Procedia Computer Science*, vol. 48, pp. 338–346, 2015.

[10] M. A. Faisal, Z. Aung, J. R. Williams, and A. Sanchez, "Data-stream-based intrusion detection system for advanced metering infrastructure in smart grid: a feasibility study," *IEEE Systems Journal*, vol. 9, no. 1, pp. 31–44, 2015.

[11] U. Ravale, N. Marathe, and P. Padiya, "Feature selection based hybrid anomaly intrusion detection system using K means and RBF kernel function," *Procedia Computer Science*, vol. 45, pp. 428–435, 2015.

[12] S. A. M. Ramona, C. M. Pompiliu, and S. L. Constantin, "Attainment of K-means algorithm using hellinger distance," *Economic Sciences Series*, vol. 17, no. 2, pp. 324–329, 2017.

[13] F. Atefeh and W. Khreich, "A survey of techniques for event detection in twitter," *Computational Intelligence*, vol. 31, no. 1, pp. 132–164, 2015.

[14] M. I. Jordan and T. M. Mitchell, "Machine learning: trends, perspectives, and prospects," *Science*, vol. 349, no. 6245, pp. 255–260, 2015.

[15] S. Haben, C. Singleton, and P. Grindrod, "Analysis and clustering of residential customers energy behavioral demand using smart meter data," *IEEE Transactions on Smart Grid*, vol. 7, no. 1, pp. 136–144, 2016.

[16] L. Chen, S. Wang, K. Wang, and J. Zhu, "Soft subspace clustering of categorical data with probabilistic distance," *Pattern Recognition*, vol. 51, pp. 322–332, 2016.

[17] D. L. Huerta-Muñoz, R. Z. Ríos-Mercado, and R. Ruiz, "An iterated greedy heuristic for a market segmentation problem with multiple attributes," *European Journal of Operational Research*, vol. 261, no. 1, pp. 75–87, 2017.

[18] G. Gan and M. K.-P. Ng, "k -means clustering with outlier removal," *Pattern Recognition Letters*, vol. 90, pp. 8–14, 2017.

[19] K. A. Prabha and N. K. K. Visalakshi, "Particle swarm optimization based K-prototype clustering algorithm," *IOSR Journal of Computer Engineering*, vol. 17, pp. 56–62, 2015.

[20] S. J. Nanda and G. Panda, "A survey on nature inspired metaheuristic algorithms for partitional clustering," *Swarm and Evolutionary Computation*, vol. 16, pp. 1–18, 2014.

[21] Z. Huang, "Extensions to the k-means algorithm for clustering large data sets with categorical values," *Data Mining and Knowledge Discovery*, vol. 2, no. 3, pp. 283–304, 1998.

[22] F. Cao, J. Liang, and L. Bai, "A new initialization method for categorical data clustering," *Expert Systems with Applications*, vol. 36, no. 7, pp. 10223–10228, 2009.

[23] S. Wu, Q. S. Jiang, and Z. X. Huang, "A new initialization method for clustering categorical data," in *Proceedings of the 11th Pacific-Asia Conference on Advances in Knowledge Discovery and Data Mining*, Nanjing, China, May 2007.

[24] G. Wang, X. a. Ma, and H. Yu, "Monotonic uncertainty measures for attribute reduction in probabilistic rough set model," *International Journal of Approximate Reasoning*, vol. 59, pp. 41–67, 2015.

Hindawi
Scientific Programming
Volume 2023, Article ID 9827619, 1 page
https://doi.org/10.1155/2023/9827619

Retraction

Retracted: Analysis of Tai Chi Ideological and Political Course in University Based on Big Data and Graph Neural Networks

Scientific Programming

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

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

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

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

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

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

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

References

[1] C. Y. Li and L. Zheng, "Analysis of Tai Chi Ideological and Political Course in University Based on Big Data and Graph Neural Networks," *Scientific Programming*, vol. 2021, Article ID 9914908, 9 pages, 2021.

Hindawi
Scientific Programming
Volume 2021, Article ID 9914908, 9 pages
https://doi.org/10.1155/2021/9914908

Research Article

Analysis of Tai Chi Ideological and Political Course in University Based on Big Data and Graph Neural Networks

Chun yan Li[1] **and Lu Zheng**[2]

[1]*College of Physical Education of Inner Mongolia University, Inner Mongolia, Hohhot 010020, China*
[2]*Institute of Foreign Languages of Inner Mongolia Normal University, Inner Mongolia, Hohho 010021, China*

Correspondence should be addressed to Chun yan Li; lichunyantiyu@163.com

Received 2 April 2021; Revised 26 April 2021; Accepted 8 May 2021; Published 25 May 2021

Academic Editor: Shah Nazir

Tai Chi martial arts education is one of the components of school education. Its educational value is not only to require students to master basic Tai movement technical skills and improve their physical fitness but also to bring students' ideological progress and cultivate students to respect teachers and lectures. Excellent moral qualities such as politeness, keeping promises, observing the rules, and acting bravely, as well as the cultivation of the spirit of unity and cooperation, and the quality of will also have a certain meaning. However, the scientific Tai Chi ideological and political courses and the construction of Wude education interactive classrooms lack relevant research. Therefore, this article builds a Tai Chi ideological and political interactive classroom system based on big data technology and graph neural network. First, the spatio-temporal graph convolutional neural network is used to reason about the relationship between Tai Chi action categories and strengthen the low-dimensional features of semantic categories and their co-occurrence expressions used for semantic enhancement of current image features. In addition, in order to ensure the efficiency of the Tai Chi scene analysis network, an efficient dual feature extraction basic module is proposed to construct the backbone network, reducing the number of parameters of the entire network and the computational complexity. Experiments show that this method can obtain approximate results, while reducing the amount of floating-point operations by 42.5% and the amount of parameters by 50.2% compared with the work of the same period, and achieves a better balance of efficiency and performance. Secondly, based on the big data of historical Tai Chi classrooms, this article constructs an interactive classroom system that can effectively improve the quality of Tai Chi ideological and political courses.

1. Introduction

At present, the characteristics of martial arts teaching are clearly written in the teaching materials of Tai Chi [1–3] in colleges and universities. The first one is "the ideological education of advocating martial arts and virtue" [4]. Tai Chi martial arts teachers should pay more attention to the ideological education of students in their daily teaching. The requirements of advocating martial arts and advocating virtue in martial arts are reflected. In actual teaching, all of this seems to have not been paid attention by everyone. The current situation of focusing on technical exercises and neglecting cultural teaching in martial arts teaching has led to the fact that the excellent martial arts culture has not been well continued [5]. How to combine traditional martial ethics with contemporary social development and realize its intrinsic value connection is a problem that needs to be considered at the moment. In the authors' opinion, it is necessary to fully excavate the new elements of moral education in martial arts teaching, to selectively absorb the beneficial parts of traditional martial arts culture, and also to eliminate the parts that are not suitable for social development. On the contrary, it is necessary to base itself on the development reality of culture so that the martial ethics of the new era will have more historical heritage and show the spirit of Tai Chi [6].

From the perspective of "curriculum ideology and politics" [7, 8], this article uses the thrust of "curriculum

ideology and politics" to explore and research the implementation of "curriculum ideology and politics" in college Tai Chi martial arts teaching. It is a good time for inheritance and promotion.

Through the research on the literature and books of Tai Chi Wushu teaching, it can be found that the results of theoretical research are quite rich. However, because the curriculum ideology and politics were put forward in recent years, the teaching concept is relatively new, the research content of martial arts in this field is less, and its related teaching system is not perfect. On the one hand, this research is conducive to broadening the research horizon of ideological and political education, promoting the reform and development of ideological and political education, extending ideological and political education to the martial arts teaching field of physical education, and innovating the exploration and research ideas of ideological and political education in colleges and universities [9, 10]. Political education has been improved in both timeliness and pertinence. On the other hand, it is conducive to perfecting the theoretical system of martial arts education in colleges and universities and makes the martial arts teaching in colleges and universities more cultural connotations. Martial arts is not only a simple technical mastery, but the ideological and political education elements contained in it are extremely rich. The martial arts teaching under the traditional teaching mode is no longer suitable for the development requirements of the current era. Martial arts teaching should keep pace with the times and show its rich cultural connotation and humanistic qualities. With the rapid development of deep learning [11–14], this paper uses big data and graph neural network technology [15–18] to construct an interactive classroom system of Taiji ideology and politics under the perspective of curriculum ideology and politics and realizes the organic combination of teaching Taiji martial arts techniques and guiding the correct values. Following are the main contributions of this paper:

(i) A novel ideological and political interactive classroom teaching model of Taiji Wushu based on big data and graph neural network: first, collect and organize historical big data for Tai Chi ideological and political classes. Secondly, the graph neural network is used to reason about the relationship between the categories of Tai Chi actions and to strengthen the low-dimensional characteristics of semantic categories and their co-occurrence expressions, so as to improve the teaching quality of Tai Chi ideological and political classrooms.

(ii) To use the node information integration and update method of the graph neural network model to perform relationship reasoning and information interaction between features to enhance feature expression.

(iii) An efficient dual feature extraction module to construct the backbone network, and further optimizes the graph model relational reasoning module to reduce the amount of calculation. Finally, a new lightweight encoding enhancement module is proposed in the codec network structure to ensure that the learned features contain both high-level semantic knowledge and sufficient spatial details.

2. Background

Incorporating ideological and political education concepts into the teaching of Tai Chi martial arts, on the one hand, is conducive to expanding the coverage of ideological and political courses, making college ideological and political education face the development of other majors, promoting the integrated development of ideological and political education and college martial arts, and promoting better ideological and political education ground penetration into martial arts teaching; on the other hand, it is beneficial to enrich martial arts teaching content, innovate Tai Chi martial arts teaching methods, and realize the innovative development of skill teaching and ideological and political education in Tai Chi martial arts teaching. At the same time, it also helps students to consciously improve their ideological and political morality and guide them to establish a correct outlook on life and values in a subtle way [19].

2.1. Courses for Ideological and Political Education. "Courses for ideological and political education" [20] refers to the in-depth exploration of the ideological and political theory education resources of various courses. All disciplines and courses set by universities must play the role of ideological and political education and build a full process, all-round, and full curriculum from a strategic height. The pattern of educating people enables various college courses and ideological and political theory courses to go in the same direction, forming a synergistic effect, and a comprehensive education concept that always runs through the fundamental task of "building morality and cultivating people."

2.2. Education of Tai Chi Wushu in Universities. Education of Tai Chi Wushu in Universities is a martial arts teaching activity that limits martial arts education within the scope of schools, takes martial arts classes as the main form, and is led by teachers. It aims to enable students to acquire martial arts technical skills, develop students' intelligence and physical strength, and inherit excellent traditional culture in a purposefully, planned, and organized manner. The education of Tai Chi Wushu [21] in universities referred to in this paper is different from social martial arts and competitive martial arts, which are widely spread in colleges and universities. The ultimate goal of college martial arts is to educate people and cultivate martial arts talents with noble moral qualities and beneficial to social development. The basic content forms of martial arts education in colleges and universities mainly include martial arts classes, martial arts amateur training, martial arts competitions, and extracurricular martial arts activities.

2.3. Graph Convolutional Neural Network. The graph convolutional neural network [22] can effectively extract feature information from non-Euclidean distance data. Suppose the

graph data format used is $G = (V, E)$, where $V = \{v_1, v_2, v_3, \ldots, v_n\}$ represents the set of nodes in the graph and E is the set of edges in the graph. Then, the graph G can be expressed as an adjacency matrix A of size $N * N$. If there is an edge connection between node i and point j, $A_{ij} = 1$. On the contrary, if there are no edges connected, $A_{ij} = 0$. The adjacency matrix is a representative description of the topological information of the graph in the matrix form. The degree of the node is obtained by summing the elements of each row of the matrix. That is, from the node, there are several edges connected to other nodes, denoted as $D_i = \sum_j^N A_{ij}$.

We can define the convolution operation on the graph as the product of a multidimensional signal x, $x \in R^N$, and a classifier $g_{\theta,}$ $\theta \in R^N$, with a parameter θ in the Fourier domain. Y represents the regularized Laplacian moment because the Laplacian matrix has a special eigen-decomposition form: $L = U \wedge U^T$, and $\wedge$ is a diagonal matrix composed of the eigenvalues of the Laplacian matrix. Then, the convolutional form of this input signal can be expressed as

$$g_\theta * x = Y g_{\theta'}(L) Y^T x, \tag{1}$$

where $*$ indicates the convolution process and $g_\theta(L)$ is a function containing Laplacian eigenvalues. Due to the high computational complexity of the above formula, Chebyshev polynomial fitting is used to simplify its complexity so that $g_\theta(L) \approx \sum_{k=0}^{K} \theta_k T_k(L)$, where T_k represents the k-order Chebyshev polynomial and takes $k = 1$ and θ is the Chebyshev coefficient. The final convolutional form of the obtained graph convolutional network is

$$g_\theta * x = \theta\left(I_N + D^{-(1/2)} A D^{-(1/2)}\right) x, \tag{2}$$

where $D = \mathrm{diag}(D_i)$, I_N are the unit matrix with the size of the number of nodes N, and $I_N + D^{-(1/2)} A D^{-(1/2)}$ can be further simplified as $D^{-(1/2)} \widehat{A} D^{-(1/2)}$, where $\widehat{A} = A + I_N$. Through the convolution of the convolution kernel (classifier) and the input signal, the local parameter sharing on the graph structure data can be realized. For the prediction and classification tasks of supervised learning, the result of parameter sharing can be used to judge the similarity between the input data features and the classifier. For unsupervised learning tasks, because the node itself does not have feature information, artificial labels are used in this article to simulate parameters that need to be shared. This artificial tag performs parameter-sharing operations when there are no tags in the surrounding neighbors. In contrast, it is a parameter propagation process.

3. Methodology

3.1. The Basic Connotation of Courses for Ideological and Political Education. Ideological and political curriculum is to incorporate ideological and political elements including ideals and beliefs, spiritual pursuits, and values in each course so that students will be subtly affected in their thoughts and behaviors on the premise of learning professional knowledge and skills. Through careful combing and careful analysis of the country's related documents on "Curriculum Ideology and Politics" and the views of some scholars, the authors found that the basic connotation of "Curriculum Ideology and Politics" is mainly embodied in six aspects: cultivating talents, peers in the same direction, all employees' participation, unity of diversity, combination of explicit and implicit, and scientific innovation. First of all, the essence of curriculum ideology and politics is outstanding. "Educating people as the foundation, moral education first," "Moral education" is an ancient topic that keeps pace with the times. Cultivating talents with both morality and ability, integrating "curriculum ideological and political" is to integrate ideological and political content into all disciplines. In the education and teaching of the course, its ultimate goal is to realize morality and foster people. Curriculum ideology and politics adhere to the moral education; first, pay attention to strengthening the cultivation of students' moral quality, and guide students to establish a correct outlook on life and values.

Walking in the same direction is another highlight of curriculum ideology. Ideological and political curriculum requires the same direction and common development of ideological and political courses and other courses. The organic combination of ideological and political education and theoretical knowledge learning is a requirement for educational work in the new era. Ideological and political curriculum education is a new concept that puts forward the "double skin" phenomenon of ideological and political work and talent training and is ensuring that the ideological and political curriculum can play a role. Under the premise of educating people, keep pace with other existing courses, strive to realize the unity of teaching and educating people of various subjects, infiltrate ideological and political education into all links of education and teaching work of various subjects, and realize mutual promotion and common development.

3.2. Tai Chi Recognition. Tai Chi recognition is a key part of constructing Tai Chi ideological and political interactive classrooms. This section will explain the Tai Chi movement recognition algorithm based on the graph neural network.

3.2.1. Skeleton Diagram. The joint points of the skeleton diagram are very important for describing the posture of Tai Chi athletes and expressing body language. It is an important data modality and has important significance, such as behavior recognition in Tai Chi classroom. Figure 1 shows a schematic diagram of the human skeleton.

In mathematics, a graph is defined as a relational structure composed of vertices and edges connecting the vertices. A vertex can represent any research object, and an edge represents a specific relationship between two objects. A graph can be expressed as a set of vertices and edges, denoted as $G = (V, E)$, where v is a set of vertices and E is a set of edges. At the same time, suppose that the number of vertices in the graph is N and the number of edges is M. An edge connecting vertex v_i is denoted as (v_i, v_j) or e_{ij}, as shown in Figure 2(a):

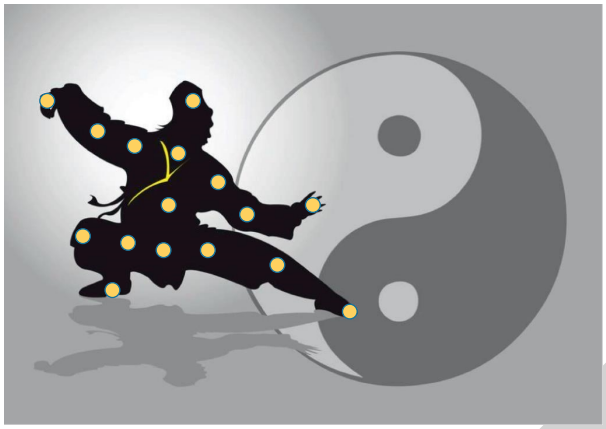


Figure 1: Schematic diagram of Tai Chi athlete's skeleton.

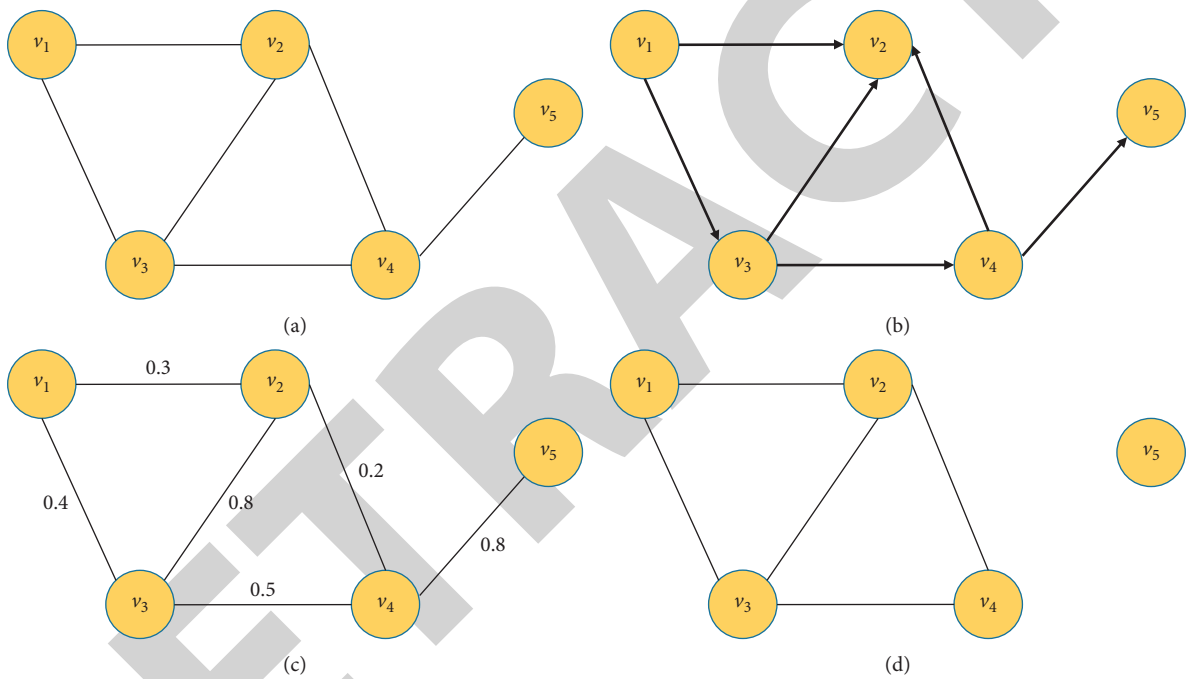


Figure 2: Different kinds of graph structures.

$$V = \{v_1, v_2, v_3, v_4, v_5\},$$
$$E = \{(v_1, v_2), (v_1, v_3)(v_2, v_3)(v_2, v_4)(v_3, v_4)(v_4, v_5)\}. \quad (3)$$

According to whether the edge has directionality, the graph can be divided into directed graph and undirected graph. If the edge in the graph has directionality, then such an edge is called a directed edge $e_{ij} = \langle v_i, v_j \rangle$, where v_i is the starting point of this directed edge and v_j is the end point of this directed edge. A graph containing a directed edge is called a directed graph, as shown in Figure 2(b). Corresponding to a directed graph is an undirected graph, the edges in an undirected graph are all undirected edges. In fact, an undirected graph can be regarded as a special symmetrical directed graph that contains two directions $e_{ij} = \langle v_i, v_j \rangle = \langle v_j v_i \rangle = e_{ji}$. According to whether the edges are assigned corresponding weights, the graph can be divided into weighted graphs and nonweighted graphs. Figure 2(c) shows a weighted graph. In actual application scenarios, the weight can be expressed as the distance between the two places or the transportation cost. Under normal circumstances, it is customary to abstract the weight as the strength of the connection between two vertices. Corresponding to it is a nonweighted graph, which can be regarded as a special weighted graph, that is, all weights are equal. In addition, according to whether the vertices are connected, the graph can be divided into connected graphs and unconnected graphs. If there are isolated points in the graph and there is no edge connected to it, it is an unconnected graph; otherwise, there is no isolation. The graph of vertices is called a connected graph. The vertex v_5 in Figure 2(d) does not have any edges connected to it, so it is a disconnected graph.

Given a graph $G = (V, E)$, where $V = \{v_1, v_2, \ldots, v_N\}$ and $E = \{e_1, e_2, \ldots, e_M\}$, the adjacency matrix is defined as

$$A_{ij} = \begin{cases} 1, & \text{if } (v_i v_j) \subseteq E, \\ 0, & \text{else.} \end{cases} \tag{4}$$

Then, the incidence matrix can be defined as

$$B_{ij} = \begin{cases} 1, & \text{if } v_i \text{ is connected to } e_j, \\ 0, & \text{else.} \end{cases} \tag{5}$$

The degree matrix is a diagonal matrix, and the diagonal elements are the degrees of their respective vertices, which can be expressed as

$$D = [d_{ii}], \quad i = 1, 2, \ldots, N. \tag{6}$$

Then, the Laplacian matrix is defined as

$$L = D - A. \tag{7}$$

The above three kinds of matrices can be expressed as follows:

$$\begin{aligned} A &= \begin{bmatrix} 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 \\ 1 & 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix}, \\ B &= \begin{bmatrix} 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}, \\ A &= \begin{bmatrix} 2 & 0 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 & 0 \\ 0 & 0 & 3 & 0 & 0 \\ 0 & 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}, \\ L = D - A &= \begin{bmatrix} 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 \\ 1 & 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix} - \begin{bmatrix} 2 & 0 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 & 0 \\ 0 & 0 & 3 & 0 & 0 \\ 0 & 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \\ &= \begin{bmatrix} -2 & 1 & 1 & 0 & 0 \\ 1 & -3 & 1 & 1 & 0 \\ 1 & 1 & -3 & 1 & 0 \\ 0 & 1 & 1 & -3 & 1 \\ 0 & 0 & 0 & 1 & -1 \end{bmatrix}. \end{aligned} \tag{8}$$

The skeleton data of Tai Chi athletes is usually represented by the 2D or 3D coordinates of each joint in each frame. From the structural diagram Figure 2(a), it can be seen that this is an undirected connected graph. Furthermore, in order to be able to capture information in the time dimension, the skeleton sequence is represented hierarchically by constructing a time-space skeleton diagram. The construction of spatio-temporal skeleton graph data is divided into two steps: the first step is to connect the joint points in each frame according to the connectivity of the human body's natural structure. Connect the nodes in the second step, in the time dimension, by doing the same thing between consecutive frames. This connection method does not require any artificial definition, so this network architecture can handle data with different numbers of joints or different joint connectivity. The constructed time-space skeleton diagram is shown in Figure 3.

The set of edges consists of two parts: $E = E_{\text{intra}} + E_{\text{inter}}$, where E_{intra} represents the connecting edges that connect the joint points according to the natural connectivity of the human body within each frame, which is defined as a

$$E_{\text{intra}} = \{(v_{ti}, v_{tj}) | v_{ti}, v_{tj} \in V_t; i, j = 1, 2, \ldots, 16\}, \tag{9}$$

where E_{inter} represents the connecting edge between consecutive frames connecting the same joint points, which is defined as

$$E_{\text{inter}} = \{v_{ti}, v_{(t+1)i}\}. \tag{10}$$

Of the two connecting edges, the former can be understood as a representation of the complete structure of the action posture, while the latter can be understood as the motion trajectory of a specific joint point in the time dimension. At the same time, the adjacency matrix can be defined accordingly:

$$A = [a_{ij}], \quad i = 1, 2, \ldots, 16; j = 1, 2, \ldots, 16. \tag{11}$$

3.2.2. Spatio-Temporal Graph Convolutional Neural Network. Spatial graph convolution has made good progress in tasks such as node classification and graph classification. In order to effectively extract spatio-temporal characteristics, it is extended to spatio-temporal graph convolution, which can construct a large amount of real-life data that can be expressed as dynamic graphs. Recalling the convolution operation for 2D images, the input image or feature map can be regarded as a 2D matrix, and the output feature map is still a 2D matrix; at the same time, if the step size is set to 1, plus the appropriate padding operation, the input can be realized. Keep the same size as the output. Given a $k \times k$ convolution kernel and the input feature map is X, the output at x can be expressed as

$$Y(x) = \sum_{h=1}^{k} \sum_{w=1}^{k} X(S(x, h, w)) * W(h, w), \tag{12}$$

where S represents the sampling function, which is used to sample the neighbors at position x and W represents the weight function, which is used to generate different weight vectors to calculate the inner product of the sampling input vector. In standard 2D convolution, the weight function has

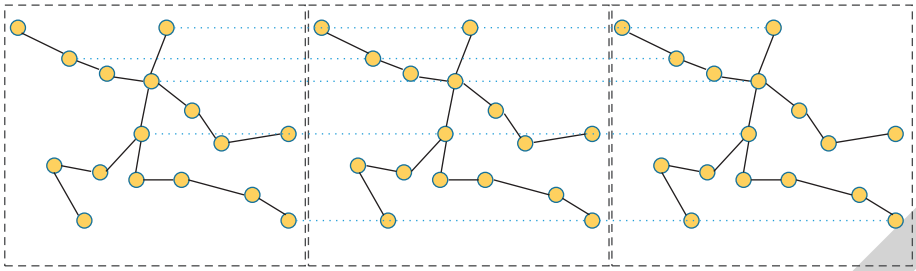


Figure 3: Time-space skeleton graph.

nothing to do with the position. Therefore, in the same picture or feature map, the weight is shared. The spatial graph convolution is mainly for the expansion of the sampling function and weight function. In the 2D convolution of the picture, the sampling function is defined as 8 areas of pixels, which can be extended to the graph data. The sampling function can be defined in the node's neighbor node set $Nb(v_{ti}) = \{v_{tj} | d(v_{ti}, v_{tj}) \leq D\}$, where $d(v_{ti}, v_{tj})$ is the shortest distance from node v_{ti} to node v_{tj}, specifically for

$$S(v_{ti}, v_{tj}) = v_{tj}. \tag{13}$$

The structure of graph data varies, and the definition of the weight function is one of the main challenges of graph convolution directly applied to the spatial domain. To this end, a weight function based on the labeling strategy is adopted. First, the nodes in the neighbor node set are divided into k subsets, and then, a digital label is assigned to each subset. The label function is defined as follows:

$$L_s(Nb(v_{ti})) = j, \quad j = 0, 1, \ldots, k-1. \tag{14}$$

Therefore, the weight function can be defined as follows:

$$W(v_{ti}, v_{tj}) = W'(L_{ti}(v_{tj})). \tag{15}$$

So the spatial graph convolution can be defined as

$$Y = \sum_{v_{tj}} \frac{1}{Z_{ti}(v_{tj})} X(S(v_{ti}, v_{tj})) * W(v_{ti}, v_{tj}). \tag{16}$$

Based on the above equations, the final output can be expressed as

$$Y = \sum_{v_{tj}} \frac{1}{Z_{ti}(v_{tj})} X(v_{tj}) * W(L(v_{tj})). \tag{17}$$

In fact, there is also the concept of neighbor nodes in the time dimension, and the set of neighbor nodes in the time domain can be defined

$$Nb(v_{ti}) = \left\{v_{qj} | d(v_{tj}, v_{qj}) \leq k, |q - t| \leq \left|\frac{\tau}{2}\right|\right\}. \tag{18}$$

The parameter τ is used to control the size of the time-dimensional convolution kernel. In the time dimension, the connection mode of the nodes is relatively fixed, so the label function can be rewritten as follows:

$$L_t(v_{qj}) = L_{ti}(v_{tj}) + \left(q - t + \left|\frac{\tau}{2}\right|\right) \times k. \tag{19}$$

Therefore, the calculation form of the single-layer spatio-temporal graph convolution is as follows:

$$Y = \Lambda^{-(1/2)}(A + I)\Lambda^{-(1/2)} XW. \tag{20}$$

3.3. Tai Chi Ideological and Political Interactive Classroom Framework Based on Big Data and Spatio-Temporal Graph Convolution. The Tai Chi ideological and political interactive classroom framework based on big data and spatio-temporal graph convolution is mainly divided into Tai Chi ideological and political interactive classroom historical big data analysis, construction of Tai Chi athlete skeleton diagram, spatio-temporal map convolution network feature extraction, classification network, and classroom interaction modules. First, extract the skeleton joint points for the input video frame sequence to obtain the position information of all skeleton joint points and the corresponding estimated confidence score; then, in the spatial dimension, the connection is made inside each frame according to the natural connectivity of the human body, and then, construct a skeleton diagram to connect the same joint points between consecutive frames in the time dimension to construct a time-space skeleton diagram; then, use the space-time diagram convolution to extract the features of the constructed space-time diagram data; finally, use the classification network to identify video clips. The framework of Tai Chi ideological and political interactive classroom is shown in Figure 4.

4. Experiments and Results

4.1. Experimental Environment. The graph neural network proposed in this paper is implemented by the deep learning framework Pytorch and trained on a workstation equipped with GTX 1080Ti GPU. The entire network is trained for 40 epochs, and the minimum batch training size is 8. The size of the input RGB image and depth image is adjusted to 256×256. Data enhancement techniques such as random horizontal flip, random rotation, and random brightness change are used in the training process to prevent overfitting. The optimizer uses Adam with a weight attenuation of 0.0001, and the initial learning rate is set to 0.0001.

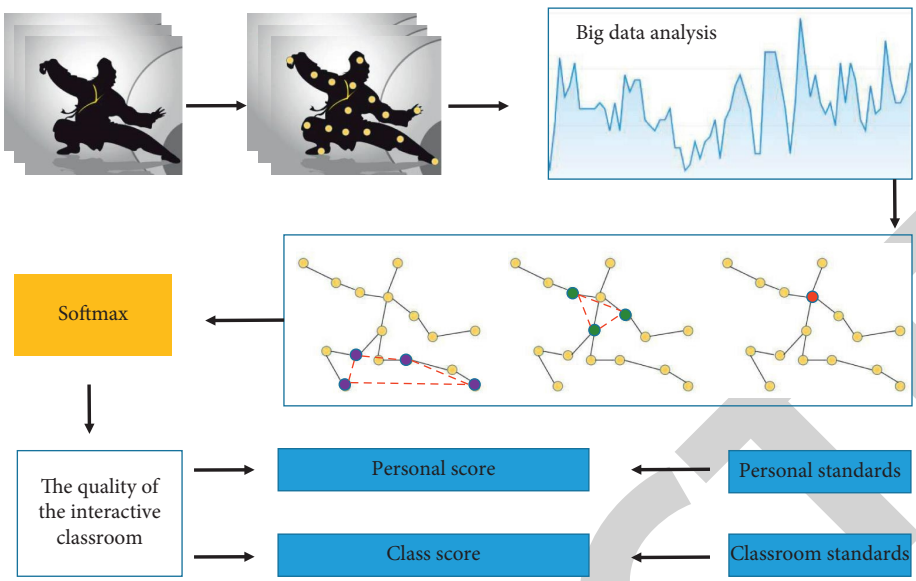


Figure 4: The Tai Chi ideological and political interactive classroom framework based on big data and spatio-temporal graph convolution.

4.2. Dataset. This paper sorted out and analyzed the big data of historical Tai Chi ideological and political classroom and produced a dataset. Due to the low efficiency of manual labeling of samples, the number of samples in the custom dataset is small, which will cause a certain degree of overfitting. Overfitting will lead to poor performance of the model on the test dataset. In order to avoid overfitting, the model is trained by means of transfer learning. The so-called transfer learning refers to the application of knowledge or patterns learned in a certain field to different but related fields; for transfer learning in deep learning, it refers to the use of previously trained models on related datasets and direct calls. Its structure and network parameters are then applied to a model training method in current research tasks.

4.3. Evaluation Index. We mainly use accuracy to evaluate the algorithm, and its calculation equation is as follows:

$$\mathrm{Acc} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FP}}. \tag{21}$$

4.4. Experimental Results. The experimental results are shown in Table 1 and Figure 5. Among the three models, the highest accuracy rate on the validation set is 79.58%, and the highest accuracy rate on the test set is 78.11%. The method of combining 2D convolution with skeleton joint point features is the worst. The main reason is that 2D convolution and skeleton joint point features are static features. Although the two features are combined, both are spatial information, and sign language actions. Time is a dynamic process that not only contains spatial information but also includes temporal motion information characteristics. In contrast, the recognition accuracy of the 3D convolution model has been significantly improved, and the accuracy has increased by nearly 10%. The reason is that the 3D convolutional neural network increases the feature extraction in the time dimension so that the final extracted features include not only the spatial structure information but also the changes in the spatial structure information in the time dimension. The skeleton joint point recognition model also considers the skeleton joint point data, but the recognition performance is not as good as the spatio-temporal graph convolution

Table 1: Experimental result.

	2D conv	3D conv	3D conv + transfer learning	Ours
Validation set	60.52	71.56	76.14	**79.58**
Test set	58.69	65.25	73.26	**78.11**

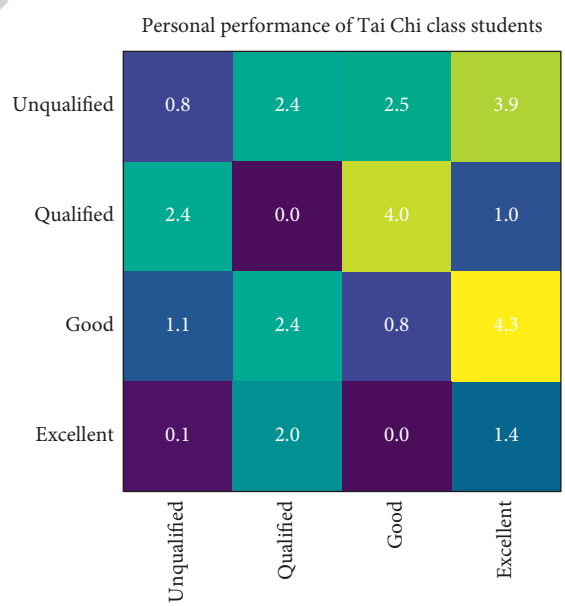


Figure 5: Confusion matrix.

model, which shows the importance of the information in the time dimension for dynamic feature extraction.

5. Conclusion

In this paper, we proposed a Tai Chi ideological and political interactive classroom system based on big data technology and graph neural network. First, the spatio-temporal graph convolutional neural network is used to reason about the relationship between Tai Chi action categories and strengthen the low-dimensional features of semantic categories and their co-occurrence expressions used for semantic enhancement of current image features. In addition, in order to ensure the efficiency of the Tai Chi scene analysis network, an efficient dual feature extraction basic module is proposed to construct the backbone network, reducing the number of parameters of the entire network, and reducing the computational complexity. Experiments show that this method can obtain approximate results, while reducing the amount of floating-point operations by 42.5% and the amount of parameters by 50.2% compared with the work of the same period, and achieves a better balance of efficiency and performance. Secondly, based on the big data of historical Tai Chi classrooms, this article constructs an interactive classroom system that can effectively improve the quality of Tai Chi ideological and political courses.

Data Availability

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

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This research was supported by the 2019 Inner Mongolia University Project "College Sports 2 (no. 1912510012)."

References

[1] L. Zou, H. Wang, Z. Xiao, Q. Fang, M. Zhang, T. ., ... Li et al., "Tai chi for health benefits in patients with multiple sclerosis: a systematic review," *PloS One*, vol. 12, no. 2, Article ID e0170212, 2017.

[2] C.-A. Cheng, Y.-W. Chiu, D. Wu, Y.-C. Kuan, S.-N. Chen, and K.-W. Tam, "Effectiveness of Tai Chi on fibromyalgia patients: a meta-analysis of randomized controlled trials," *Complementary Therapies in Medicine*, vol. 46, pp. 1–8, 2019.

[3] L. Zou, J. Han, C. Li et al., "Effects of Tai Chi on lower limb proprioception in adults aged over 55: a systematic review and meta-analysis," *Archives of Physical Medicine and Rehabilitation*, vol. 100, no. 6, pp. 1102–1113, 2019.

[4] D. U. Li-ying, "Effective promotion of ideological and political education in colleges and universities from the practical perspective," *Journal of University of Electronic Science and Technology of China (Social Sciences Edition)*, vol. 2, 2017.

[5] L. A. Lipsitz, E. A. Macklin, T. G. Travison et al., "A cluster randomized trial of Tai chi vs health education in subsidized housing: the MI-WiSH study," *Journal of the American Geriatrics Society*, vol. 67, no. 9, pp. 1812–1819, 2019.

[6] S. Sungkarat, S. Boripuntakul, S. Kumfu, S. R. Lord, and N. Chattipakorn, "Tai Chi improves cognition and plasma BDNF in older adults with mild cognitive impairment: a randomized controlled trial," *Neurorehabilitation and Neural Repair*, vol. 32, no. 2, pp. 142–149, 2018.

[7] X.-Y. Shao, "Discussion on the teaching reform of accounting major from the perspective of curriculum ideology and politics," in *Proceedings of the DEStech Transactions on Social Science, Education and Human Science, (ICESD 2020)*, Barcelona, Spain, February 2020.

[8] Y. Du, Q. Sun, H. Lv, and X. Hu, "Exploration of trinity "course ideology and politics" teaching in electrical specialty," in *Proceedings of the 2019 3rd International Conference on Education, Management Science and Economics (ICEMSE 2019)*, pp. 514–517, Atlantis Press, Singapore, August 2019.

[9] W. Chen, S. Yu, and D. Xiong, "Effects of Tai Chi intervention on perceived stress, anxiety, and sleep in college students," *Advances in Physical Education*, vol. 10, no. 1, pp. 54–67, 2019.

[10] A. K. Converse, B. P. Barrett, B. A. Chewning, and P. M. Wayne, "Tai Chi training for attention deficit hyperactivity disorder: a feasibility trial in college students," *Complementary Therapies in Medicine*, vol. 53, p. 102538, 2020.

[11] X. Ning, P. Duan, W. Li, and S. Zhang, "Real-time 3D face alignment using an encoder-decoder network with an efficient deconvolution layer," *IEEE Signal Processing Letters*, vol. 27, pp. 1944–1948, 2020.

[12] Z. L. Yang, S. Y. Zhang, Y. T. Hu, Z. W. Hu, and Y. F. Huang, "VAE-Stega: linguistic steganography based on variational auto-encoder," *IEEE Transactions on Information Forensics and Security*, vol. 16, pp. 880–895, 2020.

[13] Y. Tong, L. Yu, S. Li, J. Liu, H. Qin, and W. Li, "Polynomial fitting algorithm based on neural network," *ASP Transactions on Pattern Recognition and Intelligent Systems*, vol. 1, no. 1, pp. 32–39, 2021.

[14] W. Cai, B. Liu, Z. Wei, M. Li, and J. Kan, "TARDB-Net: triple-attention guided residual dense and BiLSTM networks for hyperspectral image classification," *Multimedia Tools and Applications*, vol. 80, no. 7, pp. 11291–11312, 2021.

[15] Z. Huang, Q. Li, H. Fang, T. Zhang, and N. Sang, "Iterative weighted nuclear norm for X-ray cardiovascular angiogram image denoising," *Signal, Image and Video Processing*, vol. 11, no. 8, pp. 1445–1452, 2017.

[16] X. Zhang, Y. Yang, Z. Li, X. Ning, Y. Qin, and W. Cai, "An improved encoder-decoder network based on strip pool method applied to segmentation of farmland vacancy field," *Entropy*, vol. 23, no. 4, p. 435, 2021.

[17] W. Cai and Z. Wei, "Remote sensing image classification based on a cross-attention mechanism and graph convolution," *IEEE Geoscience and Remote Sensing Letters*, 2020.

Hindawi
Scientific Programming
Volume 2021, Article ID 9918433, 10 pages
https://doi.org/10.1155/2021/9918433

Research Article

College English Flipped Classroom Teaching Model Based on Big Data and Deep Neural Networks

Heli Chang

Department of Foreign Languages, Liaocheng University Dongchang College, Liaocheng 252000, Shandong, China

Correspondence should be addressed to Heli Chang; changheliliaocheng@163.com

Received 2 April 2021; Revised 23 April 2021; Accepted 5 May 2021; Published 25 May 2021

Academic Editor: Shah Nazir

With the rapid development of information technology, flipped classroom as a new type of mixed teaching mode relying on computer technology has changed the traditional teaching mode and formed a teaching process of "learning first and teaching later," and it has been used in many fields of teaching. Flipped classroom reverses the sequence of traditional teaching knowledge transfer and knowledge internalization and improves students' autonomy. However, it is still in the exploratory stage of the specific impact of the flipped classroom teaching model on college students' English autonomous learning ability. Therefore, this article proposes a novel college English flipped classroom teaching model based on big data and deep neural networks. The study has selected a total of 230 students in two classes of the second-year English major of a university as the research objects. Data are utilized to investigate the changes of the two groups of students' English autonomous learning ability and English academic performance, to explore the specific changes of college students' English autonomous learning ability and its influencing factors through interviews, and to predict and effectively analyze the weight of influencing factors through the deep neural network. This research enriches the theoretical research results of college students' English autonomous learning ability under the flipped classroom teaching model, provides reference for the cultivation of college students' English autonomous learning ability, and has certain reference significance for the optimization of the flipped classroom teaching model. The proposed research will support researchers and practitioners at college and university level.

1. Introduction

The traditional college English teaching model [1–3] deprives learners of English reading ability and communication ability, and it is difficult to tap learners' learning potential. In addition, teachers spend too much time in the classroom, and learners lack opportunities for language communication and exercise. In the traditional teacher management model, the teacher is the protagonist of the college English teaching work model [4]. The learners have not developed good self-study habits; in the classroom, the learners are still passive and mechanically taking notes. As a result, many learners' interest in English learning is not high, and the effect of English teaching activities is not ideal. In order to change these problems in English teaching, it is imperative to discover more effective English teaching models. The new curriculum standard advocates the improvement of students' autonomous learning ability and innovation ability. It does not require teachers to teach knowledge content comprehensively, but to allow learners to learn independently. This is an important teaching concept in flipped classroom education practice. This article believes that using flipped classroom teaching mode to design high school English reading classes can be a good solution, making teachers become knowledge leaders [5]. With the development of flipped classroom technology, researchers have applied the flipped classroom teaching model to various subjects in teaching, but few researchers have applied it to English reading, especially in high school English teaching, so I believe that flipped classroom should be applied. Research in college English teaching is very valuable [6].

Flipped classroom is a new teaching mode [7–10]; that is, teachers can use modern electronic data and information

management technology to change the traditional English class into a need to distribute the learning materials to learners in advance, and the learners can design according to the teacher's education and teaching activities. The questions asked are studied independently before and after class. Teachers and students discuss with each other to solve difficult problems. It overturns the traditional forms of teaching and learning. To a large extent, "students listen and teachers speak" has been changed. Although there are already some problems in some high school education and teaching activities that require teachers to solve them through research and use of the flipped classroom teaching work model, there are still very few practices in English learning. In addition, the current flipped classroom is still based on platforms such as WeChat, which cannot guarantee learners' autonomous learning ability. The existing college English flipped classrooms cannot effectively solve these problems for us. Therefore, it is necessary to conduct some empirical research methods and use experiments to verify whether flipped classrooms have the ability to promote the development of English teaching [11].

Searching with the keywords of "flipped classroom" and "English teaching" indicated that there is a serious lack of research on college English teaching, and the design of the research content of English flipped classroom is also extremely incomplete, ignoring some system designs in the curriculum. In order to speed up the curriculum reform of college English education technology and improve the learning methods, skills, and class efficiency of English reading, it is essential to apply flipped classrooms to college English teaching practice activities [12].

This paper attempts to apply big data technology and deep neural network [13–16] technology to the college English flipped classroom teaching model and use this model to predict and analyze the factors affecting the quality of English flipped classrooms, which can further improve the teaching quality of flipped classrooms and become a flipped classroom. The paper also provides some valuable suggestions, including college English teaching design and research methods. The following are the main contributions points of this paper:

(i) This paper proposes a novel flipped classroom teaching model for college English based on big data and deep neural networks. First, it collects and sorts big data for flipped classrooms. Secondly, it uses deep neural networks to build a flipped classroom recommendation model to achieve personalized learning resource recommendation, which can improve the quality of college English flipped classroom teaching.

(ii) This paper combines the convolutional neural network and the gated recurrent unit to propose a college English flipped classroom prediction model. The model uses the convolutional neural network to automatically extract local features from the flipped classroom big data and is based on the learner's learning behavior. Related to timing, the model introduces a gated loop unit GRU, which enables the model to have good timing information extraction capabilities.

(iii) This paper also introduces an attention mechanism, which combines learner information and course information into the predictive model to improve the generalization ability of the predictive model.

The organization of the paper is given as follows. Section 2 briefly represents the related work to the proposed study. Section 3 of the paper shows the methodology section of the paper with in-depth details. Experiments and results of the paper are given in Section 4. The paper is concluded in Section 5.

2. Related Work

The term "flipped classroom" was first used by American economist Rach Pratt in 2000, and its meaning refers to the integration of our traditional classroom from inside to outside of the classroom. Flipped classroom is to reverse the learning process in the traditional sense, complete the teaching of knowledge before class, and complete the internalization of knowledge in class. American scholars believe that flipped classroom refers to the flipping of the classroom of "teaching in class and homework after class." The idea is to flip the traditional learning process, complete the transfer of knowledge before class, and complete the internal knowledge in the classroom. In the process of transformation, flipped classroom is also called "reversal classroom." It is relative to the classroom teaching method of traditional Chinese culture. The traditional classroom teaching is the teaching of relevant knowledge by the teacher in the classroom, and the learners are in the class. Later, the internalization of knowledge can be achieved through homework and practice [17].

The definition of "flipped classroom" in this article is as follows: learners use instructional videos or electronic textbooks produced by teachers for autonomous learning before class and conduct interactive activities between teachers and students or peers, as well as answering questions and cooperating in the classroom. As a form of education and teaching, flipped classroom is different from the traditional teaching model of "teacher speaks, students listen." Instead, it allows learners to learn by watching the microvideo prepared by the teacher before class and summarize the content in the video. The content of the video mainly includes analysis of reading, grammatical points, background knowledge of reading materials, and related consolidation exercises. After completing the preclass microvideo study, learners can have a general understanding of the reading materials, better understand the subject and content of the reading materials, and clarify their own problems and deficiencies in the reading process. In the

course of the classroom, the learners use the inquiry learning method to discuss each other, help each other to complete each other's problems, and exchange opinions so that the learners can draw lessons from the ideas of others. In addition, the learners are in the information. The communication platform can also provide better communication and discussion with teachers, which is extremely beneficial to learners. For some important and difficult points, learners can also review and consolidate knowledge through video review [18] and achieve the overall improvement of reading comprehension ability.

In flipped classroom education and teaching, the production of micro-videos is also extremely important. Learners need to summarize knowledge and key points by watching the videos themselves. After watching micro-videos, learners can have a general understanding of English reading materials. The background knowledge mentioned in the video can also help high school learners to expand their professional knowledge and consolidate their practice so that they can better understand the main ideas and content of the reading materials and ask themselves clearly the possible questions in the reading process [19].

The traditional teaching mode of reading is teacher-led and learners passively accept the knowledge. Teachers keep lecturing and learners keep taking notes. As a result, learners are limited in learning knowledge due to teaching activities, so the reform of English teaching mode is imperative. Nowadays, flipped classroom has attracted wide attention. This teaching mode can highly improve the classroom education and teaching efficiency. In classroom activities, teachers and learners form a positive atmosphere of teaching and learning. Among learners, inquiry-based learning methods are used to discuss the difficult problems with each other and help each other to complete the exchange of views, so as to enable learners to learn from others' ideas and thus improve the quality of college English learning [20].

3. Methodology

In Figure 1, in the flipped classroom teaching model, teachers need to make microvideos before class to promote good learning for learners and allow learners to carry out targeted training in education and teaching. Learners can explore solutions through group cooperative learning. For these issues, the teacher will finally give feedback and comments.

The following sections briefly show the methodology of the paper.

3.1. English Course Resource Recommendation

3.1.1. Problem Description. How to realize the personalized learning resources recommended is essential to analyze the relationship between learners and learning resources to flip the classroom teaching content. Providing personalized course is the key to big data acquisition course and to effective analysis, data mining, and other processing means, eventually recommending appropriate English curriculum resources to learners. Personalized learning resources recommended model can make use of the history of the learners' learning resources for collecting information. The information study can be used as shown in Figure 2 in the history of the matrix. R is for learning resources, L represents the learner, the shaded part represents the learning resources learned by the learner, and the white blank part represents the learning resources not learned. To solve the problem is how to implement personalized learning resources through the historical information matrix recommendations; namely, the new learning resources in the recommended resources can meet the demand of learners.

3.1.2. Recommended Model. According to the personalized learning resources, a description of the problem is recommended. This paper constructs the recommendation model. This model can be summed up in two processes, namely, training process and recommendations, as shown in Figure 3. The training process includes learning platform data processing, such as algorithm design process, based on deep learning algorithm, optimizing the depth of the neural network [21–24], and more efficient and reasonable training process. The recommended process recommended models which are obtained by training process, obtaining the personalized learning resources. The recommendation model can be divided into three levels, in which the first and second levels belong to the training process, and the third level is the personalized learning resource recommendation process. In the first stage, in order to make the deep neural network work, it is necessary to obtain its input and output. For the input, this paper designs a feature selection model based on MIFS to describe and process the historical learning data. For output, the learner-resource bipartite graph association model is constructed, and the conditions of whether learning resources are recommended or not are defined. In the second stage, the depth of the neural network learning model is optimized through the first level of input and output; in order to address the problem of customized recommendation, a variety of neural network strategies are modeled and optimized, and a recommended model is available through the training phase of complex learning tools. The third level is where qualified learning resources can test data as input to the suggested model and can be recommended to adhere to the learners' high accuracy of customized learning resources.

3.1.3. Feature Selection Module. The attributes of learners and learning resources have diversified characteristics. There are many factors that affect learners' choice of learning resources, which may include characteristics such as gender, majors, learning goals, content preferences, learning styles, cognitive levels, and learning motivations; on the other hand, learning resources may have inherent attributes such as resource styles and interaction methods. Therefore, it is necessary to find the relationship between learners and resources among many features and establish a feature selection model to complete the input process of the recommendation method.

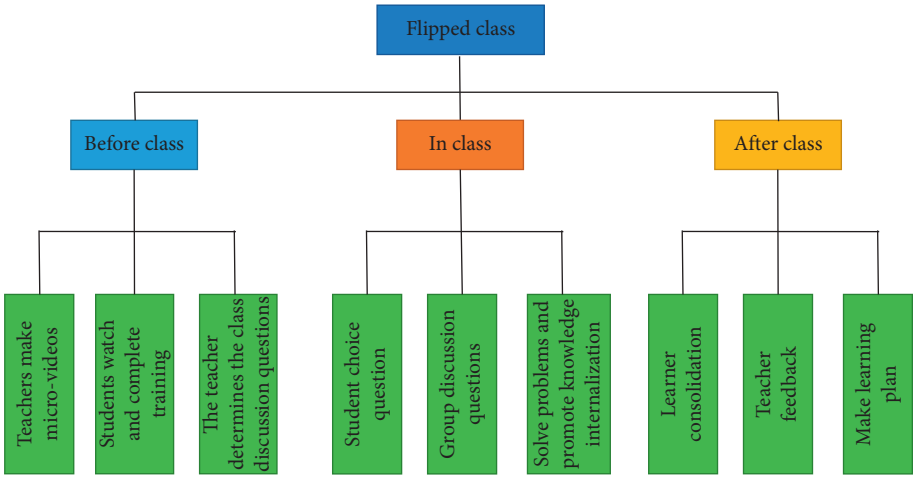


Figure 1: Teaching mode of college English flipped class.

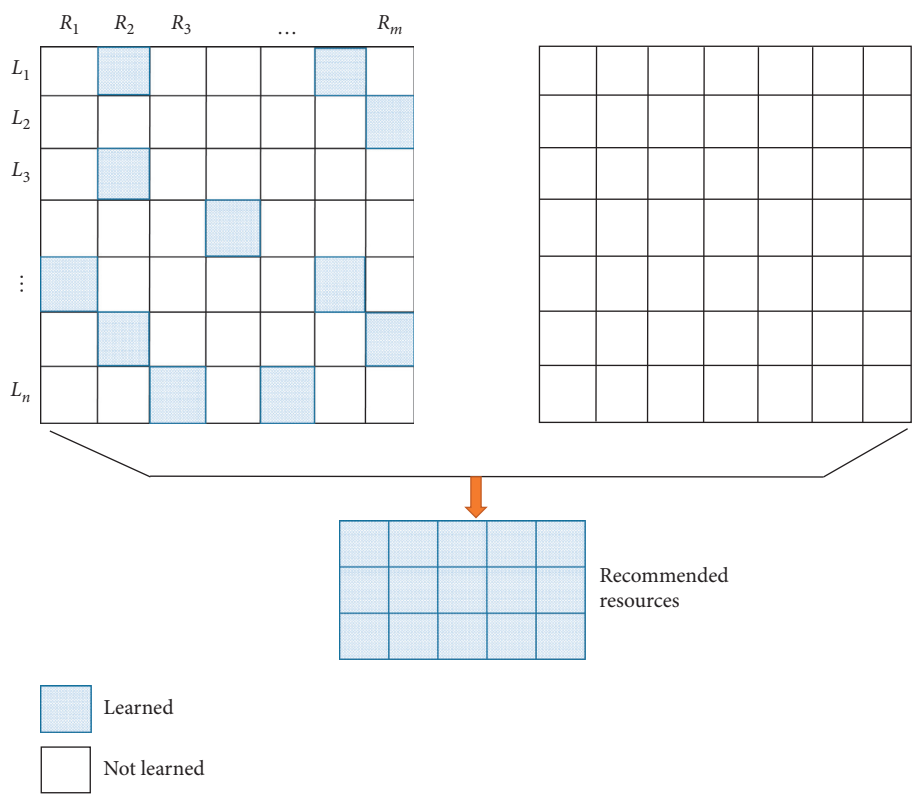


Figure 2: Schematic diagram of the recommendation problem of college English personalized learning resources.

The filtering feature selection method can usually directly use the performance of the training data to evaluate the features, and it has nothing to do with the subsequent algorithm; the speed is faster, and the relationship between the actual data is selected, so the method based on mutual information feature selection (MIFS) is selected. The information describes the strength of the association between two random attributes or features and judges the correlation between a single feature and the target category, thereby reducing the redundancy of feature dimensions. Figure 4 shows the MIFS-based feature selection model.

In the MIFS-based feature selection method, the information measurement evaluation function is very important. Although the function has various forms, the purpose is to select the feature subset that has the greatest correlation with the category. The generalized information measurement evaluation function can be expressed as

$$J(f)\alpha \times g(C, f, S) - \delta, \tag{1}$$

where S is the selected feature, f is the candidate feature, C is the category, function $g(C, f, S)$ is the amount of

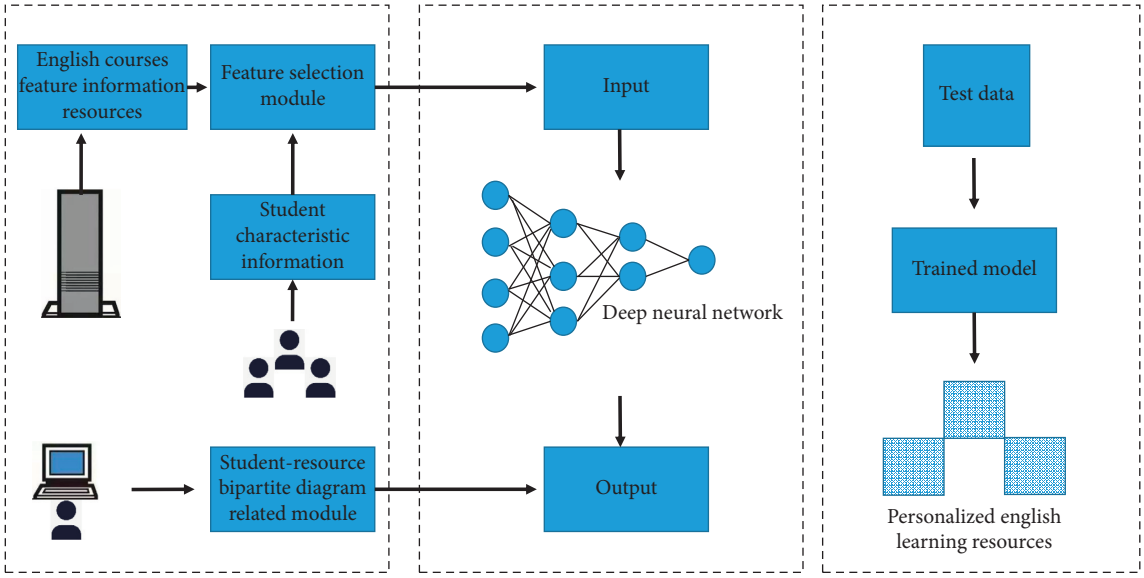


Figure 3: English resource recommendation model.

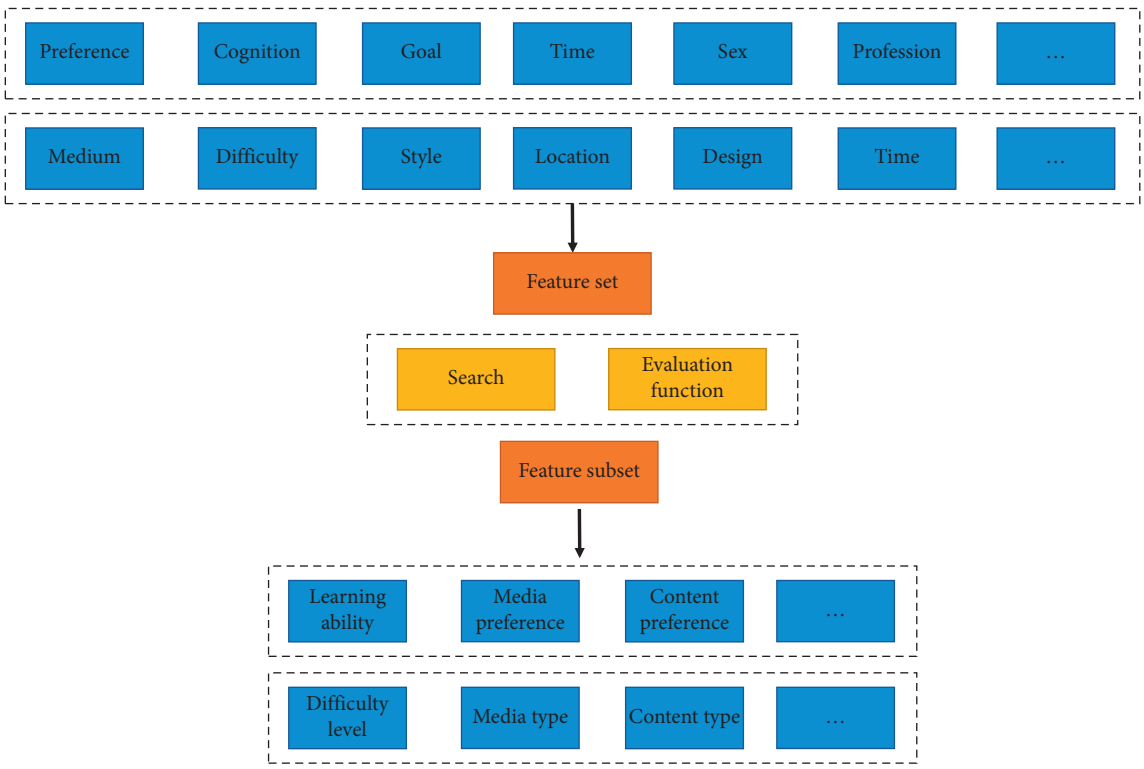


Figure 4: Feature selection model.

information between C, f, and S, that is, between the selected feature and the category after adding the candidate feature, α is the adjustment coefficient, used to adjust the degree of information brought by the addition of f, and a is the penalty factor, used for the degree of redundancy that f brings to S. Therefore, the simplest and most intuitive information measurement evaluation function can be

$$J(f) = g(C, f), \tag{2}$$

where g is mutual information and β is the regulation coefficient. Then, all the candidate features are directly calculated as the evaluation function to prevent excessive redundancy in the preprocessing. Considering the correlation between features, f is punished by the correlation

between feature f and the selected single feature s, and the final evaluation function can be expressed as

$$J(f) = g(C, f) - \beta \sum_{s \in S} g(s, f). \tag{3}$$

It is particularly important to determine the learners' preference for the features of learning resources. The selected features indicate some features that will affect learners' choice of resources, such as the knowledge content and learning duration of resources. These features can be determined according to the sorting of known studies or the investigation methods. On behalf of the temporary, feature candidates are unsure if learners can learn the characteristics of some specific tools, such as learners' age, professional history, gender, and category of the selected feature extraction, which will be used to calculate the selected features and candidate information correlation. The construction of final evaluation function is to add larger correlation feature candidates, determine the candidate characteristics' influence on the evaluation results, and filter out some redundant features. The purpose is to effectively alleviate the late use of data for deep learning training workload.

3.1.4. Student-Resource Bipartite Graph Correlation Model. This paper presents a bipartite graph association model of learner-resource and defines the learner set as $L = \{l_1, l_2, \ldots, l_m\}$. The resource set is $R = \{r_1, r_2, \ldots, r_n\}$. So can get a set by the learners and resource collections of binary relation matrix $X_{m\times n}$, row vector on behalf of the learners, column represents the learners to learn about the resources, if $X_{m\times n} = 1$, this indicates the learners to learn about the resources, if $X_{m\times n} = 0$ is not learned. In fact, this simple unitary processing cannot objectively reflect the degree of learners' awareness of the learning resources, so learning the frequency of a particular resource cannot be ignored.

$$X_{m\times n} = \begin{bmatrix} X_{11} & X_{21} & \cdots & X_{m1} \\ X_{12} & X_{22} & \cdots & X_{m2} \\ \vdots & \vdots & \cdots & \vdots \\ X_{1n} & X_{2n} & \cdots & X_{mn} \end{bmatrix}. \tag{4}$$

Learners' learning frequency of learning resources can reflect different degrees of preference. The average frequency of resource learning can be defined as

$$\overline{r}(l_i) = \frac{\sum_{j=1}^{j} k(l_i, r_j)}{n(l_i)}, \tag{5}$$

where $\overline{r}(l_i)$ represents the average number of times l_i of using learning resources, $n(l_i)$ represents the number of learning resources l_i used, and $\sum_{j=1}^{j} k(l_i, r_j)$ represents the total number of learning resources. The average number of times of using learning resources $R(L_i)$ is taken as the critical value, which is the classification analysis of the recommendation model.

$$\begin{cases} k \geq \overline{r}(l_i) & \text{recommend} \\ k \leq \overline{r}(l_i) & \text{not recommended} \end{cases}. \tag{6}$$

3.2. Deep Neural Network

3.2.1. Attention Mechanism. The input information of the model can be represented by set $X = [X^1, X^2, \ldots, X^T]$, and each element of the set X represents the input information at a certain moment. The index position of the selected information can be represented by the attention variable $z \in [1, T]$. When $z = t$, it means that the t-th information is selected in the input information set X. Given a task-related query vector q and input information X, first calculate the attention distribution (probability distribution) a_t, which represents the t input information:

$$a_t = \frac{\exp\left(X_t^T W q\right)}{\sum_{1\leq t\leq T} \exp\left(X_t^T W q\right)}. \tag{7}$$

The attention distribution can be interpreted as the degree to which the t-th information receives attention when the context queries q. The input information is encoded by a "soft" information selection mechanism, which is called the soft attention mechanism (as shown in Figure 5):

$$\text{attention} = \sum_{t=1}^{N} a_t X_t. \tag{8}$$

The main role of the attention mechanism is to effectively reduce the dimension size of the input data of the model based on the corresponding network structure without changing the original position of the input data. In this way, the design defect of the traditional codec is avoided, and the input data is not limited to a fixed length, which reduces the information loss of the input data.

3.3. GRU. As a kind of LSTM improved recursive neural network, GRU gated circulation unit has a simpler gate structure, which reduces a lot of matrix multiplication computation in the operation of neural network. At the same time, GRU networks can make better use of historical data and learn about the long-term dependence of information. The activation function of the neural network is responsible for mapping the input of the neuron to the output, which increases the nonlinear ability of the neural network model. The choice of activation function directly affects the performance of GRU network and other neural network structures. At present, activation functions commonly used in deep learning include Sigmoid function and tank function belonging to the Sigmoid system and ReLU function of the Modifying Linear Unit (ReLU) system.

But the ReLU function also has drawbacks. Firstly, in the process of neural network training, part of the input will enter the hard saturated region of ReLU function so that some neurons will never be activated by any data again, resulting in neuronal necrosis. Secondly, the output of the ReLU function has no negative value, which will cause the

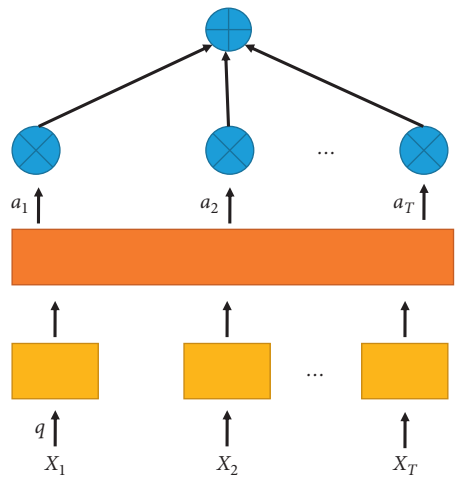


Figure 5: Schematic diagram of soft attention mechanism.

mean shift phenomenon of the output which is not the zero mean value. Mean shift is easy to cause the neurons in the lower layer to get the nonzero mean signal output by the upper layer as input, which makes it difficult to calculate the weight matrix W of the network. Therefore, this paper uses ELU to optimize the performance of the neural network. The ELU function expression and the derivative expression are as follows:

$$f(x)\begin{cases} a(e^x - 1) & x \le 0 \\ x & x > 0 \end{cases}, f(x)\begin{matrix} f(x) + a & x \le 0 \\ 1 & x > 0 \end{matrix}. \quad (9)$$

4. Experiments and Results

4.1. Experimental Environment. Since the experiment in this article needs to train a deep neural network, the scale is large, the structure is more complex, and the calculation scale is large. The programming language used is Python, the version is 3.6.5, the deep learning framework used is Pytorch 0.4, the IDE for program deployment is Pycharm, and all experiments are conducted in the same environment. All our experiments have been conducted on a desktop PC with an Intel Core i7-8700 processor and an NVIDIA GeForce GTX 1080 GPU.

4.2. Experimental Setup. To verify that the personalized learning resources recommended by the method in this paper meet the needs of learners, a series of experiments were conducted. Experimental data includes not only learning resource data, but also the historical data of learners' learning. In the existing public data sets, such as EDX, World UC, and other data sets, it provides dozens of attributes, including course data, learner information, and learner behavior data. This paper also collects the flipped classroom big data from two classes of an English major in a university. In this experiment, the initial learning rate is set to 0.001, and after 15,000 and 20,000 iterations, the learning rate becomes one-tenth of the original, and 30 training data sets are fed each time.

4.3. Evaluation Index. Accuracy rate (P), recall rate (R), and $F1$-score value (F) were used for evaluation in this paper. The calculation equation is as follows:

$$P = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FP}},$$
$$R = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FN}}, \quad (10)$$
$$F = \frac{2 \times P \times R}{P + R}.$$

4.4. Experimental Results. Table 1 is a sample used in this paper. It can be seen from the experimental results with the change of the sample size data sets that the algorithm precision (P) and the recall rate (R) will follow but are at a higher value. This paper that is constructed for personalized learning resources recommended problem model, and using machine learning algorithm is feasible and effective, with the classification of evaluation indexes. The P and R often cannot be in good value, such that sample 1 R value of the SVM algorithm is higher, but the P value is smaller, not set at this time from an indicator to judge its performance, therefore to $F1$-score values (F) as recommended comprehensive evaluation index of performance. As you can see, the depth of the neural network designed (DE) on the F value is superior to the other traditional machine learning algorithms and has high value. In order to verify the method adapted to different scale data sets, combined with Table 2, it can be seen as the cumulative learning more and long time, in different times, due to the learners or other factors affect the situation. All kinds of indexes will change, but the trend is still good, as shown in Figures 6–8. For different samples along with the increase in quantity or learning time, three of the classification indexes will increase, studying the records. It is conducive to the correlation analysis between learners and learning resources and provides recommendation performance. From the algorithm comparison level, it can be seen that the algorithm based on deep neural network model is

Table 1: Data sample descriptions.

	Sample 1	Sample 2	Sample 3
Number of learners	50	80	100
Time (days)	15	15	30
Learn the number of records	7528	12015	30021

Table 2: Experimental results.

Methods	Sample 1			Sample 2			Sample 3		
	P	*R*	*F*	*P*	*R*	*F*	*P*	*R*	*F*
DT	0.38	0.53	0.44	0.45	0.53	0.4	0.59	0.60	0.59
SVM	0.45	0.53	0.48	0.54	0.55	0.54	0.66	0.68	0.57
KNN	0.57	0.57	0.57	0.51	0.56	0.70	0.74	0.76	0.75
Ours	**0.62**	**0.61**	**0.71**	**0.60**	**0.65**	**0.75**	**0.78**	**0.78**	**0.76**

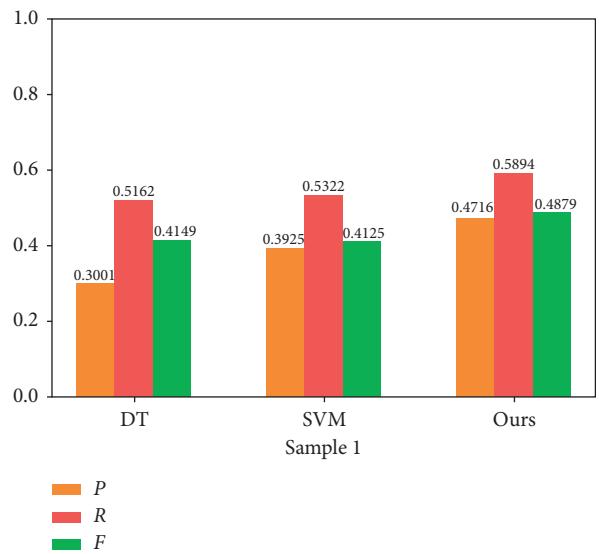


Figure 6: Histogram of evaluation index results for sample 1.

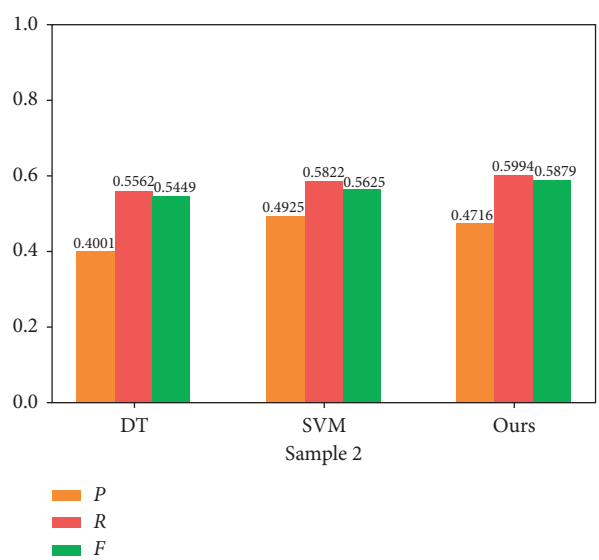


Figure 7: Histogram of evaluation index results for sample 2.

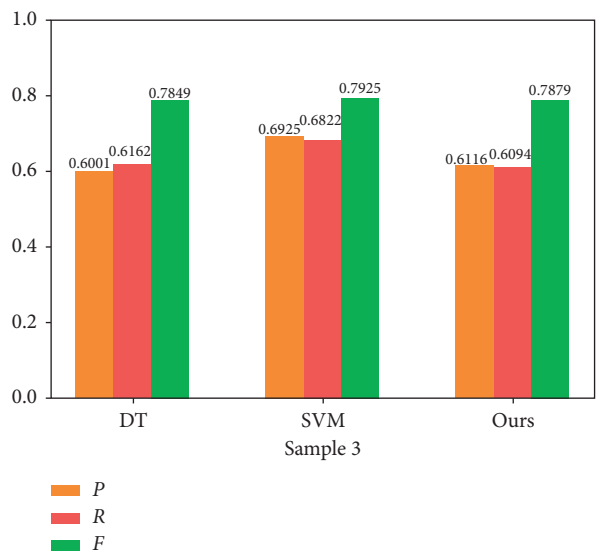


Figure 8: Histogram of evaluation index results for sample 3.

better than other algorithms, indicating that the method proposed in this paper is consistent with and targeted to the practical problem of learning resource recommendation.

It can be clearly seen that the error between the predicted result and the actual result is the smallest, indicating that the improved deep neural network has good adaptability in predicting the frequency of learners' learning resources and shows better performance than other algorithms. This is because, with the increase in learning behavior records, learners' learning rules for resources become more and more obvious. The error judgment of regression analysis is helpful in predicting learners' interests and can improve the teaching quality of college English flipped classrooms.

5. Conclusion

This paper proposes a novel college English flipped classroom teaching model based on big data and deep neural networks. The study selects a total of 230 students in two classes of the second-year English major of a university as the research objects. Data are utilized to investigate the changes of the two groups of students' English autonomous learning ability and English academic performance, to explore the specific changes of college students' English autonomous learning ability and its influencing factors through interviews, and to predict and effectively analyze the weight of influencing factors through the deep neural network. This research enriches the theoretical research results of college students' English autonomous learning ability under the flipped classroom teaching model, provides reference for the cultivation of college students' English autonomous learning ability, and has certain reference significance for the optimization of the flipped classroom teaching model.

Data Availability

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

Conflicts of Interest

The author declares no conflicts of interest.

References

[1] F. Zhang, "Quality-improving strategies of college english teaching based on microlesson and flipped classroom," *English Language Teaching*, vol. 10, no. 5, pp. 243–249, 2017.

[2] D. Qiu, "A study on english teaching model under the background of big data," *Journal of Physics: Conference Series*, vol. 1744, no. 3, Article ID 032061, 2021.

[3] D. Zhang and X. Wang, "The effects of the call model on college english reading teaching," *International Journal of Emerging Technologies in Learning*, vol. 12, no. 12, pp. 24–34, 2017.

[4] H. Zhang, S.-J. Jin, and S.-Z. Du, "Developing a curriculum model of English teaching for master's degree nursing education in a chinese medicine university," *International Journal of Nursing Sciences*, vol. 7, no. 1, pp. 99–104, 2020.

[5] M. Wiesenberg and R. Tench, "Deep strategic mediatization: organizational leaders' knowledge and usage of social bots in an era of disinformation," *International Journal of Information Management*, vol. 51, p. 10204, 2020.

[6] N. Guan, J. Song, and D. Li, "On the advantages of computer multimedia-aided english teaching," *Procedia Computer Science*, vol. 131, pp. 727–732, 2018.

[7] G. Akçayır and M. Akçayır, "The flipped classroom: a review of its advantages and challenges," *Computers & Education*, vol. 126, pp. 334–345, 2018.

[8] K. F. Hew and C. K. Lo, "Flipped classroom improves student learning in health professions education: a meta-analysis," *BMC Medical Education*, vol. 18, no. 1, pp. 1–12, 2018.

[9] P. Strelan, A. Osborn, and E. Palmer, "The flipped classroom: a meta-analysis of effects on student performance across disciplines and education levels," *Educational Research Review*, vol. 30, Article ID 100314, 2020.

[10] Z. Turan and B. Akdag-Cimen, "Flipped classroom in english language teaching: a systematic review," *Computer Assisted Language Learning*, vol. 33, no. 5-6, pp. 590–606, 2020.

[11] G.-Z. Liu, J.-Y. Chen, and G.-J. Hwang, "Mobile-based collaborative learning in the fitness center: a case study on the development of english listening comprehension with a context-aware application," *British Journal of Educational Technology*, vol. 49, no. 2, pp. 305–320, 2018.

[12] L. Shi, J. Delahunty, and X. Gao, "Constraints preventing Chinese EFL teachers from putting their stated beliefs into teaching practice," *Professional Development in Education*, vol. 45, no. 5, pp. 774–789, 2019.

[13] X. Ning, Y. Wang, W. Tian, L. Liu, and W. Cai, "A biomimetic covering learning method based on principle of homology continuity," *ASP Transactions on Pattern Recognition and Intelligent Systems*, vol. 1, no. 1, pp. 9–16, 2021.

[14] L. Zhang, X. Wang, X. Dong, L. Sun, W. Cai, and X. Ning, "Finger vein image enhancement based on guided tri-Gaussian filters," *ASP Transactions on Pattern Recognition and Intelligent Systems*, vol. 1, no. 1, pp. 17–23, 2021.

[15] S. Li, X. Ning, L. Yu et al., "Multi-angle head pose classification when wearing the mask for face recognition under the COVID-19 coronavirus epidemic," in *Proceedings of the 2020 International Conference on High Performance Big Data and Intelligent Systems (HPBD&IS)*, pp. 1–5, IEEE, Shenzhen, China, May 2020.

[16] W. Cai, B. Liu, Z. Wei, M. Li, and J. Kan, "TARDB-Net: triple-attention guided residual dense and BiLSTM networks for hyperspectral image classification," *Multimedia Tools and Applications*, vol. 80, no. 7, pp. 11291–11312, 2021.

[17] L. Buckingham and R. S. Alpaslan, "Promoting speaking proficiency and willingness to communicate in Turkish young learners of english through asynchronous computer-mediated practice," *System*, vol. 65, pp. 25–37, 2017.

[18] S. Franceschini, P. Trevisan, L. Ronconi et al., "Action video games improve reading abilities and visual-to-auditory attentional shifting in english-speaking children with dyslexia," *Scientific Reports*, vol. 7, no. 1, pp. 1–12, 2017.

[19] N. Madikiza, M. P. Cekiso, B. P. Tshotsho, and N. Landa, "Analysing english first additional language teachers' understanding and implementation of reading strategies," *Reading & Writing*, vol. 9, no. 1, pp. 1–10, 2018.

[20] D. Zou, "Research on college English teaching model based on multimedia and network," in *Proceedings of the 3rd International Conference on Social Science and Technology Education*, Wuhan, China, April 2017.

[21] X. Ning, W. Li, and J. Xu, "The principle of homology continuity and geometrical covering learning for pattern recognition," *International Journal of Pattern Recognition and Artificial Intelligence*, vol. 32, no. 12, Article ID 1850042, 2018.

[22] X. Zhang, Y. Yang, Z. Li, X. Ning, Y. Qin, and W. Cai, "An improved encoder-decoder network based on strip pool method applied to segmentation of farmland vacancy field," *Entropy*, vol. 23, no. 4, p. 435, 2021.

[23] Z. L. Yang, S. Y. Zhang, Y. T. Hu, Z. W. Hu, and Y. F. Huang, "VAE-Stega: linguistic steganography based on variational auto-encoder," *IEEE Transactions on Information Forensics and Security*, vol. 16, pp. 880–895, 2020.

[24] X. Ning, X. Wang, S. Xu et al., "A review of research on co-training," in *Concurrency and Computation: Practice and Experience*John Wiley & Sons, Hoboken, NJ, USA, 2021.

Hindawi
Scientific Programming
Volume 2021, Article ID 2846621, 6 pages
https://doi.org/10.1155/2021/2846621

Research Article

Research on the Effect of English Talents Gathering Based on Big Data Hotspot Collection Technology

Chunyan Wei

Xuchang University, Xuchang 461000, China

Correspondence should be addressed to Chunyan Wei; 12012047@xcu.edu.cn

Received 14 April 2021; Revised 6 May 2021; Accepted 10 May 2021; Published 24 May 2021

Academic Editor: Shah Nazir

Talent is the best group in human resources, and the Talents are the best group in human resources, and English-speaking talents are the most dynamic factor in productivity. In order to improve the quantitative analysis ability of the English talent aggregation effect, the English talents aggregation effect analysis model is proposed based on large-scale data collection technology. The collection information flow model of the hotspot big data of English talents aggregation effect is constructed. The high-dimensional feature grouping method is used to reconstruct the hotspot big data of the English talents aggregation effect. The piecewise linear test method is used to analyze the statistical characteristics of the hotspot big data of the English talents aggregation effect and extract the frequent vector set which reflects the hot big data category attribute of the English talents aggregation effect. According to the result of feature extraction, the fuzzy English talents aggregation is processed to realize the fusion of big data information of English talents aggregation effect hotspot. Combined with quantitative analysis method, the automatic classification of big data association rules is realized. The experimental and simulation results show that the proposed method is more accurate and effective than the traditional methods in collecting hotspot data, which is 26% and 76% higher than the traditional methods. This method has better accuracy and improved data aggregation effect in collecting hotspot data of English talents gathering and has strong collecting ability and characteristics. The research improved hotspot big data's English talents gathering effect.

1. Introduction

A large number of scientific and technological talents are introduced and trained under the guidance of the development and cultivation of strategic emerging industries and the emphasis on the growth needs of enterprises [1]. By introducing the team as a whole, the core talents led by the introduction, the development of high and new technology projects, the introduction of postdoctoral workstations, etc., it is necessary to introduce talents from overseas study abroad who are in urgent need of the development of industries, as well as famous domestic experts, and cultivate a group of academic leaders, forming a group of outstanding scientific and technological talent groups and teams. Talent is the best group in human resources. English talent is the most dynamic factor in productive forces, and the flow of English talents among regions can form a phenomenon of English talents gathering in a certain region [2]. It can improve the level of scientific research and education of English talents and reduce the transaction cost. Its main characteristics are spatial, clustering, and scale. The phenomenon of English talents aggregation can produce two kinds of effects: one is the economic effect of English talents gathering, and the other is the diseconomic effect of English talents aggregation. The economic effect of English talents aggregation refers to the aggregate effect of English talents gathering in a certain region in accordance with certain internal relations and under the action of harmonious internal and external environment [3].

The main reasons for the formation of diseconomic effect are that the accumulation of English talents does not form the internal relationship between English talents, the environment of English talents gathering is not ideal, the mobility is not strong enough, and all kinds of conflicts occur in the organization. In order to study conveniently, this paper classifies the diseconomic effect of English talents

aggregation as the phenomenon of English talents aggregation, and it classifies the economic effect as English talents gathering effect. The overall characteristics of English talents aggregation can be divided into organizational effect, environmental effect and self-effect of English talents aggregation [4]. The specific characteristics can be divided into eight characteristics: information sharing effect, knowledge spillover effect, innovation effect, collective learning effect, incentive effect, time effect, regional effect, and scale effect. The following are the contributions of the proposed study.

(i) In order to improve the ability of quantitative analysis of English talents aggregation effect, a model of English talents aggregation effect analysis based on big data hotspot collection technology is proposed.

(ii) The collection information flow model of the hotspot big data of English talents aggregation effect is constructed, and the grouping reconstruction of the hotspot big data of English talents aggregation effect is carried out by using the high-dimensional characteristic grouping method.

(iii) By using the piecewise linear test method, the statistical characteristics of the hotspot big data of English talents aggregation effect are analyzed, and the frequent vector sets reflecting the hot big data category attributes of the English talents aggregation effect are extracted.

(iv) According to the result of feature extraction, the fuzzy English talents aggregation is processed to realize the fusion of big data information of English talents aggregation effect hotspot. Combined with the method of quantitative regression analysis, the automatic classification of hotspot big data association rules of English talents aggregation effect is realized.

(v) Finally, the performance test is carried out by the simulation experiment, and the validity conclusion is obtained.

2. Analysis on the Aggregation Effect of English Talents and the Countermeasures of Talent Attract

English talents receive long years of education and have a higher level. Whether they can give full play to their talents often depends on the cultural background of science and education in the region where the English talents are located. Good regional science and education culture can make English talents have a stage in this environment, show their "fantastic ideas," stimulate their creative inspiration, and produce the knowledge spillover effect and innovation effect of talent gathering. In the English talents aggregation effect, the talent aggregation mode is mainly reflected in the following aspects:

(1) Strengthen the flow of talents and open the channel of tacit knowledge sharing. On the one hand, it is necessary to break the existing administrative authority of personnel unit ownership and departmental ownership, to change "unit person" into "social person," so that talent can flow freely throughout the country and promote the formation of talent gathering phenomenon. After the gathering of talents, the opportunities of knowledge sharing and information exchange of English talents increase. Through information exchange, more and more knowledge and experience can be obtained, thus forming a more reasonable knowledge structure. Reasonable knowledge structure is the basis of forming innovation ability, without solid knowledge accumulation, and it is difficult to form rich association and creative thinking. On the other hand, to establish innovative learning organizations, English talents can be influenced and cultivated in a specific learning environment by participating in activities, reflection, conversation, communication, cooperation, problem solving, and so on. For example, through academic debates, symposia, expert consultations, and so on, these forms can promote the exchange and sharing of tacit knowledge to a certain extent, so that participants can share time and space. All expressions such as expressions, words, actions, and eyes, all sensory organs are receiving information. At the same time, communication is stimulated and guided by timely stimulation and feedback. Many tacit knowledge which is difficult to communicate can be met in this kind of communication, and tacit knowledge is exchanged and shared in these activities. Compared with explicit knowledge, tacit knowledge has a logical primacy, is the guide of mastering clear knowledge, is the internal motive force of human's cognitive behavior and ultimately determines the creative performance of human beings.

(2) Optimizing the creative environment of talents and transforming government functions. To create a good social atmosphere of introducing, emphasizing, knowing, loving, cultivating, and using talent, to create favorable conditions for promoting the strategy of strengthening the country by talents, and to optimize the environment of innovation and entrepreneurship of talents to ensure that people use their talents to the best of their ability. First, it is necessary to build a career platform for talent innovation and entrepreneurship. Persist in using industry to gather talents, attract talents with projects, retain talents with career, make talents have the opportunity to start a business, have stage for officers, and have room for development. Second, we should pay close attention to and protect talented people, actively guide all kinds of talents to establish a correct world outlook, outlook on life, values, consciously strengthen moral cultivation, further carry forward the spirit of dedication and hard work, unity and cooperation, honesty and trustworthiness,

constantly improve themselves in practice, improve themselves in competition, and enrich yourself in your struggle. Third, relax the field of English talents venture capital. All kinds of English talents are allowed to invest in entrepreneurship in all fields not restricted by the state in various forms, English talents in enterprises and institutions are encouraged to stay without pay, leading the establishment of economic entities and striving to improve the investment and financing environment for English talents to innovate and start businesses. Fourth, it is necessary to change the functions of the government, to build a service-oriented government, to provide services for the gathering of talents, and to enable talented people to work happily in a harmonious environment, thus stimulate their enthusiasm for knowledge innovation.

(3) Improve the ability of both spillovers and recipients through training. The knowledge spillover subject and the recipient subject directly determine the quality of knowledge dissemination and acceptance, and the quality of knowledge dissemination and acceptance directly determines whether knowledge can produce all kinds of innovative performance. Organizations can organize various kinds of training, make English talents enrich knowledge, perfect knowledge structure, improve the quality of knowledge dissemination and acceptance, and promote the production of innovation effect [5–7].

3. Hotspot Big Data Collection and Feature Parameter Extraction of English Talents Gathering Effect

The hot big data information flow of English talents aggregation effect is constructed, and the high-dimensional characteristic grouping method is used to reconstruct the hotspot big data of English talents aggregation effect. x_{n-i} represents the fuzzy distribution of the hotspot big data attribute set of the English talents aggregation effect and η_{n-j} is expressed as the association rule distribution set of the big data attribute characteristic vector of the English talents aggregation effect hotspot, and then the English talents gather [8]. The information flow model of the set effect hotspot big data attribute set is expressed as follows:

$$x_n = a_0 + \sum_{i=1}^{M_{AR}} a_i x_{n-i} + \sum_{j=0}^{M_{MA}} b_j \eta_{n-j}, \tag{1}$$

where a_0 is the sampling amplitude of the big data time series of the initial English talents gathering effect and b_j is the statistical characteristic of hotspot big data of English talents aggregation effect which is analyzed by using piecewise linear test method. $x(t)$ is the optimal association rule distribution attribute of hotspot big data of English talents aggregation effect, $t = 0, 1, \ldots, n-1$. For the scalar time series of hotspot big data of English talents aggregation effect, the finite set of big data set distribution of English talents aggregation effect hotspot is obtained by using the method of piecewise test:

$$\frac{1}{2\pi m} \sum_{k=-q/2}^{q/2} b_k \sum_{i=0}^{p} a_i (n + c_k m)^i = \frac{1}{2\pi} \sum_{i=0}^{p} a_i n^{i-1}. \tag{2}$$

According to the result of vector reconstruction, the feature extraction of association rules of big data, a hotspot of English talents aggregation effect, is carried out [9].

4. Big Data Classification of English Talents Aggregation Effect

4.1. Feature Extraction of Association Rules. On the basis of constructing the data information flow model of English talents aggregation effect, using the high dimension feature grouping method to reconstruct the hot data of the English talents aggregation effect, this paper carries out the automatic classification and optimization design of the hot data of the English talents aggregation effect. This paper puts forward the social people based on the association rules [10]. According to the autocorrelation matching constraints, the detection statistics of the hot data classification of the English talents aggregation effect are expressed as follows:

$$\begin{aligned} &\min F(x) = \left(f_1(x), f_2(x), \ldots, f_m(x)\right)^T, \\ &s.t. \quad g_i \le 0, \quad i = 1, 2, \ldots, q, \\ &\qquad h_j = 0, \quad j = 1, 2, \ldots, p, \end{aligned} \tag{3}$$

where T is the time variable of the hotspot of English talents gathering effect. In the solution space of association rule distribution, the optimal English talents gathering center search is carried out, and the judgment function of the hot big data category judgment of English talents aggregation effect is obtained as follows:

$$Thre = \max_b \left| \int r(t) \frac{1}{\sqrt{a'}} f^*\left(\frac{t-b}{a'}\right) dt \right| = \max_b \left| W_f r(a', b) \right| \underset{<H_0}{\overset{H_1>}{}} \lambda_2, \tag{4}$$

where λ_2 is the second decision threshold, and the initial feature vector $x(t)$ of the hotspot big data output of English talents aggregation effect includes the association rule vector, and the feature extraction method of association rule is adopted [11]. The results of big data feature extraction of English talents aggregation effect are expressed as follows:

$$C(l) = \sum_{j=1}^{k} \sum_{k=1}^{n_j} \left(\left\| x_k^j - A_j(L) \right\| \right)^2. \tag{5}$$

Under the influence of disturbance error, the English talents gathering center is satisfied with the disturbance error $\|C(l) - C(l-1)\| < \xi$, and the iterative formula of English talents gathering center in finite vector set is calculated in the disturbed range. The English talents

aggregation is the hotspot of English talents gathering effect big data. The iterative formula of English talents gathering is

$$A_j(L+1) = 1/n_j \sum_{i=1}^{k} X_i^j \quad j = 1, 2, \ldots, k. \tag{6}$$

According to the above feature extraction results, combined with the fuzzy English talents aggregation algorithm, the big data automatic classification design of the hotspot of English talents aggregation effect is carried out [12].

4.2. Characteristics of English Talents Gathering Effect Automatic English Talents Gathering. On the basis of extracting the association rule vector set which reflects the hot big data category attribute of English talents aggregation effect, the fuzzy English talents aggregation processing is carried out according to the feature extraction result [13]. By using the grouping attribute correlation information of the hotspot big data of English talents aggregation effect, the measurement information of the hotspot big data of English talents aggregation effect is judged, and the classification measures are obtained as follows:

$$E = \sum_{p=1}^{n} E_p = \frac{1}{l} \sum_{p=1}^{n} \sum_{k=1}^{l} \left[r_p(k) - y_p(k) \right]^2. \tag{7}$$

In the rule vector set, the English talents gathering center with the hot big data characteristics of English talents aggregation effect is expressed as

$$u = \sum_{j=0}^{n} a_j \tanh^j \xi. \tag{8}$$

Euler distance is used to represent the collection center of the two groups of English talents aggregation vectors in the hotspot big data classification of English talents aggregation effect [14], which are expressed as follows:

$$\Delta E = -\eta \left[\left(\frac{\partial E}{\partial \omega} \right)^2 + \left(\frac{\partial E}{\partial b} \right)^2 \right] \tag{9}$$

and

$$f(x) = \begin{cases} f(x), x \in Levf \\ a, x \in Levf \end{cases} \tag{10}$$

It can be seen that the convergence function of the error function of the hotspot big data classification of English talents gathering effect is

$$\left| f(x_0 + h) - P_n(h) \right| \le A|h|^a. \tag{11}$$

By using the fuzzy C-means method of English talents aggregation, the hotspot big data information fusion of English talents aggregation effect is carried out, the hotspot big data classification of English talents aggregation effect is realized, and the big data hotspot ability of English talents gathering effect is improved [15].

5. Simulation Experiment

In order to test the application performance of this algorithm in big data collection and clustering analysis, simulation experiments are carried out. The data processing algorithm is designed with MATLAB. The scale of big data sampling is 1024, the scale of training sample set is 200, the time period of hotspot data collection is 0.25 ms, the simulation time is 100 s, and the English talents aggregation effect is hot. The initial sampling frequency of the point big data $f_s = 40$ KHz, in different sampling periods, the test set of different hotspots of English talents gathering effect big data is obtained, and the result of big data hotspot collection is shown in Figure 1.

Taking the big data hotspot of English talents gathering collected in Figure 1 as the research object, the feature analysis of English talents gathering is carried out, and the result of feature extraction is shown in Figure 2.

Figure 2 shows that this method can accurately collect the big data hotspots of English talents, and the clustering of feature extraction is good. The accuracy of big data hotspot collection with different methods is tested, and the results are shown in Figure 3. Figure 3 shows that the three methods are compared with the accuracy of big data's hotspot collection and English talents collection, and the aiming time is 10 : 00. The recall rate of the method in this paper is 933. The recall rate of the proposed method in reference [5] is 67. The recall rate of the method in this paper is 26 times higher than that in reference [5]. The recall rate of the proposed method in reference [7] is 76. The method in this paper is more effective than that in [7]. France's recall rate increased by 17%. When the aiming time is 50, the recall rate of our method is 99%, and the recall rate of the method proposed in reference [5] is 822. The method in this paper is more effective than that in reference [5]. The recall rate is increased by 17%, the recall rate of the method proposed in [7] is 89%, the recall rate of this method is 10% higher than that of the method in [7]; when the aim time is 100, the recall rate of the method in this paper is 100, the recall rate of the proposed method in reference [5] is 933. The recall rate of the method in this paper is 7 percent higher than that of the method in reference [5], the recall rate of the method in reference [7] is 100, and the method in this paper and the method in [7] have been proposed. France has the same recall rate. Experimental data show that the proposed method is more accurate and effective in collecting hotspot data.

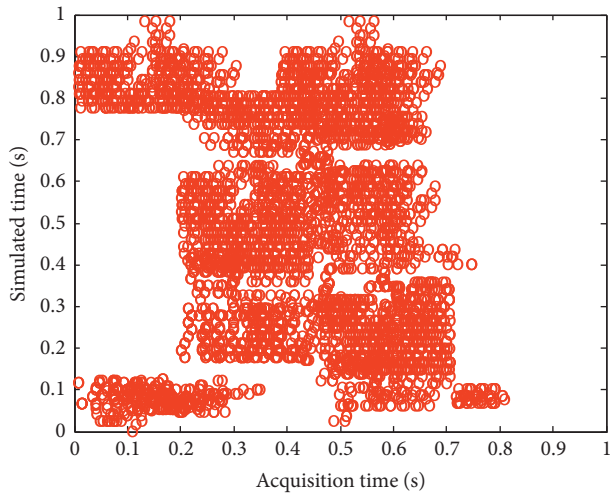


Figure 1: Big data hotspot collection results of English talents gathering.

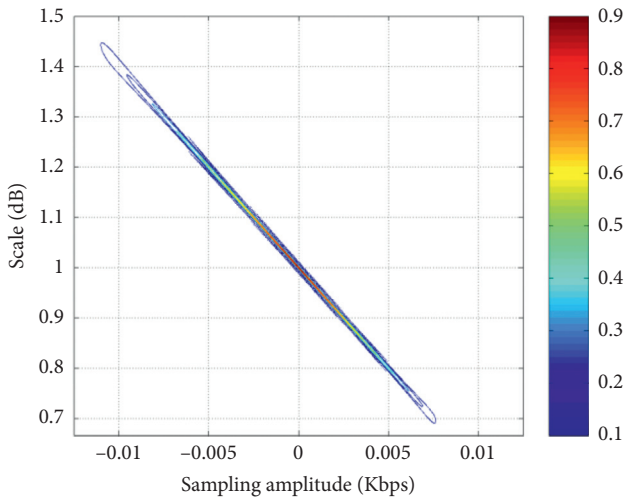


Figure 2: Feature extraction results.

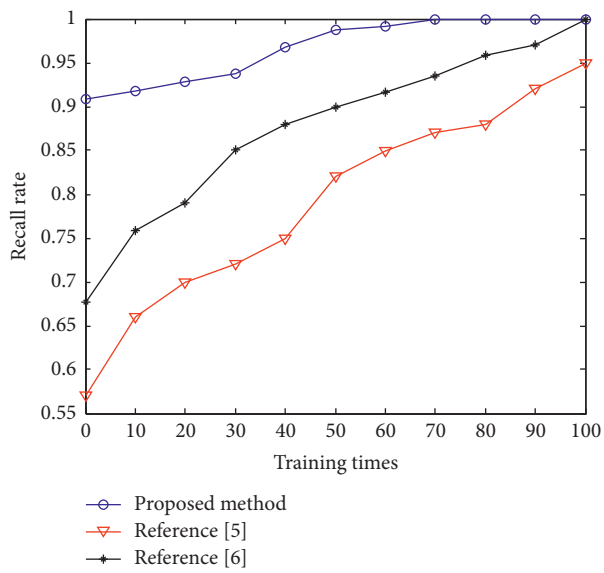


Figure 3: Accuracy comparison of big data hotspot collection with English talents gathering.

6. Conclusions

In this paper, an English talents aggregation effect analysis model is proposed based on large-scale data collection technology. The piecewise linear test method is used to analyze the statistical characteristics of the hotspot big data of English talents aggregation effect and extract the frequent vector set which reflects the hot big data category attribute of English talents aggregation effect. According to the experimental result of feature extraction, the fuzzy English talents aggregation is processed to realize the fusion of big data information of English talents aggregation effect hotspot. Combined with the method of quantitative regression analysis, the automatic classification of hotspot big data association rules of English talents aggregation effect is realized. The experimental and simulation results of the study show that this approach has better accuracy and improved data aggregation effect in the collection of hotspot data of English talents gathering, and the ability of collecting and feature extraction of hotspot big data of English talents aggregation effect is improved. It has good application value in analysis of English talents gathering.

Data Availability

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

Conflicts of Interest

The author declares that there are no conflicts of interest.

Acknowledgments

The study was supported by 2020 First-Class Undergraduate Course Construction Project of Xuchang University (no. PX-1620989) and 2017 Higher Education Reform Research and Practice Project in Henan Province (no. 2017SJGLX441).

References

[1] C. H. Ju and J. B. Zou, "An incremental classification algorithm for data stream based on information entropy diversity measure," *Telecommunications Science*, vol. 31, no. 2, pp. 86–96, 2015.

[2] Y. X. Lyu, C. Y. Wang, C. Wang et al., "Online classification algorithm for uncertain data stream in big data," *Journal of Northeastern University (Natural Science Edition)*, vol. 37, no. 9, pp. 1245–1249, 2016.

[3] Y. Chen and L. J. Li, "Very fast decision tree classification algorithm based on red-black tree for data stream with continuous attributes," *Journal of Nanjing University of Posts and Telecommunications (Natural Science Edition)*, vol. 37, no. 2, pp. 86–90, 2017.

[4] Y. Wu, B. Shen, and H. Ling, "Visual tracking via online nonnegative matrix factorization," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 24, no. 3, pp. 374–383, 2014.

[5] M. Ye, Y. Qian, and J. Zhou, "Multitask sparse nonnegative matrix factorization for joint spectral-spatial hyperspectral imagery denoising," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 53, no. 5, pp. 2621–2639, 2015.

[6] L. R. Hu, J. G. Wu, and L. Wang, "Application and method for linear projective non-negative matrix factorization," *Computer Science*, vol. 40, no. 10, pp. 269–273, 2013.

[7] Y. Komai, D. H. Nguyen, T. Hara et al., "KNN search utilizing index of the minimum road travel time in time-dependent road networks," in *Proceedings of the 2014 IEEE 33rd International Symposium on Reliable Distributed Systems Workshops*, pp. 131–137, IEEE, Piscataway, NJ, USA, 2014.

[8] S. Ke, J. Gong, S. Li, Q. Zhu, X. Liu, and Y. Zhang, "A hybrid spatio-temporal data indexing method for trajectory databases," *Sensors*, vol. 14, no. 7, pp. 12990–13005, 2014.

[9] S. Ma, H. Wu, L. I. U. Zheng et al., "Method for emitter TDOA sorting based on recursive extended histogram," *Journal of National University of Defense Technology*, vol. 34, no. 5, pp. 83–89, 2012.

[10] J. Lin, "Brute force and indexed approaches to pairwise document similarity comparisons with mapreduce," in *Proceedings of the 32nd International ACM SIGIR Conference on Research and Development in Information Retrieval*, pp. 155–162, ACM, Boston, MA, USA, July 2009.

[11] R. Vernica, M. J. Carey, and C. Li, "Efficient parallel set-similarity joins using mapreduce," in *Proceedings of the 2010 ACM SIGMOD International Conference on Management of Data*, pp. 495–506, ACM, Indianapolis, IN, USA, 2010.

[12] C. Rong, W. Lu, X. Wang, X. Du, Y. Chen, and A. K. H. Tung, "Efficient and scalable processing of string similarity join," *IEEE Transactions on Knowledge and Data Engineering*, vol. 25, no. 10, pp. 2217–2230, 2013.

[13] D. Miorandi, S. Sicari, F. De Pellegrini, and I. Chlamtac, "Internet of things: vision, applications and research challenges," *Ad Hoc Networks*, vol. 10, no. 7, pp. 1497–1516, 2012.

[14] S. K. Chong, M. M. Gaber, S. Krishnaswamy, and S. W. Loke, "Energy conservation in wireless sensor networks: a rule-based approach," *Knowledge and Information Systems*, vol. 28, no. 3, pp. 579–614, 2011.

[15] L. M. Oliveira and J. J. Rodrigues, "Wireless sensor networks: a survey on environmental monitoring," *Journal of Communications*, vol. 6, no. 2, pp. 143–151, 2011.

Hindawi
Scientific Programming
Volume 2021, Article ID 9921551, 7 pages
https://doi.org/10.1155/2021/9921551

Research Article

Design and Implementation of Human-Computer Interaction System in Parallel Digital Library System Based on Neural Network

Jun Cao

Zaozhuang University Library, Zaozhuang 277100, China

Correspondence should be addressed to Jun Cao; 101024@uzz.edu.cn

Received 9 March 2021; Revised 26 March 2021; Accepted 18 April 2021; Published 17 May 2021

Academic Editor: Shah Nazir

Information and communication technologies are well thought-out as probable assets for the development of socioeconomics in developing countries. Studies have shown that enhanced infrastructure of telecommunication has facilitated means for underserved population development by various ways. Among the existing applications of ICT, the digital library systems provide with better solutions and respond to a variety of unmet needs of research institutions, scientific communities, and developments. With the development of digital library technology, the parallel database system has become the main tool for efficient information processing in the digital library system. On this basis, based on the parallel environment of the computer cluster, by coordinating the communication in the parallel environment, the coordinator, the collection machine, and the query processor can complete the operation of distribution, load, and maintenance, which has high efficiency and saves much precious time, supports the digital library to meet user requirements effectively, and meets the digital library's performance requirements for data, and also, the key problem in the parallel algorithm has been solved. The experimental results show that this parallel technique has very good performance and efficiency.

1. Introduction

Information and communication technologies (ICT) are well thought-out as possible assets for the development of socioeconomics in developing countries. Recent studies have exposed that higher infrastructure of telecommunication has facilitated means for underserved population development by various ways [1]. Among the existing applications of ICT, the digital library systems provide with better solutions and respond to a variety of unmet needs of research institutions, scientific communities, and developments. With the development of information technology, more and more information needs to be stored and disseminated, and the types and forms of information are more and more multiple. The mechanism of the traditional library is obviously unable to meet these needs [2]. Therefore, people put forward the idea of the digital library. The digital library is an electronic information storage, which can store a large number of various forms of information. Users can easily access it through the network to get these information, and its information storage and user access are not limited by geographical restrictions [3]. The digital library integrates all kinds of information, such as the data of information, storage management, and query and message posting, in which multimedia is involved so that information can be spread on the Internet to make information be used to maximum [4]. Through multimedia database technology and hypermedia technology and aiming at the characteristics of various media in the digital library, an effective and feasible management retrieval scheme in image retrieval, video-on-demand, and literature has been proposed [5, 6]. The digital library is an innovation based on the traditional library in the information era. It contains not only the traditional library functions, providing the corresponding service to the public, but also integrates some functions of other information resources (such as museums and archives); some of the features provide comprehensive public information service [7]. In other words, the digital library will become the public information center and hub of the future society [8].

Based on the parallel environment of the computer cluster, the communication in the parallel environment is controlled by the coordinator, so that the coordinator, collection machine, and query processor can complete the operation of distribution, load, and maintenance, which has high efficiency and saves much precious time, supports the digital library to meet user requirements effectively, and meets the digital library's performance requirements for data, and also, the key problem in the parallel algorithm has been solved [9]. The proposed study has considered the design and implementation of the human-computer interaction system in the parallel digital library system based on the neural network [10].

The paper is organized as follows. Section 2 shows the related work to the design and implementation of the human-computer interaction system in the parallel digital library system. Section 3 describes the methodology section of the paper. Section 4 briefly describes the results' analysis and discussion of the paper. The paper is concluded in Section 5.

2. Related Work

Our digital library workers have also done some work in the digital library, such as the research of the personalized active service system in the digital library system and query optimization algorithm. Although there are some works having been done in the field of digital library at home and abroad, the work is relatively scattered and preliminary [11]. The research that seemed the digital library as an independent and universal system tool has not been done and lacks of a comprehensive understanding of the digital library [12]. A lot of research is based on the use of the traditional database management system to implement the digital library system. It is noteworthy that there is no such report about the research of the parallel digital library [13]. Now, loading is a new research topic in the world. The research on parallel text loading is still at the initial stage, both at home and abroad. Although there are already some data-loading system prototypes at home and abroad, the research is mainly focused on the realization of two-dimensional relational data-loading technology of the single machine, and there is no parallel text-loading system. Moreover, the data loaded by these existing systems are mainly relational tables whose structure has been determined. No system can load fixed structure data according to the tables of the changing structure [14].

3. Methodology

3.1. Algorithm of the Collection Machine When the Digital Library System Adds Class. As the amount of data increases in the system, a new category needs to be added to the system at a certain stage. At this time, the classification system is dynamic. Which processor the new category should belong to is not predetermined, so we must first determine which processor the new category should belong to [15]. When the processor that a new class belonging to is determined, data related to the new class is sent to the corresponding processor instead of sending any message to other processors. When adding class, a special case is needed to be solved: the case of a minimum class on many data acquisition machines. In this way, each query processor should derive the thread according to the data acquisition machine, and these threads need to simultaneously receive data collected from different machines, more importantly, when the data is inserted into the same Oracle table, because multiple threads cannot simultaneously insert data to the same table, so it is necessary to with the help of intermediate files. These intermediate files are not merged into one big file until all threads input the data to the intermediate file, respectively; then, data merged is inserted into the Oracle table [16]. If the file on the query processor is a sharing file, such as many data acquisition machines input data into a certain file on the query processor at the same time, it also needs to input data to temporary files separately; finally, these temporary files are merged to form the final file [17].

The idea of the algorithm is that the data acquisition machine does not do any operation if the addition is not the smallest class, that is, it is not the leaf node in the classification pattern table. If the class is a leaf node, when the acquisition machine is receiving command of the adding class, it can also receive which query processor the new class should belong to, which is judged by the front-end machine. After the data is extracted, the extracted data is sent to the related query processor [18].

The function of the algorithm is to add new classes from the acquisition machine to the query processor. Input is a class of coding, processor number belonging to, and a string of number in the class composed of the number of text. The output returns 0 when succeeds. The process is to separate the parameters into class code, the processor number that the class should belong to, and the number of text in the class and store it in the array list of the structure [19]:

$$P(X_i, Y_i)P(Y_{i+1}|X_{i+1}) = P(X_i, Y_{i+1})P(Y_I|X_I). \quad (1)$$

Then, read the network configuration file; the processor that the new add class should belong to is read, connecting to the IP address of the processor. If a new type of class needing to add files is only one, send the file name and the contents of the file in order; if a new type of class to add the files is multiple, first find the first file. After loading the text, you can view the detailed steps of the processor operation [20]:

$$\operatorname{sim}(u_i, u_j) = \frac{\sum_{c \in I_{ij}} (R_{ic} - \overline{R})}{\sqrt{\sum_{c \in I_{ij}} (R_{ic} - \overline{R})^2}}. \quad (2)$$

For the analysis of algorithmic complexity, the number of minimum classes in the system is set as n, and the average amount of data under each minimum class is x, and the number of data acquisition machines is c. The number of the query processors is p. The time complexity of the worst case of the algorithm is o(x), and the time complexity in the best case is O(x/cP):

$$P(S) = \frac{\mathrm{Sim}(T_{rw}, S_{rw})}{\mathrm{Sim}(T_{uw}, S_{uw})}. \tag{3}$$

With a random process $\{X_n, n \in T\}$, if for any integer $n \in T$ and arbitrary $i_{0,} i_1, \cdots, i_n \in I$, the conditional probability satisfies

$$\begin{aligned} &P\{X_{n+1} = i_{n+1} * X_0 = i_0, X_1 = i_1, \ldots, X_n = i_n\} \\ &\quad = P\{X_{n+1} = i_{n+1} * X_n = i_n\}. \end{aligned} \tag{4}$$

Then, $\{X_n, n \in T\}$ would be called the MARKOV chain, which indicates that the probability of the system will be transferred to j at the moment of $n + 1$ when the system is in the state of i at time n; put P_{ij} in order to obtain the following matrix:

$$P = \begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1n} \\ p_{21} & p_{22} & \cdots & p_{2n} \\ & & & \\ p_{n1} & p_{n2} & \cdots & p_{nn} \end{bmatrix}. \tag{5}$$

This matrix is called the transition probability matrix. The state space I of any system can be decomposed into the following disjoint subsets:

$$I = C_1 + C_2 + \cdots + N. \tag{6}$$

In which, P_{ij} is composed of all the very return to the state set and $C_i\,(i = 1, 2, \cdots)$ is the reciprocal of the return of the state that is often composed of the closed set. If j is the aperiodic normal return state, then

$$\lim_{n \longrightarrow \infty} P_{ij}^{(n)} = \frac{1}{U_t}. \tag{7}$$

The variable U_j in this formula is the average return time of the state j.

We call the probability distribution $\{\pi_j, j \in I\}$ a smooth distribution of MARKOV chains, where I is the state space if it satisfies the following conditions:

$$\pi_i = \sum_{i \in I} \pi_i P_{ij}, \tag{8}$$

$$\sum_{i \in I} \pi_j, \pi_j > \; = 0. \tag{9}$$

If $\{\pi_j, j \in I\}$ is the smooth distribution of the MARKOV chain, then

$$\lim_{n \longrightarrow \infty} P_j(n) = \frac{1}{U_j} = \pi_j. \tag{10}$$

So from the data sequence changes, Start at time t_1 of the state, only the state at time t_n can predict the probability, the data sequence is divided into several states, recorded as $i_1, i_2, \cdots, i_n$, the probable transfer time be recorded as $t_1, t_2, \cdots, t_n$, the transition probability of the data sequence of the state at time i_k transitioning to step is expressed as $P_{ij}^{(m)}$:

$$P_{ij}^{(m)} = \frac{M_{ij}^{(m)}}{M_j}, \tag{11}$$

where M_{ij} is the number of times the state i_k transitions to the state i_y after m steps and M_i is the number of occurrences of the state i_k. As the uncertainty of the final state of the data sequence is steering, the last one data should be removed when calculating M_1.

3.2. The Maintenance of Data in the Parallel Digital Library System. The parallel data manipulation subsystem (PDOP)S provides basic query operations based on multiple data distribution strategies and parallel storage structures, such as one-dimensional data partition, multidimensional data partition, and compressed multidimensional array storage structure and attribute partitioning storage structure [21]. The parallelism of all algorithms is based on parallel data, and it is easy to be implemented in the computer parallel cluster environment. All the implementation environment of the parallel algorithm is the computer cluster parallel environment that is composed of several ordinary PC computers; through the parallel high-speed, network processors are connected together, one of them as the front-end processor and coordination machine. The random processor is used as the back-end data acquisition machine, and the arbitrary processor is the final processor of the system administrator [22]. The front-end machine does not store any data and is only used to receive operation requests from multiple users and coordinate the execution of commands by threads on each back-end machine. The data is transferred from the acquisition machine to store in the query processor on the back-end machine according to a certain data distribution method. The back-end machine is responsible for the specific execution. When the parallel algorithm is executed, the scheduling module on the front-end machine will coordinate all the back-end query processors and the operation execution modules on the data acquisition machine to work in parallel. This parallelism is achieved by performing the same operation on different data items, so it is a data parallel [23].

The implementation of the parallel data-loading algorithm is divided into two phases. The first phase is the data division phase. The second phase is the operation execution phase. In the first phase, the scheduling module first receives the execution information from the system administrator and calculates the data distribution strategy locally. The purpose of data distribution is to uniformly distribute data objects on a certain data acquisition machine to multiple query processors so that parallelization of the system can be fully realized during query processing. Data distribution is an important and active field in the research of the parallel database system at present. There are several methods of data distribution in a parallel database system. The one-dimensional data distribution method is the simplest way of data distribution. By partitioning the domain values of one attribute, the whole relationship is partitioned, and a set of subrelationships is obtained, and then, these subrelationships are distributed

among multiprocessors. At present, the one-dimensional data distribution method mainly includes Round−Robin, Hash, Range−partition, and Hybrid−Range−Partition. The one-dimensional data distribution method has a common problem: it is not able to effectively support queries with choice predicates on nonpartition attributes. In order to solve this problem, some multidimensional distribution methods have been put forward, including CMD method, ECC data distribution method, BM data distribution method, FX data distribution method, data distribution method based on the iHlbert curve, and BERD multidimensional data distribution method. Under the control of data distribution strategy, the threads of operation execution modules are derived from all the back-end query processors and data acquisition machines, and the operation execution information is broadcast to the threads of each back-end machine. After the operation information is received by the thread of the operation execution module on the data acquisition machine, the local data is transferred to the query processor.

Metadata is an important part of the digital library. The quality of metadata determines the quality management of the whole digital library. Metadata is stored in the metadata table. It is not only an important data material of the digital library to be used by workers but also can be used for querying of users so that the structure of the digital library can be better understood and its use level can be improved. The idea of the algorithm is that when modifying metadata, new metadata of modified metadata is stored in the Oracle tables of each back-end machine query processor, and new values in the table replace the old values of metadata tables. Because the new value table and metadata table both exist on all query processors, the algorithm only involves every query processor, and the algorithm on each query processor is exactly the same.

The algorithm modifies the back-end machine metadata information, and input is empty. The output returns 0 when it succeeds. Then, the name of the machine is obtained. Open the network configuration file and read it from the second records. If the name of the machine is the same as the name of a certain machine recorded of the network configuration file, the logical name of the machine is obtained. Using a cursor from the machine to select data for the table prepared for the update metadata, the condition is that the logical name of the machine is the same as the logical name of the machine. When data can be extracted from the cursor, update the title of the machine by using the following items, including document identifier, abstract, author, department, other information of the author, publisher, publication time, input time, page number, ISBN, and category code. Finally, return successfully.

4. Result Analysis and Discussion

4.1. Test One. The first experiment of the multithread parallel text-loading algorithm and serial algorithm in the query processor number is fixed, and the experimental results when the amount of data changing occurs is in the first experiment, and the back-end machine we use includes two sets of data acquisition machine and four sets of query processors; each back-end machine configuration is IG memory and 70G hard disk. The amount of data in this group of experiments is constantly changing, and the number of the back-end machines is fixed. The purpose of the experiment is to compare the efficiency of serial and parallel algorithms when the amount of data is increased. The following are the data of the test and the performance analysis table according to the test data. Performance analysis table is shown in Table 1.

Table 1 is the experimental data of the multithread parallel text-loading algorithm described in this article on 4 node machines. Row is the data volume, line is the parallel, serial different loading algorithms, and table content describes the running time of the algorithm. Comparison of parallelism and serialization when data quantity changes is shown in Figure 1.

We can see from the chart that when using the serial algorithm, although we use four queries for data processor loading at the same time, the query processor is working in the serial mode, that is to say, when the work of a query processor is over, the other one can work. In this way, the total load time will be the sum of the loading time of all processors, not only does the speed do not improve but also the communication between the processors should also be considered. When using parallel algorithms, four query processors will work at the same time, so the total time of loading will be the slowest processor loading time. As we can see in the graph, when the number of processors is constant, the parallel algorithm is much faster than the serial algorithm with the increase of data volume, and the total cost is about 1/3 of the serial algorithm.

4.2. Test Two. Experiment two is the experimental results of a multithread parallel text-loading algorithm and a serial algorithm when the data is fixed and the number of the query processor changes. When the amount of data to be loaded and the number of the query processors change, we use the 19.7 GB data, and the results are shown in the chart below. Line is the number of processors, row is parallel, serial different loading algorithms, and the table content describes the running time of the algorithm. Algorithm running time is shown in Table 2. Comparison of multiple threads and single thread when number of processors changes is shown in Figure 2.

The experimental results show that when using the serial text-loading algorithm, the more the number of the query processor is used, the lower the efficiency of the algorithm is. Since using the serial algorithm, the query processor still needs to load data one by one, and the latter query processor must wait until the last query processor is finished. And, the overhead of the algorithm also includes the communication time between the processors. But when using the parallel algorithm, the more the query processor is used, the higher the efficiency of the algorithm is because all the data query processor will be loaded at the same time, and the loading time is less than the communication time between processors, so increasing the communication overhead between

Table 1: Performance analysis table.

	624 mega bytes	1380 mega bytes	1710 mega bytes	1970 mega bytes
Parallelism serialization	5.65 minutes and 12.55 minutes 3886 mega bytes	21.5167 minutes and 38.05 minutes 5668 mega bytes	23.2833 minutes and 50.433 minutes 13877 mega bytes	31.9833 minutes and 117.45 minutes 19564 mega bytes
Parallelism serialization	49.67 minutes and 142.8 minutes	73.54 minutes and 216.433 minutes	135.85 minutes and 396.7 minutes	233.76 minutes and 723.4 minutes

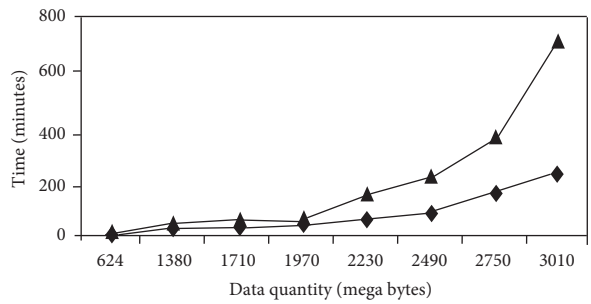


Figure 1: Comparison of parallelism and serialization when data quantity changes.

Table 2: Algorithm running time.

	Multithreads	Single thread
2	50.1	59.733
4	31.983	64.1
6	26.088	73.45
8	22.57	83.77
10	19.66	95.89

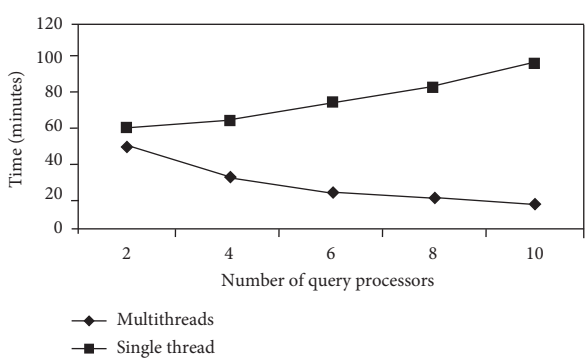


Figure 2: Comparison of multiple threads and single thread when number of processors changes.

processors is worth it. As we can see from the results, the efficiency of the parallel algorithm will be significantly improved when the number of processors increases.

4.3. Test Three. The experimental results of the multi-threaded text parallel loading algorithm and single-thread loading algorithm are fixed in the number of query processors and data volume, and the number of data acquisition machine changes. When the amount of data to be loaded and the number of query processors are fixed and the number of data acquisition machines changes, we use 19.7 GB data and 4 query processors. The experimental results are shown below. Among them, line is the number of data acquisition machines, the row is whether each data acquisition machine and query processor can derive multiple threads and single-threaded loading algorithms, and tabular content describes the running time of the algorithm. The running time of the algorithm is shown in Table 3. Comparison of multiple threads and single thread when number of processors changes is shown in Figure 3.

The experimental results show that when the single threaded and multithreaded text parallel loading algorithms are used, the total execution time decreases with the increase of the number of data acquisition machines. When using a single thread loading algorithm, four sets of query processors are used; each query processor only derived a thread to receive the data of a data acquisition. Each data acquisition machine is derived from a thread to a query processor data; after a data acquisition machine finishes sending, the query processor receives the data from second sets; when using the multithread loading algorithm, each query processor derived M threads according to the data acquisition machine number; each data acquisition machine derived N threads according to the number of query processor so that each data acquisition machine with N threads simultaneously sends data to N query processors. And, each query processor with M threads simultaneously receives M data acquisition data, so the multithread loading algorithm will be much faster than the single thread algorithm. The running time prediction of the algorithm is shown in Table 4. Processor number prediction results were compared and are shown in Figure 4.

Table 3: The running time of the algorithm.

	Multithreads	Single thread
1	31.983	64.1
2	26.76	57.8
4	19.85	50.77
6	15.7	41.99
8	14.39	34.34
10	9.6	29.37

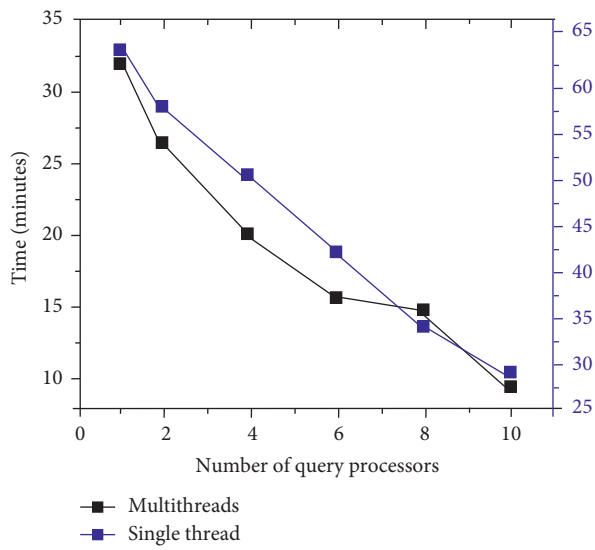


Figure 3: Comparison of multiple threads and single thread when number of processors changes.

Table 4: The running time prediction of the algorithm.

	Multithreads	Single thread
1	33	69.22
2	28.69	62.30
4	25.14	56.15
6	15.69	45.36
8	18.69	36.33
10	18.66	36.21

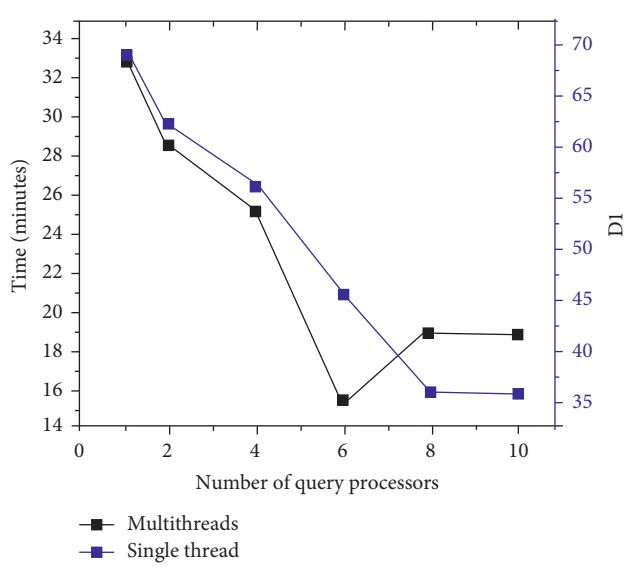


Figure 4: Processor number prediction results were compared.

5. Conclusion

With the development of digital library technology, the parallel database system has become the key tool for efficient information processing in the digital library system. Data-loading operation is a significant part of the digital library, which is well known that the data-loading operations are time-consuming. Data loading, especially the loading of parallel text data, is a new field of research. In this paper, a novel data operation algorithm based on a new parallel digital library is proposed, and all the data operation algorithms are implemented in the prototype system. A multithread parallel text data-loading operation and maintenance algorithm has been proposed in this paper, which has no research about it up to now. A large number of experiments show that the algorithm proposed in this paper is more efficient than the existing algorithms and has high practical value. Considering the performance and price ratio, the parallel algorithm has high practical value and benefit. To sum up, the author thinks it is still a large area having many problems that need to be solved. There are many works to do, hoping that experts and scholars will pay enough attention to it.

Data Availability

The datasets used and/or analyzed during the current study are available from the corresponding author upon reasonable request.

Additional Points

This is a research involving human participants and/or animals.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

[1] N. Yu, S. Li, Y. Zhao et al., "Design and implementation of a dexterous human-robot interaction system based on haptic shared control (in Chinese)," *Chinese Journal of Scientific Instrument*, vol. 38, no. 3, pp. 602–611, 2017.

[2] Y. Park, H. Yang, T. Dinh et al., "Design and implementation of a container-based virtual client architecture for interactive digital signage systems," *International Journal of Distributed Sensor Networks*, vol. 13, no. 7, p. 155, 2017.

[3] Y. J. Zhang, X. J. Meng, and G. Wang, "Design and implementation of 3D electronic sand table system based on gesture interaction," *Command Control & Simulation*, vol. 49, no. 21, pp. 706–712, 2016.

[4] W. K. Liou, K. K. Bhagat, and C. Y. C. Y. Chang, "The design, implementation, and evaluation of a digital interactive globe system integrated into an earth science course," *Educational Technology Research & Development*, vol. 53, no. 11, pp. 1201–1209, 2018.

[5] R. Habel, F. Silber-Chaussumier, F. Irigoin et al., "Combining data and computation distribution directives for hybrid parallel programming: a transformation system,"

International Journal of Parallel Programming, vol. 44, no. 6, pp. 1–28, 2016.

[6] D. Huang, D. Han, J. Wang et al., "Achieving load balance for parallel data access on distributed file systems," *IEEE Transactions on Computers*, vol. 108, no. 99, p. 1, 2017.

[7] M. Hermanns and E. Cramer, "Likelihood inference for the component lifetime distribution based on progressively censored parallel systems data," *Journal of Statistical Computation and Simulation*, vol. 87, no. 3, pp. 607–630, 2016.

[8] T. C. Pan, P. Flick, C. Jain et al., "Kmerind: a exible parallel library for k-mer indexing of biological sequences on distributed memory systems," *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, vol. 16, 2019.

[9] X. Shen, L. Yi, Y. Yi et al., "Dynamic identifying protein functional modules based on adaptive density modularity in protein–protein interaction networks," *BMC Bioinform*, vol. 16, no. 12, pp. 5–12, 2015b.

[10] B. P. Tu, A. Kudlicki, M. Rowicka et al., "Logic of the yeast metabolic cycle: temporal compartmentalization of cellular processes," *Science*, vol. 310, no. 5751, pp. 1152–1158, 2005.

[11] S. Wang and W. Guo, "Sparse multigraph embedding for multimodal feature representation," *IEEE Transactions on Multimedia*, vol. 19, no. 7, pp. 1454–1466, 2017a.

[12] S. Wang and W. Guo, "Robust co-clustering via dual local learning and high-order matrix factorization," *Knowledge-Based Systems*, vol. 138, pp. 176–187, 2017b.

[13] J. Wang, X. Peng, M. Li, and Y. Pan, "Construction and application of dynamic protein interaction network based on time course gene expression data," *Proteomics*, vol. 13, no. 2, pp. 301–312, 2013.

[14] J. Wang, J. Liang, X. Zhao, and W. Zheng, "Overlapping protein complexes detection algorithm based on assortativity in PPI network," *Computer Sciences*, vol. 46, pp. 294–300, 2019.

[15] X. Xiao, J. Ji, and C. Yang, "Fireworks algorithm for functional module detection in protein–protein interaction networks," *Journal of Harbin Institute of Technology*, vol. 51, pp. 57–66, 2019.

[16] Y. Zhang, K. Jia, and A. Zhang, "Consistent protein functional module detection from multi-view of biological data," *Acta Electronic Sinica*, vol. 42, no. 12, pp. 2337–2344, 2014.

[17] Y. Zhang, H. Lin, Z. Yang et al., "A method for predicting protein complex in dynamic PPI networks," *BMC Bioinformatics*, vol. 17, no. 7, pp. 229–239, 2016.

[18] B. Zhao, H. Xiong, W. Ni et al., "Improved weighted-network based algorithm for predicting protein complexes," *Computer Science*, vol. 41, no. 6, pp. 231–234, 2014a.

[19] B. Zhao, J. Wang, M. Li et al., "Detecting protein complexes based on uncertain graph model," *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, vol. 11, no. 3, pp. 486–497, 2014b.

[20] B. Zhao, X. Li, S. Hu et al., "Prediction of protein functions based on essential functional modules mining," *Acta Automatica Sinica*, vol. 44, no. 1, pp. 183–192, 2018.

[21] W. P. Zheng, J. Y. Li, and J. Wang, "Protein complex recognition algorithm based on genetic algorithm," *Journal of Computer Science and Technology*, vol. 12, no. 5, pp. 794–803, 2018.

[22] J. Zhong, J. Wang, W. Peng et al., "A feature selection method for prediction essential protein," *Tsinghua Science and Technology*, vol. 7, no. 10, pp. 491–499, 2015.

[23] W. Zhu, W. Guo, Z. Yu, and H. Xiong, "Multitask allocation to heterogeneous participants in mobile crowd sensing," *Wireless Communications and Mobile Computing*, vol. 2018, Article ID 7218061, 10 pages, 2018.

Hindawi
Scientific Programming
Volume 2021, Article ID 9956710, 10 pages
https://doi.org/10.1155/2021/9956710

Research Article

Machine Vision and Big Data-Driven Sports Athletes Action Training Intervention Model

Hui Jiang,[1] Ping wang,[2] Lei Peng,[3] and Xiaofeng Wang[4]

[1]*Institute of Physical Education, Dezhou University, Dezhou 253023, Shandong, China*
[2]*Baoding Vocational and Technical College, Baoding, 071000, China*
[3]*College of Physical Education, Hengshui University, Hengshui 053000, Hebei, China*
[4]*Sports Department of Hebei Vocational College of Rail Transportation, Tianjin, Hebei, China*

Correspondence should be addressed to Xiaofeng Wang; wangxiaofengdeqq@163.com

Received 8 March 2021; Revised 8 April 2021; Accepted 26 April 2021; Published 17 May 2021

Academic Editor: Shah Nazir

In recent years, athlete action recognition has become an important research field for showing and recognition of athlete actions. Generally speaking, movement recognition of athletes can be performed through a variety of modes, such as motion sensors, machine vision, and big data analysis. Among them, machine vision and big data analysis usually contain significant information which can be used for various purposes. Machine vision can be expressed as the recognition of the time sequence of a series of athlete actions captured through camera, so that it can intervene in the training of athletes by visual methods and approaches. Big data contains a large number of athletes' historical training and competition data which need exploration. In-depth analysis and feature mining of big data will help coach teams to develop training plans and devise new suggestions. On the basis of the above observations, this paper proposes a novel spatiotemporal attention map convolutional network to identify athletes' actions, and through the auxiliary analysis of big data, gives reasonable action intervention suggestions, and provides coaches and decision-making teams to formulate scientific training programs. Results of the study show the effectiveness of the proposed research.

1. Introduction

Estimation of sports athletes' action pose [1–3] is an emerging frontier research direction. At present, most of them are based on traditional image processing methods. The research steps of human body posture estimation based on neural network are firstly to detect the human body on the input sports athlete images, secondly to estimate the athlete's action posture, and finally evaluate the estimation results and compare the correct posture to give action intervention [4] suggestions. The human body detection of sports athletes is an important step in constructing a neural network for motion posture estimation. It usually uses related algorithms to determine whether there is a target research object in the input image. If there is, mark its specific location in the picture and use a border or red box subject circled. The research of motion pose estimation first needs to detect the athlete's human body, and the input picture is used to extract the bounding box of the athlete's portrait through a specific human body detector.

Since the single-person pose estimation target has only a single object, the recognition technology is relatively mature and the recognition speed is faster. TOSHEV et al. [5] first proposed the DeepPose method based on neural networks, which proposed a cascaded DNN-like regression, which can achieve high-precision estimation. The advantage of this method is to estimate the attitude in a holistic manner and has good generalization performance; TOMPSON et al. [6] proposed a new hybrid architecture, which is composed of deep convolutional neural networks and Markov random fields. Including an additional "torso joint heat map" to merge data to select the correct feature activation in the chaotic scene, YANG et al. [7] presented a new end-to-end human pose estimation framework, combining DCNN with parts. Expressible deformations are mixed together, DCNN is used to return the heat map of each body part, and the

structured output of deep learning of neural network is used to further simulate the relationship between body joints.

The above work proves the powerful performance of machine vision in the recognition of athletes' action gestures, so it is effective to use machine vision neural networks to participate in the intervention of sports athletes' action recognition. However, the existing machine vision-based methods rarely consider the combination of sports athletes' historical training and competition big data for analysis. Therefore, based on this shortcoming, we propose a novel combination of machine vision and big data analysis of athletes' postures. The estimation method can accurately estimate the motion posture. In addition, we also calculate the loss of the estimation result and the correct action posture. Following is the main innovative points of this paper:

This paper recommends a novel machine vision method to recognize sports athletes' actions. Among them, the machine vision method is realized by spatiotemporal attention map convolutional network.

This paper combines machine vision methods with big data technology. Because machine vision methods are used alone, even if they can accurately recognize action gestures, it is difficult to give intervention suggestions. Therefore, this paper considers athletes' historical training and competition big data. Auxiliary analysis is helpful for the formulation of training action intervention programs.

On the basis of the machine vision method, this paper proposes to use big data analysis to analyze athletes' historical training and competition data and to recognize losses in computer actions in order to develop scientific intervention strategies.

We have conducted sufficient comparative experiments and ablation studies. The athlete's movement intervention method based on machine data and big data analysis can provide a scientific basis for sports coaches to develop a reasonable training plan.

The organization of the paper is given as follows: Section 2 of the paper represents the related research in the area. Section 3 describes the methodology of the paper with the details of the research done. Section 4 shows the experiments and results of the study conducted for the proposed study. The paper is concluded in Section 5.

2. Related Research

Song et al. [8] proposed a deeply structured model to predict human pose sequences in unconstrained videos. The model can be effectively trained in an end-to-end manner and can simultaneously represent the appearance of body joints and their temporal and spatial relationships. Hossain et al. [9] used the time information on the 2D joint position sequence to estimate the 3D pose sequence. A sequence-to-sequence network composed of layered standardized LSTM units is also designed. The network has shortcut connections to connect the input to the output on the decoder side and imposes time smoothing constraints during the training process. Pavllo et al. [10] proved that a complete convolution model based on temporal convolution on 2D key points can effectively estimate the 3D pose in the video. It also introduces back projection, which can start from the 2D key point prediction of the unmarked video, then estimate the 3D pose, and finally back-project it to the input 2D key point.

CAO et al. [11] proposed an effective method for detecting multiperson poses in images. This method has high accuracy on multiple public benchmarks and greatly exceeds the multiperson detection level of MPII dataset in performance and efficiency. This method expresses bottom-up association scores through partial affinity fields (PAF). PAF is a set of two-dimensional vector fields that can encode the position and direction of limbs in the image domain, allowing bottom-up analysis, and can maintain high precision in real time at the same time and is not affected by the number of people in the image. Chen et al. [12] used a top-down approach to propose a new neural network structure called cascaded pyramid network (CPN), which includes two stages: GlobalNet and RefineNet. GlobalNet is a functional pyramid network that can accurately locate "simple" joint points (such as eyes and hands) but cannot accurately identify occluded or invisible joint points; RefineNet integrates all the feature representations of GlobalNet and mines the loss of online joint points and explicitly handles hard joint points.

In view of the above-mentioned research, the literature proves that machine vision has excellent performance in human body posture, which is superior to traditional image processing algorithms. Therefore, this also proves the rationality of the machine vision method proposed in this paper.

3. Methodology

Estimation of sports athletes' action pose [1–3] is an emerging frontier research direction. At present, most of them are based on traditional image processing methods. As shown in Figure 1, the general procedure is to first obtain the front and side images of the person and then extract the outline of the human body through image processing. Identify the key size points, then establish a function model of the human body dimension curve through statistical analysis and curve fitting, and import the complete athlete action data record table into the large sports action database after measurement by related auxiliary tools. With the fast development of computer vision technology [13–18], human body posture estimation has begun to be researched with neural network models [19–23], which has significantly improved the accuracy and robustness of human body posture estimation, has expanded the scope of application, and has been deeply integrated into sports competition and sports training.

Figure 2 is a flowchart of the overall architecture of our MVBD-Net algorithm. First, we extract the human pose from the motion images of sports players. Secondly, the adjacency matrix of the graph is constructed to feed the graph convolutional neural network to obtain the posture feature, and then it is input into the fully connected layer and the Softmax function to obtain the output of the action posture estimation. Secondly, this article also uses big data analysis technology to obtain the weight of the pose estimation and perform feature

Figure 1: Example of human pose estimation.

fusion with the output of the graph convolution network and finally calculate the loss with the label.

3.1. Human Detection. For football player training image data, this article first conducts human body detection. Human body detection is an important step in constructing a neural network for human posture estimation. It usually uses related algorithms to determine whether there is a target research object in the input image. If there is, mark its specific location in the picture, and circle the target object with a frame or red box stand up. Human body pose estimation research first needs human body detection. The input picture is used to extract the bounding box of the person through a specific human body detector. Common human body detectors include Yolo and R–CNN.

Since multiperson pose estimation cannot determine the specific location and total number of people in the image, it is much more difficult to implement than single-person pose estimation. There are usually two methods to achieve multiperson pose estimation: (1) first provide a human detector, then estimate each component separately, and finally get the pose of each person. This method is a top-down method; detect all the parts in the image, and then associate and group different parts. This method is a bottom-up method. Therefore, this article uses the YOLO algorithm for human detection (as shown in Figure 3).

YOLO loss calculation during training is shown in the following equation:

$$\begin{aligned}
&\lambda_{\text{coord}} \sum_{i=0}^{S^2} \sum_{j=0}^{B} \Pi_{ij}^{\text{obj}} \left[(x_i - \hat{x}_i)^2 + (y_i - \hat{y}_i)^2 \right] \\
&+ \lambda_{\text{coord}} \sum_{i=0}^{S^2} \sum_{j=0}^{B} \Pi_{ij}^{\text{obj}} \left[\left(\sqrt{w_i} - \sqrt{\hat{w}_i} \right) + \left(\sqrt{h_i} - \sqrt{\hat{h}_i} \right)^2 \right] \\
&+ \sum_{i=0}^{S^2} \sum_{j=0}^{B} \Pi_{ij}^{\text{obj}} \left(C_i - \hat{C}_i \right)^2 \\
&+ \lambda_{\text{coord}} \sum_{i=0}^{S^2} \sum_{j=0}^{B} \Pi_{ij}^{\text{noobj}} \left(C_i - \hat{C}_i \right)^2 \\
&+ \sum_{i=0}^{S^2} \Pi_{ij}^{obj} \sum_{c \in \text{classes}} \left(p_i(c) - \hat{p}_i(c) \right)^2,
\end{aligned} \tag{1}$$

where x, y, w, C, p is the network predicted value, $\hat{x}, \hat{y}, \hat{w}, \hat{C}, \hat{p}$ is the labeled value, Π_i^{obj} indicates that the object falls into the grid i, and Π_{ij}^{obj} and Π_{ij}^{noobj}, respectively, indicate that the object falls and does not fall into the jth bounding box of the grid i.

3.2. Spatiotemporal Attention Graph Convolutional Network

3.2.1. Graph Convolutional Network. The human body motion posture is in the form of graphics, rather than two-dimensional or three-dimensional grids, which makes it difficult to simply use convolutional networks. GCN (as shown in Figure 4) is a general and effective framework for learning to represent graph structure data. Various GCN variants have achieved the most advanced results on many tasks. For skeleton-based human behavior recognition, skeleton-based data can be obtained from motion capture devices or pose estimation algorithms in videos. Usually the data is a frame sequence, and each frame will have a set of joint coordinates. On the basis of a given sequence of human joints in the form of two-dimensional or three-dimensional coordinates, a spatio-temporal graph with joints as graph nodes and the natural connectivity of human body structure and time as graph edges is constructed. Graph convolution can be defined as follows:

$$f_{\text{out}}(v_{ti}) = \sum_{v_{ti} \in N} \frac{1}{z_{ti}(v_{ti})} f_{in}(v_{ti}) W(l_{ti}(v_{ti})), \tag{2}$$

where f_{in} is the feature vector input of node v_{ti} and W is a weight function and is mapped from the graph label l_{ti}: V_{t} from K, which can be used to assign a label to each graph node v_{ti}.

3.2.2. Spatiotemporal Attention. In the algorithm model of this paper, a new spatiotemporal attention mechanism is proposed, which captures the dynamic spatiotemporal correlation on the athlete's action network. This mechanism contains two kinds of attention, namely, spatial attention and temporal attention.

Space Attention. In the spatial dimension, the conditions of nodes at different locations influence each other, and the mutual influence is very dynamic.

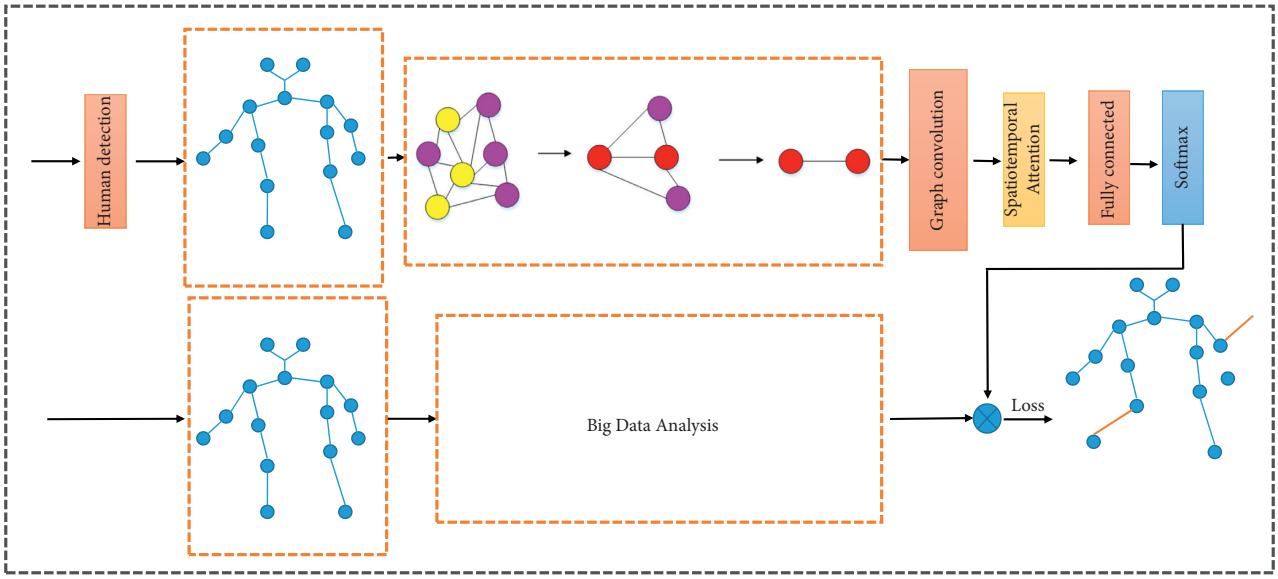


Figure 2: The flowchart of the overall architecture of our MVBD-Net algorithm.

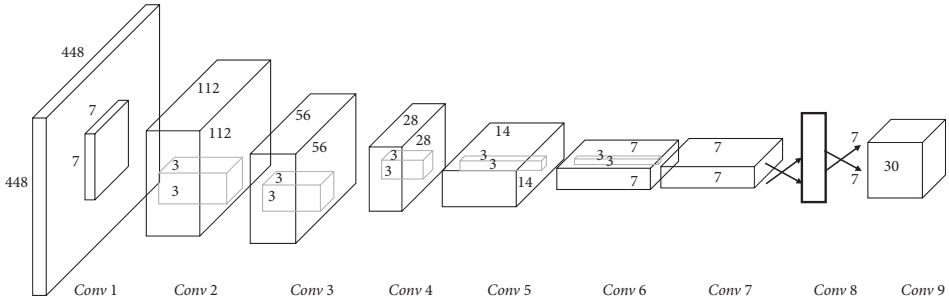


Figure 3: The flowchart of the overall architecture of Yolo.

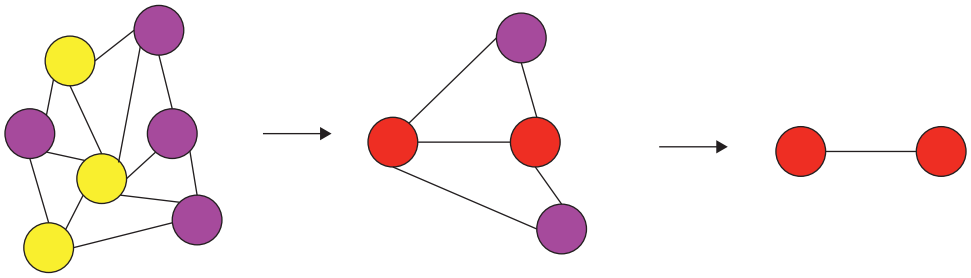


Figure 4: Graph convolutional neural network process.

$$S = V_s \times \sigma\left(\left(x_h^{r-1} w_1\right) w_2 \left(w_3 x_h^{r-1}\right)^T + b_s\right), \tag{3}$$

$$S_{ij}' = \frac{\exp\left(S_{ij}\right)}{\sum_{j=1}^{N} \exp\left(S_{ij}\right)}, \tag{4}$$

where x_h^{r-1} represents the input of the r th space-time block and $_C^{r-1}$ is the number of channels of input data in the r th layer. The attention matrix S is dynamically calculated based on the current input of this layer. T_{r-1} represents the length of the time dimension in the r th layer, V_s, b_s, w_1, and w_2, and w_3 represent the learning parameters, and σ represents the activation function.

Time Attention. In the time dimension, there is a correlation between the conditions of human joints in different time periods, and the correlation is also different in different situations. Similarly, this paper uses the attention mechanism to adaptively give different attention to the data.

$$E = V_e \times \sigma\left(\left(x_h^{r-1}\right)^T u_1\right) u_2\left(u_3 x_h^{r-1} + b_e\right), \tag{5}$$

$$E_{ij}' = \frac{\exp\left(E_{ij}\right)}{\sum_{j=1}^{T_{r-1}} \exp\left(E_{ij}\right)}, \tag{6}$$

where V_e, b_e, u_1, u_2, u_3 is the learning parameter and the time correlation matrix E is determined by the changing input. The value of an element E_{ij} j in E semantically represents the strength of dependence between time i and time j, and finally E is normalized by the Softmax function.

3.2.3. Spatiotemporal Attention Fusion. The spatial attention network generates the spatial attention heat map to guide the action classification network of football players to extract effective spatiotemporal features from the spatial region of interest. The time attention mechanism automatically mines discriminative time-domain video clips from the original training or competition videos and uses these video clips for network training, while eliminating the interference of other video clips on the classifier. The proposed spatiotemporal attention model separately trains two network models in RGB video frames and optical flow sequences, namely, spatial subnetwork (SN) and temporal subnetwork (Temporal Network, TN). Then, the weighted fusion of the Softmax prediction scores of the two networks is used as the basis for behavior classification, which can effectively improve the classification robustness of the network. It should be noted that when training on RGB video frame data, the parameters of the spatial attention network need to be pretrained on the optical flow prediction database in advance, and when training on optical flow sequence data, the spatial attention network and the behavior classification network share weights.

4. Experiments and Results

In this article, we collected 1000 training images of sports figures from the Leeds Sports Pose dataset and annotated 14 joints. Those images are challenging due to the different appearance and strong sharpness. The images in the Leeds Sports Pose dataset have been scaled so that the most prominent figures are approximately 150 pixels tall. Although each image in Leeds Sports Pose may contain multiple people, standard preprocessing for human detection has been performed to extract a single person. As in previous works, we use the subimages of these detected individuals as training and testing samples. In this way, the training and testing data contains only one person, and as mentioned earlier, in the testing phase, we only use the entire image (for the Leeds Sports Pose dataset, this means the entire subimage of a person) as a body patch.

4.1. Evaluation Methods. oks (object key point similarity) is the evaluation index of the commonly used human bone key point detection algorithm. This index is inspired by the IoU index in target detection. The purpose is to calculate the truth value and predict the similarity of the key points of the human body. The formula is as follows:

$$O_p = \begin{cases} KS = e^{\dfrac{\left\|\hat{\theta}_i^P - \theta_i^P\right\|_2^2}{2s^2k_i^2}}, \\ OKS = \dfrac{\sum_i \exp\left\{-d_{pi}^2/2S_p^2\sigma_i^2\right\}\delta\left(v_{pi} > 0\right)}{\sum_i \delta\left(v_{pi} > 0\right)}, \end{cases} \tag{7}$$

where p represents the person with id p among all ground truth pedestrians in the current picture, $p \in (0, M)$, and M represents the number of pedestrians in the current picture. Since the training and test data contains only one person, M here is set to 1. i represents the key point with id i. d_{pi} represents the Euclidean distance between the key point with id i in the set of key points currently detected and the key point with id p in the ground truth pedestrian.

$$d_{pi} = \sqrt{\left(x_i^{\sim} - x_{pi}\right)\left(y_i^{\sim} - y_{pi}\right)}, \tag{8}$$

where $(x_i^{\sim}, y_i^{\sim})$ is the current key point detection result and (x_i, y_i) is the ground truth.

S_p represents the scale factor of the person with id p in the ground truth pedestrian, and its value is the square root of the area of the pedestrian detection frame:

$$S_p = \sqrt{wh}, \tag{9}$$

where w and h are the width and height of the detection frame.

σ_i represents the key point normalization factor of type i. This factor is the standard deviation between the ground truth key points in all the sample sets and the true value manually marked. v_{pi} represents the visibility of the i key points of the pedestrian with id p in the ground truth.

$\delta(*)$ means if the condition $*$ holds, then $\delta(*) = 1$; otherwise, $\delta(*) = 0$.

AP (average precision) is used to calculate the accuracy percentage of the test set. In single-person pose estimation, only one pedestrian is estimated at a time, that is, $M = 1$ in the *oks* indicator, so the ground truth in a picture is a pedestrian (GT), and a set of key points will be obtained after the key point detection of this pedestrian (DT), and finally calculate the similarity *oks* between GT and DT as a scalar, and then artificially give a threshold T, and then AP can be calculated from the *oks* of all pictures:

$$AP = \frac{\sum_p \delta\left(oks_p > T\right)}{\sum_p 1}. \tag{10}$$

PCP is the Percentage of Correct Parts. If the key distance between the positions of the two joint points and the real limb reaches at most half the length of the real limb, the joint point is considered to be correctly predicted.

4.2. Experimental Results of Different Methods. In order to verify the competitiveness of MVBD-NET in the task of human skeleton action recognition, the model trained on the

Table 1: PCP comparison on Leeds Sports Pose dataset using oks evaluation method.

Methods	Arm		Leg		Torso	Head
	Upper	Lower	Upper	Lower		
Dantone	0.53	0.35	0.74	0.71	0.82	0.78
Tian	0.45	0.38	0.52	0.69	0.81	0.65
Johnson	0.52	0.31	0.64	0.52	0.71	0.74
Wang	0.54	0.45	0.74	0.65	0.76	0.80
Pishchulin	0.43	0.55	0.68	0.71	0.89	0.81
Ours	**0.86**	**0.70**	**0.89**	**0.86**	**0.92**	**0.88**

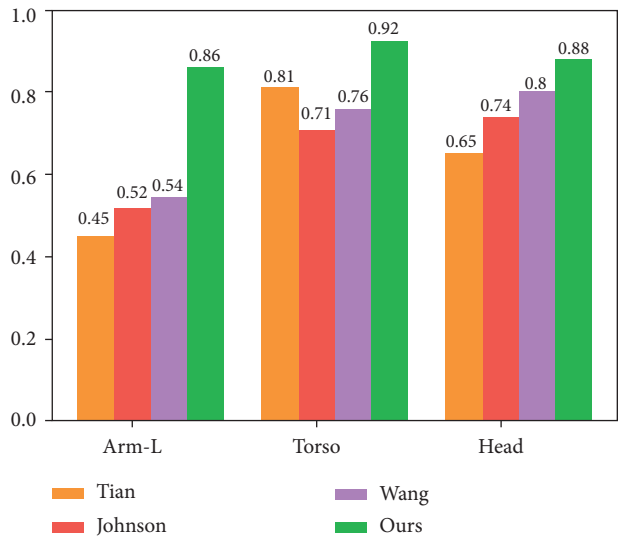


Figure 5: PCP comparison on LSP.

graph convolutional network based on the spatiotemporal attention mechanism is compared with the current state of the art (SOA). The methods were compared experimentally. Compare the proposed MVBD-Net model with Dantone [24], Tian [25], Johnson [26], Wang [27], and Pishchulin [28]. The results of the different datasets are shown in the following Table 1 and Figure 5.

The experiment is mainly to verify the effectiveness of the time attention mechanism on the action recognition data set. This set of experiments designed two different convolution models based on time attention mechanism. The models are all based on the ST-GCN benchmark model with time attention mechanism. The temporal attention mechanism convolution model uses different video time-domain segmentation and selection methods to verify the influence of different time-domain segmentation methods and different segmentation parameters on the accuracy of behavior recognition. The training process of the benchmark model in this set of experiments is the same as the previous set of experiments. The experimental results are shown in Table 1. Among them, the average fusion mode (AFM) is to directly average the prediction results of each time-domain segment as the prediction result of the entire video; and the discrimination weighted (Discriminative Confidence Weighted, DCW) represents a time-domain fusion method based on weighted prediction credibility; that is, the most reliable video segment is selected according to the prediction credibility of each input segment, and the prediction results of these segments are weighted as the prediction result of the entire action.

4.3. Experimental Results of Ablation Studies on Spatiotemporal Attention Graph Convolutional Network. In this section, five sets of experiments are designed to verify in detail the effectiveness of the graph convolutional neural network based on the spatiotemporal attention mechanism in the task of skeletal action recognition. The first set of experiments shows the visualization results of the attention heat map of the corresponding frame learned by the attention model network [29, 30]. The second and third sets of experiments are slice experiments, which are used to independently verify the effectiveness of the spatial attention mechanism and the temporal attention mechanism. The fourth set of experiments is to integrate spatial attention mechanism and temporal attention mechanism into a graph convolutional neural network to achieve end-to-end training and apply it to action recognition tasks. This experiment is to verify the effect of introducing spatiotemporal attention mechanism into the graph convolutional neural network on the task of action recognition.

The experimental results are shown in Table 2. MVBD is a spatiotemporal attention network model established on the

Table 2: Average precision of joint detection on LSP.

LSP	Ankle	Knee	Hip	Wrist	Elbow	Shoulder	Neck	Head	mAP
P_{TF}	0.25	0.34	0.23	0.23	0.43	0.32	0.21	0.34	0.31
P_{HSV}	0.37	0.31	0.36	0.41	0.43	0.33	0.32	0.76	0.36
MVBD-net	**0.47**	**0.49**	**0.39**	**0.38**	**051**	**0.52**	**0.61**	**0.88**	**0.69**

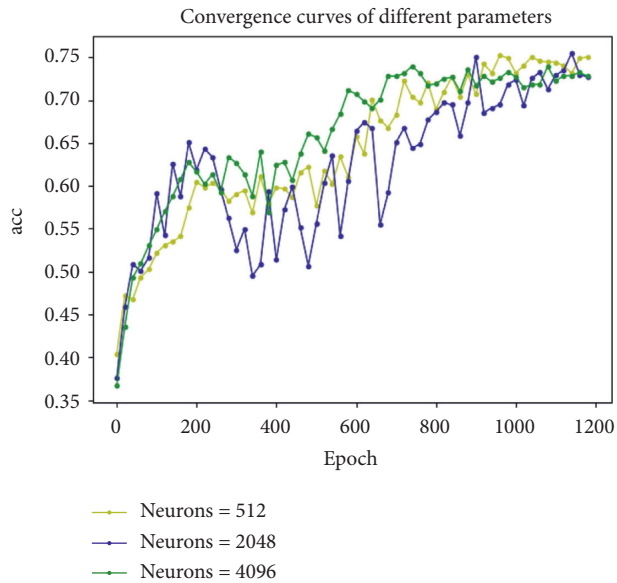


Figure 6: Convergence curves of different parameters.

basis of a recurrent neural network with long- and short-term memory. It can selectively focus on the joint differences of input frames and give different degrees of attention to the output of different frames, so it can extract distinguishable. The spatiotemporal features help action recognition. ST-GCN [21] breaks through the limitations of previous bone modeling methods and applies graph convolution to human skeleton action recognition, and the proposed model has strong generalization ability. AS-GCN [22] combines A-links and S-links into a generalized pose graph, further establishes a behavior structure graph convolutional network model, learns spatial and temporal characteristics, and can capture different action patterns more accurately and in detail.

4.4. Experimental Results of Ablation Studies on Different Parameters. Considering that there are a large number of parameters that can be optimized in the network structure designed in this paper, the use of different parameter settings will have different effects on the accuracy and operating efficiency of the model, so this paper conducts ablation experiment analysis on different parameter configurations. Since the three-dimensional pose estimation in this paper is implemented using a fully connected network, and the number of neurons in the fully connected layer is different, the number of model parameters and the prediction effect are also different. So in Figure 6, a different number of neurons (that is, Linear_size, representing fully connected layers) are analyzed. By comparing the loss function (loss) curve of the number of neurons, it is found that the loss value shows a gradually decreasing trend as the number of neurons increases from 256 to 4,096, which indicates the accuracy of the number of neurons. There was a positive correlation between model training and the number of neurons. The positive correlation indicates that the convergence of the model training will be higher with the increase of the number of neurons. As shown in Figure 7, this paper randomly selects 50 batch test data sets from the test set of the data set Human3.6 M to verify the prediction effect of the model. The batch-size is 64, and a total of 3 200 test data sets are used in this experiment. The abscissa in the figure represents the number of neurons in each layer, the parameters are set to 512, 1024, 2 048, and 4 096, and the ordinate represents the average value (mm) of node errors in the calculation of the batch of test data. By comparing the number of neurons with different numbers, the prediction result found that the error value of the number of neurons of 4 096 is significantly lower than that of the number of neurons of 2 048. Therefore, this article sets the number of neurons in the fully connected layer to 4096. The convergence curves of different parameters are shown in Figure 6.

4.5. Experimental Results of Ablation Studies on Different Submodule. Since the MVBD-NET algorithm uses the Yolo algorithm to extract the LSP human body pose and uses big data technology to do auxiliary analysis; in this section, we

Figure 7: The visualization results of athlete training action recognition on the LSP dataset.

Table 3: Average precision of different submodule on LSP.

LSP	Ankle	Knee	Hip	Wrist	Elbow	Shoulder	Neck	Head	mAP
Yolo-submodule	0.26	0.39	0.22	0.27	0.41	0.33	0.28	0.36	0.33
BigData-Submodule	0.41	0.48	0.39	0.32	0.49	0.46	0.47	0.72	0.56
MVBD-net	**0.47**	**0.49**	**0.39**	**0.38**	**051**	**0.52**	**0.61**	**0.88**	**0.69**

will analyze the impact of these strategies on the experiment. The results are shown in Table 3.

4.6. Visualization of Results. This section shows the visualization results of the MVBD-NET algorithm on the LSP data set. Due to the introduction of an effective spatiotemporal attention model and training strategy, which can extract discriminative spatiotemporal features, the MVBD-Net proposed in this paper has obtained the best classification accuracy among current similar methods on this data set. Figure 7 visually shows the results of the algorithm.

5. Conclusion

Nowadays, the athlete action recognition has become a significant research area for showing and recognition of athlete actions. Movement recognition of athletes can be accomplished through a variety of modes, such as motion sensors, big data analysis, and machine vision. This paper proposes a graph convolutional neural athlete motion recognition algorithm based on big data analysis and machine vision, which is used for athlete's motion gesture recognition and intervention. Among them, this article first uses the Yolo model to extract the athlete's action posture features and uses the spatiotemporal attention map convolutional neural network to estimate the posture. Secondly, we analyzed the historical training and competition data through big data analysis technology, obtained the posture weight, and performed feature fusion with the output of the graph convolutional neural network. Experimental results show that our MVBD-Net algorithm is effective. It can formulate reasonable training plans for coaches and athletes of sports teams, provide scientific basis, and improve training efficiency and training effects. The results achieved from the proposed study show the effectiveness of the study.

Data Availability

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

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

[1] D. C. Luvizon, D. Picard, and H. Tabia, "2d/3d pose estimation and action recognition using multitask deep learning," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 5137–5146, Salt Lake City, UT, USA, June 2018.

[2] M. Liu and J. Yuan, "Recognizing human actions as the evolution of pose estimation maps," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1159–1168, Salt Lake City, UT, USA, June 2018.

[3] H. H. Pham, H. Salmane, L. Khoudour, A. Crouzil, S. A. Velastin, and P. Zegers, "A unified deep framework for joint 3d pose estimation and action recognition from a single rgb camera," *Sensors*, vol. 20, no. 7, p. 1825, 2020.

[4] G. Yang, L. Wang, X. Xu, and J. Xia, "Footballer action tracking and intervention using deep learning algorithm," *Journal of Healthcare Engineering*, vol. 2021, 2021.

[5] A. Toshev and C. Szegedy, "Deeppose: human pose estimation via deep neural networks," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1653–1660, Ohio, Columbus, June 2014.

[6] J. Tompson, A. Jain, Y. LeCun, and C. Bregler, "Joint training of a convolutional network and a graphical model for human pose estimation," 2014, https://arxiv.org/abs/1406.2984.

[7] W. Yang, W. Ouyang, H. Li, and X. Wang, "End-to-end learning of deformable mixture of parts and deep convolutional neural networks for human pose estimation," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 3073–3082, Las Vegas, NV, USA, June 2016.

[8] J. Song, L. Wang, L. Van Gool, and O. Hilliges, "Thin-slicing network: a deep structured model for pose estimation in videos," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 4220–4229, Honolulu, HI, USA, July 2017.

[9] M. R. I. Hossain and J. J. Little, "Exploiting temporal information for 3d human pose estimation," in *Proceedings of the European Conference on Computer Vision (ECCV)*, pp. 68–84, Munich, Germany, September 2018.

[10] D. Pavllo, C. Feichtenhofer, D. Grangier, and M. Auli, "3d human pose estimation in video with temporal convolutions and semi-supervised training," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 7753–7762, Long Beach, CA, USA, June 2019.

[11] Z. Cao, T. Simon, S. E. Wei, and Y. Sheikh, "Realtime multi-person 2d pose estimation using part affinity fields," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 7291–7299, Honolulu, HI, USA, July 2017.

[12] Y. Chen, Z. Wang, Y. Peng, Z. Zhang, G. Yu, and J. Sun, "Cascaded pyramid network for multi-person pose estimation," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 7103–7112, Salt Lake City, UT, USA, June 2018.

[13] X. Ning, S. Xu, W. Li, and S. Nie, "FEGAN: flexible and efficient face editing with pre-trained generator," *IEEE Access*, vol. 8, pp. 65340–65350, 2020.

[14] W. Cai and Z. Wei, "PiiGAN: generative adversarial networks for pluralistic image inpainting," *IEEE Access*, vol. 8, pp. 48451–48463, 2020.

[15] X. Ning, P. Duan, W. Li, Y. Shi, and S. Li, "A CPU real-time face alignment for mobile platform," *IEEE Access*, vol. 8, pp. 8834–8843, 2020.

[16] W. Cai and Z. Wei, "Remote sensing image classification based on a cross-attention mechanism and graph convolution," *IEEE Geoscience and Remote Sensing Letters*, vol. 7, 2020.

[17] X. Ning, F. Nan, S. Xu, L. Yu, and L. Zhang, "Multi-view frontal face image generation: a survey," *Concurrency and Computation: Practice and Experience*, Article ID e6147, 2020, in Press.

[18] W. Cai, B. Liu, Z. Wei, M. Li, and J. Kan, "TARDB-Net: triple-attention guided residual dense and BiLSTM networks for hyperspectral image classification," *Multimedia Tools and Applications*, vol. 80, pp. 1–22, 2021.

[19] X. Ning, W. Li, and J. Xu, "The principle of homology continuity and geometrical covering learning for pattern recognition," *International Journal of Pattern Recognition and Artificial Intelligence*, vol. 32, no. 12, Article ID 1850042, 2018.

[20] Z. Wang, C. Zou, and W. Cai, "Small sample classification of hyperspectral remote sensing images based on sequential joint deeping learning model," *IEEE Access*, vol. 8, pp. 71353–71363, 2020.

[21] X. Zhang, Y. Yang, Z. Li, X. Ning, Y. Qin, and W. Cai, "An improved encoder-decoder network based on strip pool method applied to segmentation of farmland vacancy field," *Entropy*, vol. 23, no. 4, p. 435, 2021.

[22] Z. L. Yang, S. Y. Zhang, Y. T. Hu, Z. W. Hu, and Y. F. Huang, "VAE-Stega: linguistic steganography based on variational auto-encoder," *IEEE Transactions on Information Forensics and Security*, vol. 16, pp. 880–895, 2020.

[23] X. Ning, X. Wang, S Xu et al., "A review of research on co-training," *Concurrency and Computation: Practice and Experience*, vol. 24, 2021.

[24] Y. Kim and D. Kim, "A CNN-based 3D human pose estimation based on projection of depth and ridge data," *Pattern Recognition*, vol. 106, Article ID 107462, 2020.

[25] M. Dantone, J. Gall, C. Leistner, and L. Van Gool, "Human pose estimation using body parts dependent joint regressors," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 3041–3048, Portland, OR, USA, June 2013.

[26] Y. Tian, C. L. Zitnick, and S. G. Narasimhan, "Exploring the spatial hierarchy of mixture models for human pose estimation," in *Proceedings of the European Conference on Computer Vision*, pp. 256–269, Springer, Berlin, Heidelberg, October 2012.

[27] S. Johnson and M. Everingham, "Learning effective human pose estimation from inaccurate annotation," in *Proceedings of the IEEE Conference*, pp. 1465–1472, IEEE, Shanghai, China, June 2011.

[28] L. Pishchulin, M. Andriluka, P. Gehler, and B. Schiele, "Poselet conditioned pictorial structures," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 588–595, Portland, OR, USA, June 2013.

[29] X. Yu, F. Jiang, J. Du, and D. Gong, "A cross-domain collaborative filtering algorithm with expanding user and item

features via the latent factor space of auxiliary domains," *Pattern Recognition*, vol. 94, pp. 96–109, 2019.

[30] X. Yu, Y. Chu, F. Jiang, Y. Guo, and D. Gong, "SVMs classification based two-side cross domain collaborative filtering by inferring intrinsic user and item features," *Knowledge-Based Systems*, vol. 141, pp. 80–91, 2018.

Hindawi
Scientific Programming
Volume 2021, Article ID 7488025, 7 pages
https://doi.org/10.1155/2021/7488025

Research Article

Efficient Segmentation of Lymphoblast in Acute Lymphocytic Leukemia

Syed Ijaz Ur Rahman,[1] **Misbah Jadoon**,[2] **Sikandar Ali**,[3,4] **Hizbullah Khattak**,[5] **and Jiwei Huang**[3,4]

[1]*Department of Computer Science, Islamia College University Peshawar, Peshawar, Pakistan*
[2]*Department of Computer Science, Women University Swabi, Khyber Pakhtunkhwa, Swabi 23430, Pakistan*
[3]*Department of Computer Science and Technology, China University of Petroleum, Beijing 102249, China*
[4]*Beijing Key Lab of Petroleum Data Mining, China University of Petroleum, Beijing 102249, China*
[5]*Department of Information Technology, Hazara University Mansehra, Khyber Pakhtunkhwa, Pakistan*

Correspondence should be addressed to Sikandar Ali; sikandar@cup.edu.cn and Jiwei Huang; huangjw@cup.edu.cn

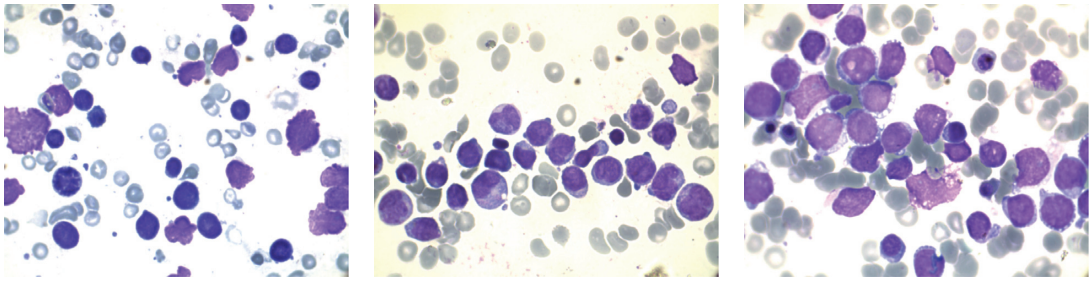
Received 6 April 2021; Accepted 22 April 2021; Published 3 May 2021

Academic Editor: Shah Nazir

Microscopic examination of peripheral blood smears and bone marrow is the preliminary step for the diagnosis of several life-threatening diseases. Acute lymphocytic leukemia (ALL) is the most common disease in children that also needs an early diagnosis for on-time treatment as it spreads rapidly in the blood and forms immature lymphocytes. This might cause death in some weeks if left untreated. Manual methods in clinical laboratory being applied for the diagnosis of these diseases are inefficient and expensive, and the results are less accurate. A computer-aided system is the need of the day in which the most important step is segmenting the region of interest in blood or bone marrow for the detection and cure of the diseases which is the most challenging task. This study aims to propose a simple threshold-based segmentation technique by processing the S component of the HSV color space to segment the lymphoblasts in the bone marrow images of ALL patients. The technique was applied to 230 RGB bone marrow images having all the three types of ALL, i.e., L1, L2, and L3 resulted in the overall accuracy of 96.8%.

1. Introduction

Blood and its components are the key substances for the microscopic examination of most life threat diseases such as leukemia, HIV, anemia, tumor, cancer, and thalassemia [1–5]. Leukemia is the type of cancer starting in bone marrow from immature lymphocytes called the lymphoblasts and then spreads through the blood in all over the body. The on-time treatment has a good chance of recovery for the patient. Lymphoblasts are mainly present in the bone marrow or peripheral blood slides of a patient having acute lymphoblastic leukemia [6–8]. According to French American British (FAB) classification, ALL is subclassified as L1, L2, and L3 on the basis of morphological structures exhibited in Figure 1. L1 type consists of nucleus and may or may not have a basophilic cytoplasm and have a small size, L2 having irregular nucleus and having a cytoplasm. L3 is large in size and have vacuoles in the cytoplasm. To diagnose ALL, it is totally dependent on the detection and morphological examination of blast cells in bone marrow and peripheral blood images. This process is performed by the hematologists' manually in clinical laboratories, which is a slow process, and it also depends on the experience of the laboratory expert and his physical condition. Hence, to overcome the limitations of these manual processes, fast and automated system is the need for the diagnosis of these lives in danger diseases. For the designing of this automated system, the most essential part is the segmentation of the region of interest, and it is the most difficult task to be done as the whole system depends on it. The aim of this study is to segment the blasts from bone marrow images and to introduce an effective segmentation approach. This can



(a) (b) (c)

Figure 1: Three types of lymphoblast in acute lymphoblastic leukemia. (a) L1 blasts. (b) L2 blasts. (c) L3 blasts.

help the experts in the designing of the whole diagnostic system.

The further paper is organized into four main sections: Section 2 explores related work, Section 3 presents materials and methods reported in the paper, and Section 4 exhibits experimental results and discussion. Finally, Section 5 presents the conclusion.

2. Related Studies

There is no simple technique for the segmentation process as most of the researchers addressed this problem and reported their techniques in state of the art [5, 9]. Sadeghian et al. [10] proposed a segmentation method for ALL blast in peripheral blood images for which canny edge detection algorithm was used with gradient vector flow GVF for the detection of the nucleus and then zack threshold was applied for cytoplasm. However, this method is applied to a single WBC image. Osowski et al. [11] recognized myoblasts in bone marrow by segmenting it and applied watershed transformation algorithm, but due to the crowded marrow, it resulted in oversegmentation. Madhloom et al. [12] proposed a method in which they used contrast stretching and arithmetic for the localization and segmenting white blood cells with 95% accuracy; however, the method was used for the nucleus component, and that's why this is not identifying the leukemia cells as it required nucleus as well as the cytoplasm. In this study, we have proposed a method based on thresholding with the HSV color model, which could accurately segment the lymphoblasts for the detection of ALL in microscopic bone marrow images.

Tabrizi et al. [13] proposed snake contour for the segmentation of leukocytes; LVO, PCA algorithms, and SVM classifier is used for the identification. Mohamed and Far [14] segmentation was intensity based on Gram–Schmidt orthogonalization. Abbas and Mohamad [8] proposed a convolution mask along with Otsu method for the segmentation of nuclei of the lymphocytes for the detection of leukemia having the same limitation as faced by Madloom et al. [12] to detect leukemia nucleus, cytoplasm is needed. Mohapatra et al. [15] separated the interested region using K-means clustering and then used shadowed C-means (CSM) on the CIELAB color model for the segmentation of nuclei and cytoplasm of the leukocytes and then used an ensemble of classifiers for the classification. Adollah et al. [16] reviewed different methods for the segmentation of leukocytes in which different authors addressed the problem with different techniques, i.e., Otsu threshold and circular histogram were used for the segmentation of WBCs, entropy with higher order is used as a feature, gray level threshold, filters with color matching, different morphological operators, shape information of binarizing and generating maximum intensity, watershed with seeds, and combining filters with scale-space were used for the segmentation purpose.

Comaniciu and Meer [17] recommend an image-guided decision support (IGDS) for the classification of different lymphomas and uses mean shift algorithm for the accurate segmentation; however, the system is not checked for acute leukemia. Abbas and Dzulkifli Mohamad [8] applied $2*2/6$ convolution mask for repressing high values of RGB, and then Otsu method is applied to obtain nuclei, small areas were removed, and nuclei have been dilated for the required results of segmenting nuclei and detecting leukemia resulting in accuracy of 96.5%.

Bhattacharjee and Saini [18] used a watershed algorithm for the segmentation of blast cells to detect ALL, and the segmentation process was followed with morphological operators, binary search tree, and Gaussian mixture model to classify the blasts. Pan et al. [19] employed extreme learning machine to segment the leukocytes on the basis of gradient threshold and entropy for the segmentation of multicolor objects on the HIS color model with Otsu's thresholding to properly segment leukocytes. Abbas et al. [6] applied a convolution filter of 3×3 on the red channel of the RGB image and then converted that image into the binary mask and on the analysis of the histogram, and after that, both the binary images were added which results in nuclei of leukocytes. Amin et al. [20] used K-means clustering to segment the lymphoblasts cells from the blood images. Geometric and statistical features were employed to classify that lymphoblast into subtypes of ALL using support vector machine (SVM).

Dhanachandra et al. [21] enhanced the contrast of the images for which the author used partial contrast stretching, then used subtractive clustering method followed by K-means clustering and calculating potential for each pixel, then finds the center cluster using Euclidean distance, and then reshapes the image. For noise removal, they use a median filter. Rawat et al. [22] proposed histogram equalization followed by global threshold and morphological

opening to segment the nucleus of the blast cell and finally subtracted an image from the preprocessed image to obtain cytoplasm. To classify the images, geometrical, statistical, and chromatic features were extracted and fed to classifiers PCA-KNN, PCA-PNN, PCA-SVM, and PCA-ANFIS in hierarchical; this took a long time to classify the images. Nighat Bibi et al. [23] used dense convolution neural network and residual CNN with cloud framework for the detection and classification of all types of leukemia.

3. Materials and Methods

The main purpose of the segmentation process is to collect the lymphoblasts from the slides containing all the blood components. It will help in the detection and classification of acute leukemias [7]. The original images obtained from digital camera microscope are in the RGB color model. However, for speedy and accurate segmentation, it is converted to the HSV color model to process the S component.

3.1. Image Database. Images for this study were taken in the department of hematology lab at Saidu Medical College, Swat, KPK, Pakistan, using Euromax digital camera microscope. The slides were stained with Giemsa under the supervision of Dr. Amreek Lal, and the images were taken with 100x lens and oil immersion. The total images of $L1$ type 102, $L2$ type 120, $L3$ type 30 were kept in different folders classified according to (FAB) with the help of concerned hematologist.

3.2. Proposed Methodology. A simple method for the segmentation of lymphoblasts is proposed. The images were first converted to the HSV color model, and then the S component is processed further to achieve good segmentation results. There are the subtypes of acute lymphoblastic leukemia according to the French American British classification, i.e., $L1$, $L2$, and $L3$ having different morphologies [20]. So, we tested different threshold values for each class according to the morphological structures of the lymphoblast. In the S component image, we found the maximum threshold value to find out the value (T_{rh} value) in which the blast could only segment, while the rest of the objects should be removed.

$$\mathrm{TT} = \mathrm{Max}(\mathrm{Max}(S\ (i,\ j))) - \mathrm{T}_{rh}. \quad (1)$$

Different values of T_{rh} were checked with equation (1) for all classes and then take that value on which the lymphoblasts were segmented accurately. After this, a MATLAB command Imfill holes were applied to all the images to restore the lost information and then find out a seeded value of RGB color of the blast in the original input image to convert back the image into its original RGB color. The proposed research framework is exhibited in Figure 2.

3.3. Pseudocode. The proposed algorithm processed the image and segmented lymphoblast according to the following pseudocode:

(i) Read the image img(i, j) from the directory in RGB color space

(ii) Convert RGB image img(i, j)to HSV color space img_seg(i, j) for further processing

(iii) Get the S component of the HSV color model, $S(i, j)$ as they have more information about blasts and segmentation occur easily

(iv) Find out the maximum threshold in $S(i, j)$ and then subtract that value on which other small objects remove and only the lymphoblasts remain

(v) If $S(i, j)$ is greater or equal to the threshold value, find out in the previous step and then give the value 1, otherwise 0

(vi) For the lost information function, Imfill holes are used to fill the gaps and information

(vii) Wherever the value is 1, convert to its original color in RGB img(i, j)

4. Experimental Results

Different T_{rh} values were checked for the accurate segmentation in different experiments. Firstly, the value for L1 blasts was kept 0.40 but could not produce efficient results and then changed to 0.45 in which it is segmented accurately over 102 images. It has an irregular shape nucleus and in rare cases have a cytoplasm. For L2 blast, due to the morphology of having cytoplasm, we checked the value 0.47 and 0.48, but the results were not satisfactory. On $\mathrm{T_{rh}}$ value 0.50, it produced good segmentation results. This was applied to 120 images having L2 morphology. L3 blasts are large in size and having vacuoles in the cytoplasm part of the cells for that the value was set higher than that of L2 0.52 to segment the blasts with vacuoles to turn out satisfactory results. After hole filling is applied to fill the gaps which occurred during the whole process, results are exhibited in Figures 3–5 for ALL subtypes L1, L2, and L3, respectively.

The segmentation results are shown in Figures 3–5 for lymphoblasts of types L1, L2, and L3, respectively, giving efficient results as compared to other methods. Performance accuracy of the technique used in our study is shown in Figure 6.

4.1. Performance Analysis. The proposed technique was also compared in terms of segmentation accuracy for the evaluation with three different existing methods which is shown in Figure 7. It is used for the segmentation of five types of white blood cells, and those methods were proposed by Madloom et al. [12]; Tabrizi et al. [13]; and Mohamed and far [14] presented in Figure 7. The only limitation in this study is that sometimes it results in oversegmentation and sometimes the very important information is missing especially in the case of L3 subtype of the ALL which affects the overall accuracy. Further segmentation results of all the three types of lymphoblasts are shown in Figure 8. Finally Table 1 presents accuracy for each subtype. The algorithms were also checked with ground-truth data, and the result is shown in Table 2.

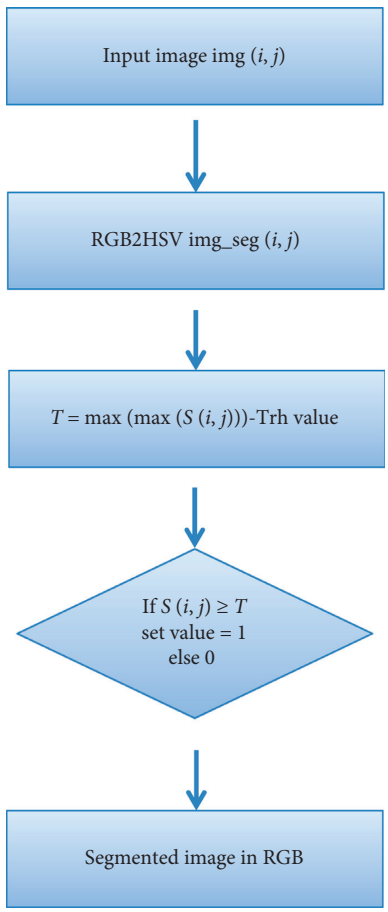
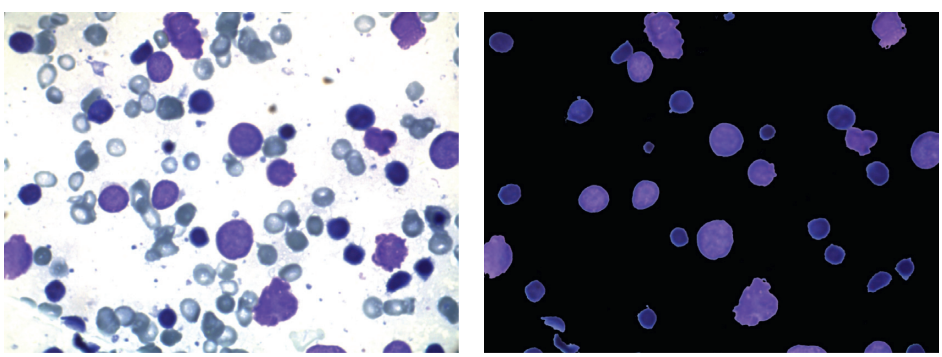
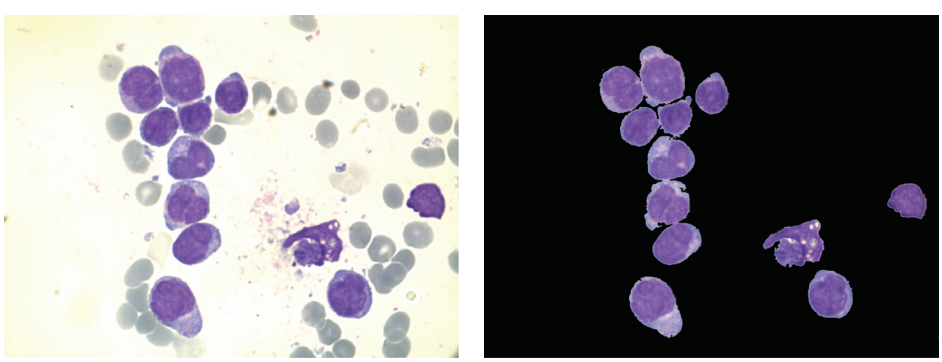


Figure 2: Proposed research framework.



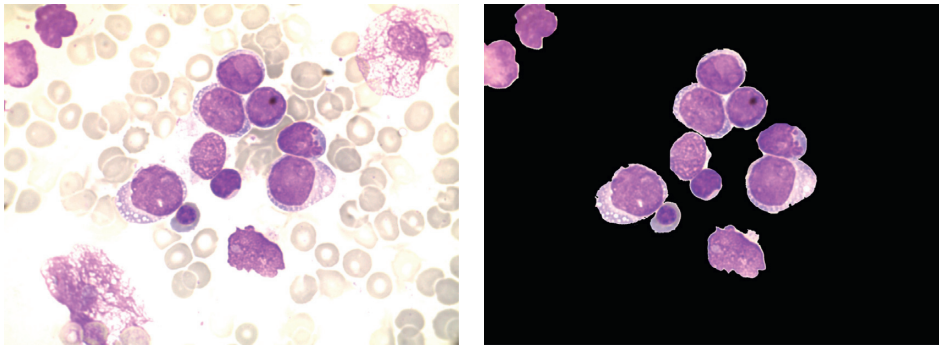
(a) (b)

Figure 3: ALL subtype L1. (a) Original image. (b) Segmented lymphoblast.



(a) (b)

Figure 4: ALL subtype L2. (a) Original image. (b) Segmented lymphoblast.



(a) (b)

Figure 5: ALL subtype L3. (a) Original image. (b) L3 segmented lymphoblast.

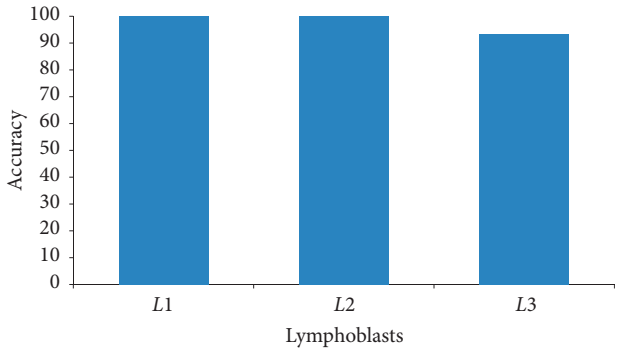


Figure 6: Performance accuracy of the proposed system.

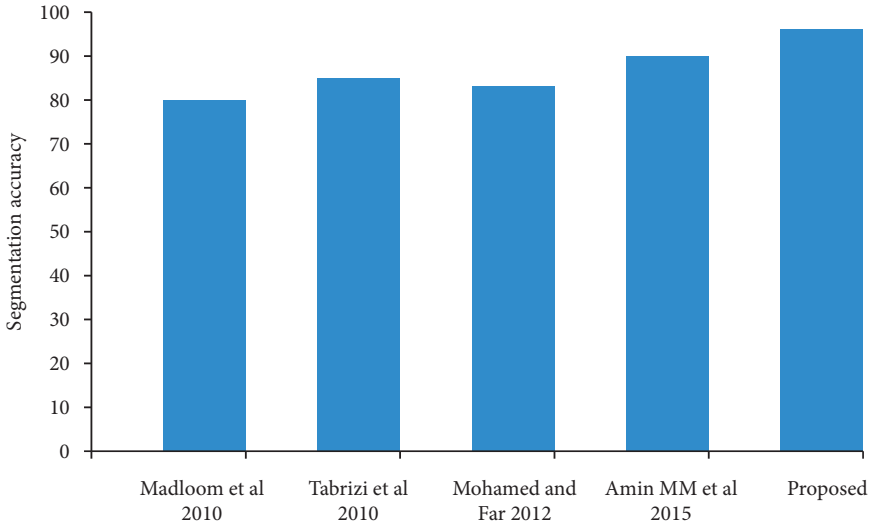
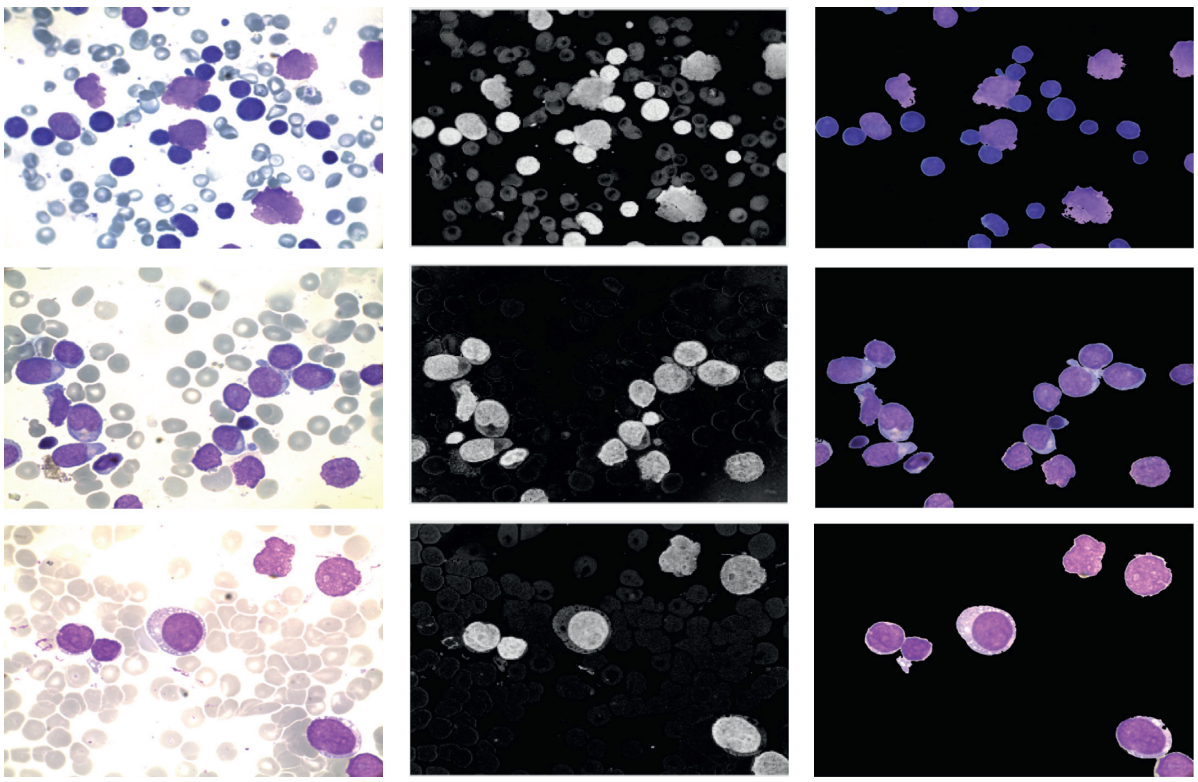


Figure 7: Performance evaluation of the proposed method with different techniques in state of art.



(a) (b) (c)

Figure 8: Experimental results. (a) Original images. (b) S component of HSV color space. (c) Segmented images.

Table. 1: Accuracy for each subtype.

ALL subtype	$L1$	$L2$	$L3$
Total images	102	120	30
Accurate segmented	102/102	120/120	28/30
Accuracy	100%	100%	93.3%

Table 2: Proposed method evaluation with ground-truth data.

Lymphoblast type	Recall	Precision	F-measure
L1	0.9622	0.9811	0.9703211
L2	0.9501	0.9712	0.96010211
L3	0.9211	0.9322	0.9313221

5. Conclusion

This paper has presented an effective approach to segment lymphoblast in a robust and accurate manner for the detection of leukemia by using simple threshold-based method with the HSV color model and attained an overall accuracy of 96.8%. The proposed segmentation technique is easy to implement and then applied to 230 RGB images of bone marrow slides images of ALL patients. The segmentation accuracy for the subtypes L1 and L2 was 100%; for the third subtype L3, the accuracy was 93.3% because of the vacuoles present in the cytoplasm area. Furthermore, the method is compared for performance evaluation with three other existing methods in terms of accuracy. We can also use the proposed method in general for the segmentation of blast cells in other types of leukemia. This is witnessed from our experimental results that the proposed method in this paper is more helpful in the segmentation of lymphoblasts specially for ALL. For future work, segmentation accuracy for L3 subtype of ALL could be improved and limitations will be covered.

Data Availability

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

Additional Points

Summary. Acute lymphocytic leukemia (ALL) is a crucial disease, and manual methods are inefficient and expensive. Automated system is proposed to identify ROI in blood/marrow for the detection and cure of the disease by processing S component of HSV of lymphoblast.

Conflicts of Interest

The authors declare no conflicts of interest.

Acknowledgments

This study was supported by the National Key Research and Development Plan (no. 2016YFC0303700), National Natural Science Foundation of China (no. 61972414), Beijing Natural Science Foundation (no. 4202066), Beijing Nova Program

(no. Z201100006820082), and Fundamental Research Funds for Central Universities (nos. 2462020YJRC001 and 2462018YJRC040).

References

[1] S. Iqbal, M. U. Ghani, T. Saba, and A. Rehman, "Brain tumor segmentation in multi-spectral MRI using convolutional neural networks (CNN)," *Microscopy Research and Technique*, vol. 81, no. 4, pp. 419–427, 2018.

[2] S. Iqbal, M. Usman Ghani Khan, T. Saba, and A. Rehman, "Computer-assisted brain tumor type discrimination using magnetic resonance imaging features," *Biomedical Engineering Letters*, vol. 8, no. 1, pp. 5–28, 2017.

[3] A. Jamal, M. H. Alkawaz, A. Rehman, and T. Saba, "Retinal imaging analysis based on vessel detection," *Microscopy Research and Technique*, vol. 80, no. 7, pp. 799–811, 2017.

[4] B. Mughal, M. Sharif, N. Muhammad, and T. Saba, "A novel classification scheme to decline the mortality rate among women due to breast tumor," *Microscopy Research and Technique*, vol. 23, 2017.

[5] T. Saba, S. T. F. Bokhari, M. Sharif, M. Yasmin, and M. Raza, "Fundus image classification methods for the detection of glaucoma: a review," *Microscopy Research and Technique*, vol. 18, 2018.

[6] A. N. Mohamad, A. H. Abdullah, T. Al-Rodhaan, and M. Al-Dhelaan, "Nuclei segmentation of leukocytes in blood smear digital images," *Pakistan Journal of Pharmaceutical Sciences*, vol. 28, no. 5, pp. 1801–1806, 2015.

[7] A. Rehman, N. Abbas, T. Saba, S. Mehmood, and Z. Kolivand, "Classification of acute lymphoblastic leukemia using deep learning," *Microscopy Research & Technique*, vol. 81, 2018.

[8] N. Abbas and D. Mohamad, "Automatic color nuclei segmentation of leukocytes for acute leukemia," *Research Journal of Applied Sciences, Engineering and Technology*, vol. 7, no. 14, pp. 2987–2993, 2014.

[9] A. Norouzi, M. S. M. Rahim, A. Altameem, T. Saba, A. E. Rada, and A. Rehman, "Medical image segmentation methods, algorithms," *Applications IETE Technical Review*, vol. 31, no. 3, pp. 199–213, 2014.

[10] F. Sadeghian, Z. Seman, A. R. Ramli, B. H. A. Kahar, and M. I. Saripan, "A framework for white blood cell segmentation in microscopic blood images using digital image processing," *Biological Procedures Online*, vol. 11, no. 1, p. 196, 2009.

[11] S. Osowski, T. Markiewicz, B. Marianska, and L. Moszczyński, "September. Feature generation for the cell image recognition of myelogenous leukemia," in *Procedings of the Signal Processing Conference, 2004 12th European*, IEEE, Vienna, Austria, September 2004.

[12] H. T. Madhloom, S. A. Kareem, H. Ariffin, A. A. Zaidan, H. O. Alanazi, and B. B. Zaidan, "An automated white blood cell nucleus localization and segmentation using image arithmetic and automatic threshold," *Journal of Applied Sciences*, vol. 10, no. 11, pp. 959–966, 2010.

[13] P. Tabrizi, S. Rezatofighi, and M. Yazdanpanah, "Using PCA and LVQ neural network for automatic recognition of five types of white blood cells," in *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC*, Berlin, Germany, August 2010.

[14] M. Mohamed and B. Far, "An enhanced threshold based technique for white blood cells nuclei automatic segmentation," in *Proceeding of the IEEE 14th International Conference on E-Health Networking*, pp. 202–207, Munich, Germany, September 2012.

[15] S. Mohapatra, D. Patra, and S. Satpathy, "An ensemble classifier system for early diagnosis of acute lymphoblastic leukemia in blood microscopic images," *Neural Computing and Applications*, vol. 24, no. 7-8, pp. 1887–1904, 2014.

[16] R. Adollah, M. Y. Mashor, N. M. Nasir, H. Rosline, H. Mahsin, and H. Adilah, "Blood cell image segmentation: a review," in *Proceedings of the 4th Kuala Lumpur International Conference on Biomedical Engineering 2008*, pp. 141–144, Kuala Lumpur, Malaysia, June 2008.

[17] D. Comaniciu and P. Meer, "Cell image segmentation for diagnostic pathology," in *Advanced Algorithm Approaches to Medical Image Segmentation: State-of-The-Art Application in Cardiology*Springer, Berlin, Germany, 2001.

[18] R. Bhattacharjee and L. M. Saini, "Detection of Acute Lymphoblastic Leukemia using watershed transformation technique," in *Proceedings of the Signal Processing, Computing and Control (ISPCC), 2015 International Conference on 2015*, pp. 383–386, Solan, India, September 2015.

[19] C. Pan, D. S. Park, Y. Yang, and H. M. Yoo, "Leukocyte image segmentation by visual attention and extreme learning machine," *Neural Computing and Applications*, vol. 21, no. 6, pp. 1217–1227, 2012.

[20] M. M. Amin, S. Kermani, A. Talebi, and M. G. Oghli, "Recognition of acute lymphoblastic leukemia cells in microscopic images using k-means clustering and support vector machine classifier," *Journal of Medical Signals and Sensors*, vol. 5, no. 1, pp. 49–58, 2015.

[21] N. Dhanachandra, K. Manglem, and Y. J. Chanu, "Image segmentation using K -means clustering algorithm and subtractive clustering algorithm," *Procedia Computer Science*, vol. 54, pp. 764–771, 2015.

[22] J. Rawat, A. Singh, H. S. Bhadauria, J. Virmani, and J. S. Devgun, "Classification of acute lymphoblastic leukaemia using hybrid hierarchical classifiers," *Multimedia Tools and Applications*, vol. 76, no. 18, pp. 19057–19085, 2017.

[23] N. Bibi, M. Sikandar, I. Ud Din, A. Almogren, and S. Ali, "IoMT-based automated detection and classification of leukemia using deep learning," *Journal of Healthcare Engineering*, vol. 12, pp. 1–12, 2020.

Hindawi
Scientific Programming
Volume 2021, Article ID 9925165, 11 pages
https://doi.org/10.1155/2021/9925165

Research Article

A Video Production Method of Microclass Combined with MOOC

Qingbo Yan

Changshu Institute of Technology, Changshu 215500, China

Correspondence should be addressed to Qingbo Yan; tg2021@cslg.edu.cn

Received 11 March 2021; Revised 26 March 2021; Accepted 13 April 2021; Published 26 April 2021

Academic Editor: Shah Nazir

In recent years, video production has become one of the key challenges for researchers in the area. In the face of the poor video effects of earlier video production methods and approaches, this study has considered the MOOC-based microclass combined with the MOOC video production method for improving the video production effect. According to the idea of MOOC, based on the early design stage and shooting preparation stage of video development, the video development team prepares the text, manuscript, and course type according to the initial design stage of video development and formulates the recording process. In the video recording stage, all activities of classroom teaching are recorded through the multislot mode for enhancing the sense of scene of the video picture. In the later stage of video production, video editing and feature processing are implemented by means of film and television to improve the effect of video production. The video production method based on the Hidden Markov model is used to ensure the continuity of front and back shots and complete online open video production and release to the cloud. The experimental results of the study show that the application effect of this method is effective, and the average satisfaction is as high and measured as 91.36%.

1. Introduction

At present, driven by the concept of global open education resource sharing, massive open online courses (MOOC) are developing rapidly. As a new online teaching mode, it has attracted great attention of scholars at home and abroad. It pays attention to the careful design and development of teaching resources. In particular, the MOOC video course developed with a variety of design, thinking, and expression techniques subverts the traditional classroom teaching video, constructs a teaching form of face-to-face conversation between teachers and online learners, and shortens the spatial distance between teachers and distance learners [1]. From the traditional distance education to today's online open education, knowledge dissemination is inseparable from the carrier media. Compared with monotonous documents and audio, video, which has both image and sound, can more intuitively and vividly interpret the teaching content and promote the wave of teaching video construction at home and abroad. In recent years, with the rapid rise of video open courses in well-known foreign universities in the world, Ted, Khan College, and other institutions have launched their own video lectures and teaching videos, which are loved and sought after by many online learners. At the same time, online education platforms such as OpenCourseWare Alliance, Udacity, Coursera, and edX have also developed a large number of unique video courses. It redesigns the course itself, divides a complete course into several short videos ranging from 5 to 15 minutes according to the knowledge points, and designs them jointly by teachers and professional teams to make the traditional classroom into an online classroom for online learners, which once again sets off a new upsurge in the construction of online video courses abroad. Under this influence, China has also launched a large-scale construction of online video courses, with the emergence of online video courses based on excellent resource sharing courses and Video Open Courses of Chinese University [2, 3]. However, looking at the existing video course resources at home and abroad, it is not difficult to find that compared with the foreign MOOC, which is a learner-centered concept of video course resource development, there is a big gap in the performance and design thinking of online video course in China. The development of the video courses still remains in the classroom

record level of face-to-face teaching. From the perspective of video quality, the picture quality and sound effect of online video courses are uneven, and the duration of video courses is generally about 30–50 minutes, which is not conducive to the long-term visual attention and knowledge memory of online learners, and it is not suitable for learners to use mobile terminals to implement fragmented learning; from the perspective of the video expression form, the phenomenon of "big head picture" of the lecturer in the picture is relatively serious [4], the screen visual attention guidance is not enough, the design of information is not in place, the content of the course is repetitive and lengthy, and the main teacher's performance is not strong. The main reason is that we lack careful design of course content, image design of teachers, and visual packaging of screen information in the process of network video course development.

Microclass refers to a video and audio section of the targeted explanation on a knowledge point within five minutes by using multimedia technology according to the idea of teaching design. These knowledge points can be the explanation of teaching materials, the explanation of question types, the introduction of materials, and the induction of test points; they can also be the explanation and display of teaching experience and methods. Microlesson is a kind of video course designed for a certain subject, knowledge point, or teaching link. Because of this kind of targeted and single characteristic, the video production method of microlesson is very changeable. Only after the video producers and teachers fully communicate and understand the teaching needs, can they use the appropriate shooting method to produce effective teaching videos (of course, if teachers make videos themselves, there will not be such a problem). Generally speaking, the length of teaching video of microclasses is not more than 10 minutes, and there may be mass production demand for different topics, so "electronic green board" and "EFP multi-machine shooting" are very suitable for making teaching videos of microclasses. The operation of "electronic green board" is simple, and teachers can work independently. After all, it is the teacher who knows the students best. Only teachers know how to play in front of the camera, so that students can effectively absorb it. Therefore, the advantages of "electronic green board" can be shown here; "EFP multi-position shooting" is suitable for real scene shooting, such as experiments and sports courses, which cannot be replaced by virtual objects, and multicamera multiangle shooting, will not miss any detail, can be a perfect picture of specific knowledge, in line with the microclass's targeted, single characteristics. Of course, although the microclass is simple, it does not mean that it is a crude video. If we want to produce excellent microclass courses, "virtual studio" can also be an option, but we need to spend more time on scripts and subscenes.

As for the production of teaching video, the first important idea to establish is that the school is no better than the TV station, and it is far inferior to the professional TV station in shooting specialty, budget, production manpower, and other aspects. If the video producers blindly ask the school to provide assistance beyond the campus hardware or teachers' professional degree or the school tries to catch up with the professional equipment of the TV station, it will not produce good results. Therefore, recognizing each other's needs and understanding each other's differences is a key point for film and television producers and school units to make teaching videos. It is also a basic concept for professional film and television circles and education circles to run in with each other. With this level of cognition, it will not simply pursue the level of hardware and equipment, but fail to produce teaching videos corresponding to school teaching needs. Compared with the traditional video, the traditional video class mainly depends on the teacher's explanation, which is difficult to modify, and the resources are fixed and closed. Microclasses and MOOC are rich in resources, easy to search, easy to spread, and widely used.

Video cutting is essential in video production. Video cutting is a video editing method to cut the video material and simulate multiview effects. It aims to generate multiview video from limited views. Fixed position camera images usually cannot provide enough visual interest or guidance for the viewer. Kong et al. applied virtual photography to a fixed lens speech scenes, according to the photography rules, used computer vision and signal processing methods to select the appropriate lens, used the image synthesis method to generate new images from the original material, and formed a variety of different lenses and visual effects, finally to guide the audience's attention and maintain the audience's visual interest. Later, scholars extended this method to live broadcast and other activities [5]. Ren et al. studied the automatic editing method of automatically generating multiview video from single-view video and automatically generating multishot video from single-view video. Their main idea is to capture the whole field of view of the event with a single fixed camera and simulate the camera's translation, tilt, and zoom movement by clipping and zooming the original video material. In this method, L1 regularization optimization was used to calculate the composition of each clip, and the editing practice was encoded as a single cost function. Finally, the editing problem was transformed into an optimization problem. In addition, there is eye movement-based clipping [6]. Cabo et al. studied a method to find the clipping position and calculate the clipping size by collecting the eye data of the viewer, which can also achieve the effect of camera motion such as translation, tilt, zoom, and so on [7]. Zhu and Zhou further optimized the cutting window moving path algorithm to obtain better results. This kind of automatic video editing requires very strict material, which must be recorded by wide-angle static cameras. In a sense, cutting and editing is through the completion of the material in exchange for camera motion and lens switching simulation effects [8].

Therefore, to improve the effect of video production, microclass combined with the MOOC video production method is considered for reconstructing teachers from the traditional classroom into the course design thinking of network classroom and traditional classroom video to adapt to the fragmented learning needs of the mobile Internet era of shooting technology, which has a practical significance on

how to develop high-quality video courses and solve its existing problems. The proposed study can help researchers to devise a new solutions in the area of research.

The study is organized as follows: Section 2 shows the MOOC-based video production method of microclass integrated with MOOC. Section 3 represents results of the study. The study is concluded in Section 4.

2. MOOC-Based Video Production Method of Microclass Combined with MOOC

The following subsections show the details of this section.

2.1. MOOC-Based Video Development of the Microclass Combined with MOOC

2.1.1. General Process of Video Development. In recent years, with the development of multimedia technology, the development methods and means of video teaching resources present multiple styles, and the video compilation process is also more simplified. People pay more attention to the use of video course shooting technology and despise video production. As a result, most videos on the current network platform have many shortcomings, and the video course development process is increasingly confused. As a network video course evolved and developed from traditional educational TV, its development process can be based on the production mode of educational TV and form its own development mode [9]. The development of MOOC-based microclasses combined with MOOC video focuses on learner-centered design. Therefore, video design is particularly important. The former design conforms to the learners' teaching content, while the latter design conforms to the teaching content, video form, and diversified recording schemes and production means make the video form richer. Therefore, this study constructs the development process of MOOC-based microclass combined with MOOC video, as shown in Figure 1. The figure shows generic flowchart of video development containing different steps.

MOOC-based microclass combined with MOOC video development includes the preliminary design stage, shooting preparation stage, recording stage, and postproduction stage to complete the process of video production and release to the cloud. Among them, the video postproduction stage is crucial.

2.1.2. Recording Process of Video Development

(1) Shooting. The video recording process is an important work before video shooting. A relatively complete process of professional video course recording can improve the video recording level and work efficiency [10]. The video development team should prepare the script and course type according to the preliminary design stage and formulate the recording process. Based on microclass and the MOOC's video production method based on MOOC, the production

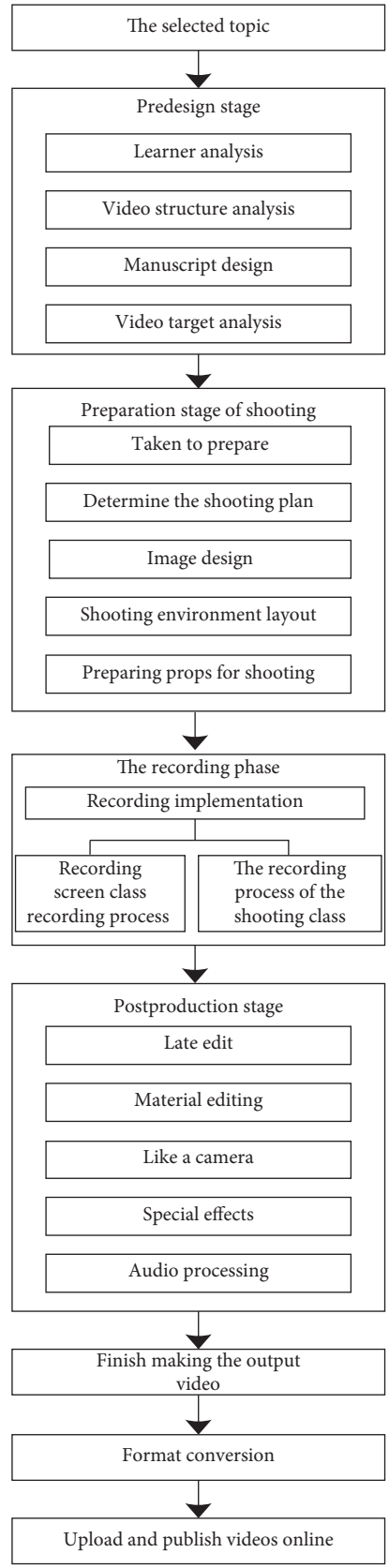


Figure 1: The general flow of video development.

of excellent resource sharing class is mainly based on shooting. The general professional level video recording process is mainly composed of the camera, switchers, audio, teacher, computer, display screen, and nonlinear editing (as shown in Figure 2). The camera, tuning system, and teacher computer are connected to the switcher at the same time. Then, the recorded video is imported into the nonlinear editing system by switching the pilot system to implement the postediting work of the video. At the same time, the teacher's computer is connected to the display screen to present the content picture on the computer screen. The video recording process is shown in Figure 2 which consists of various phases.

In the process of video recording, the installation of several cameras is the key link. According to the different types of video courses, a single camera, two cameras, or three cameras can be used to shoot the course. Stand-alone shooting: when using stand-alone shooting, the camera lens should shoot the teacher's lecture activities in front and record the whole process of the teacher's lecture [11]. The lens scene can use medium range or close range to highlight the teacher's main image. This shooting method can create a picture feeling that the teacher is giving a lesson to the viewer instead of the viewer watching the teacher giving a lesson to others. However, the disadvantages of single shot are single scene, easy to create long-term viewer, visual fatigue, lack of classroom teaching, atmosphere of interaction between teachers and students, and single scenes in the later editing process, such as the same scene group will appear frame skipping reality, which is more difficult. Two-camera shooting: when two-cameras shooting is adopted, a panoramic view can be taken from the front of No.1 camera, close-up views of teachers' lectures or students' listening and answering questions can be taken from No.2 camera, the blackboard and display screens can also be taken, and scenes switching can reasonably use panoramic view, medium view, or close-up view and other lens language according to the teaching plot [12]. This kind of shooting method can record all activities of classroom teaching completely. The video picture has a strong sense, and the scenes are rich. The smoothness of lens assembly is good, which is more in line with the process of human visual thinking activities. However, attention should be paid to the lens assembly of the two machine positions to avoid the phenomenon of cross-axis. Shooting with three cameras: when shooting with three cameras, it is generally according to the panoramic position in the middle of the classroom, with two sides on each side. In one machine setting, this shooting method can fully record all aspects of the whole teaching process, with rich scenes, large amount of information, high flexibility, but large investment. The development conditions of the online video courses are limited. Without switching the platform, a camera can be used to capture the content of the display screen, and the content of the corresponding period can be inserted according to the teaching rhythm in the later editing process. The teacher's presentation can

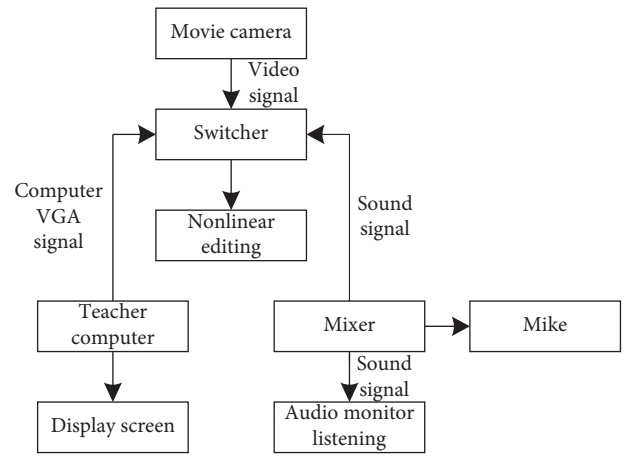


Figure 2: Video recording process.

also be a directly generated video or image by the software and edited to the corresponding time node to form a complete video in line with the learning logic courses.

(2) Screen Recording. The video production process of screen recording is simple, and the personnel input is small. Generally, a computer screen recording software, external drawing board, and cameras (including computer integrated camera) can be competent for the production of screen recording video courses. Video recording courses can be divided into simple slide and slide + teacher images. Simple slide type video courses only have slide pictures and the voice of teachers, but not the image of teachers and students [13]. Teachers can use the sketchpad function to write the drawing or detailed derivation process in PPT courseware and then guide students to pay attention and think just like the video course production form of Khan Academy. Slide + teacher image video generally needs to connect a camera or computer-integrated camera on the computer, and teachers record slides through a screen recording software, but also take their own images. In the postproduction, according to the needs of the teacher image and slide implementation synthesis, recording slide video is easy to make students feel boring, and occasionally, a teacher image can attract students to pay more attention to learning to improve learning efficiency.

2.1.3. Video Editing. The postediting of a video course is to import video materials into a nonediting software and use film and television techniques to implement editing and feature processing to improve the effect of video production. The video editing process is shown in Figure 3 which contains various steps.

In the figure, first, it needs to prepare the video and media materials, open the selected nonlinear video editing software, establish the project engineering file, and set the engineering parameters according to the video course production technical standards; second, it imports the materials into the video and audio tracks for editing and special effects processing [14]; during the editing process, the relevant media materials are inserted according to the content expression needs; after editing the first draft, the

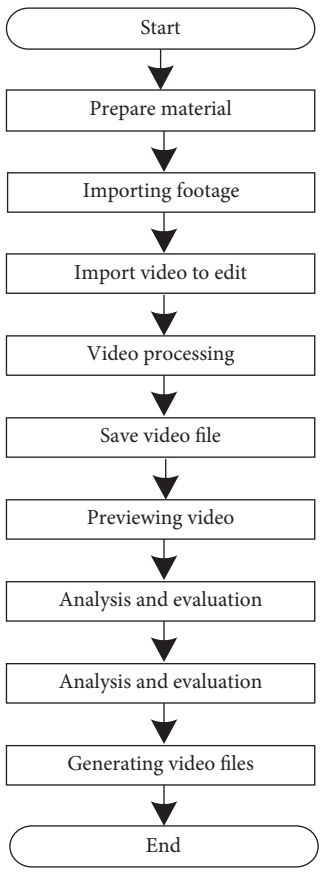


Figure 3: Video editing process.

instructor and director should preview the video course, analyze and evaluate the video course, and edit the unsatisfied parts again until the desired effect is achieved.

2.2. Video Production Method of Shot Assembly Based on the Hidden Markov Model. Shot assembly refers to the logical connection of video images to describe the occurrence of something. In the existing work, automatic shot assembly is mostly organized according to the timeline of the event, which is used to express the story or documentary, and cannot edit the video material without an obvious time relationship [15]. This study compares the traditional shot assembly method based on natural timeline and proposes an artificial time line based on information arrangement logic, which logically connects the subshots to achieve the purpose of display. At the same time, the global similarity constraint is established to improve the richness of video information. In the process of shot combination, the coherence editing rules and industry experience are coded, and the hidden Markov model is established to ensure the coherence of the front and back shots.

2.2.1. Introduction of the Hidden Markov Model. The hidden Markov model is a probability graph model, which is the simplest dynamic Bayesian network. Its standard model includes a group of state variables and a group of observation variables [16]. It is usually assumed that the state variable is hidden and unobservable $\{y_1, y_2, y_3, \ldots, y_n\}$, where $y \in Y$ represents the state at time i, and observed variable is $\{x_1, x_2, x_3, \ldots, x_n\}$, where $x \in X$ represents the observed value at time i. The state of time i is only determined by time $i-1$ and has nothing to do with other time $i-2$. The observation variables of time i are only related to the state of this time and have nothing to do with the state and observation variables of other times. Each observation variable x_i comes from the observation space $U = \{u_1, u_2, \ldots, u_M\}$, and each hidden state y_i comes from the hidden space $T = \{t_1, t_2, \ldots, t_M\}$. The hidden Markov model is described by the $N*1$-dimensional initial probability matrix $P(c_0^i)$, the $N*N$-dimensional state transition matrix $A(c_t^i, c_{t+1}^j)$, and the $N*M$-dimensional conditional probability $P(y \mid x)$ defined by the output observation matrix $B(c_t^i)$. The definitions of the three probability matrices are as follows:

State transition probability: the transition probability of the model among various states is usually recorded as matrix $A = [a_{ij}]_{N*N}$, where a_{ij} is

$$a_{ij} = P\left(y_{t+1} = s_j \mid y_t = s_i\right), \quad \sum_j a_{ij} = 1, \quad 1 \le i, j \le N, \tag{1}$$

$$a_{ij} \propto A\left(c_t^i, c_{t+1}^j\right). \tag{2}$$

Output observation probability: the probability of each observation value obtained by the model according to the current state is usually recorded as the matrix $B = [b_{ij}]_{N*M}$, where b_{ij} is

$$b_{ij} = \left(x_t = o_j \mid y_t = s_i\right), \quad \sum_j b_{ij} = 1, \quad 1 \le i, j \le N, \tag{3}$$

$$b_{ij} \propto B\left(c_t^i\right). \tag{4}$$

Initial probability: the probability of each state of the model at the initial time, usually is denoted as $\pi = [\pi_i]_{1*N}$, where π_i is

$$\pi_i = P\left(y_i = s_i\right), \quad 1 \le i \le N, \tag{5}$$

$$\pi_i \propto P\left(c_0^i\right). \tag{6}$$

2.2.2. Editing Rule Coding. In this study, the shot assembly problem is modeled as a hidden Markov model. To solve the specific calculation problem of parameters in the model, this study adopts a coding method similar to dialogue-driven automatic editing and dance scene automatic editing and implements coding based on editing rules [17]. c_t^i means that the state at time t is i. Each rule defines the initial probability, observation probability, and transition probability.

(1) Avoid skipping: avoid the short-term interruption or change of the front and back shots during assembly. In this study, it is coded as the similarity problem between adjacent shots. When the front and back similarity $\mathrm{sim}(c_t^i, c_{t+1}^j)$ of adjacent shots is higher than the given threshold τ, it is considered as skipping, and the corresponding state transition probability is set to 10^{-5}. If it is not skipping, the probability is 1. It does not affect the observation and initial probability:

$$\begin{aligned} A\left(c_t^i, c_{t+1}^j\right) &= \begin{cases} 1 & \mathrm{sim}\left(c_t^i, c_{t+1}^j\right) \le \tau, \\ \in & \text{otherwise}, \end{cases} \\ B\left(c_t^i\right) &= 1, \\ P\left(c_0^i\right) &= 1, \end{aligned} \tag{7}$$

(2) Picture motion coherence: the picture motion intensity at the end of the front shot and the beginning of the next shot should be as close as possible. In this study, the main body motion intensity matching is coded as the difference $d_{\mathrm{move}}(c_t^i, c_{t+1}^j)$ of the motion intensity value. The larger the difference, the lower the probability of the corresponding state transition and the higher the contrary. It does not affect the observation and initial probability:

$$\begin{aligned} A\left(c_t^i, c_{t+1}^j\right) &= \frac{1}{d_{\mathrm{move}}\left(c_t^i, c_{t+1}^j\right) + 1}, \\ B\left(c_t^i\right) &= 1, \\ P\left(c_0^i\right) &= 1. \end{aligned} \tag{8}$$

(3) Main body position coherence: when assembling, the main body positions at the end of the front shot and the beginning of the next shot should be as close as possible. In this study, the main body position matching is coded as the difference $d_{\mathrm{pos}}(c_t^i, c_{t+1}^j)$ of the main body's x-direction position value. The larger the difference, the lower the probability of corresponding state transition and the higher the contrary. It does not affect the observation and initial probability:

$$\begin{aligned} A\left(c_t^i, c_{t+1}^j\right) &= \frac{1}{d_{\mathrm{pos}}\left(c_t^i, c_{t+1}^j\right) + 1}, \\ B\left(c_t^i\right) &= 1, \\ P\left(c_0^i\right) &= 1. \end{aligned} \tag{9}$$

(4) Picture tone coherence: the picture color temperature at the end of the front shot and the beginning of the next shot should be as close as possible. In this study, the picture color temperature matching is coded as the difference $d_{\mathrm{color}}(c_t^i, c_{t+1}^j)$ of the color temperature value. The larger the difference, the lower the probability of the corresponding state transition and the higher the contrary. It does not affect the observation and initial probability:

$$\begin{aligned} A\left(c_t^i, c_{t+1}^j\right) &= \frac{1}{d_{\mathrm{color}}\left(c_t^i, c_{t+1}^j\right) + 1}, \\ B\left(c_t^i\right) &= 1, \\ P\left(c_0^i\right) &= 1. \end{aligned} \tag{10}$$

2.2.3. Layout Constraints of Lens Information. Because there is no obvious time relationship between the video materials, there is no natural timeline. It needs to manually set the timeline of the video arrangement. To maximize the information richness of the video, it is necessary to implement constraints on the logic of information arrangement, that is, the time line arrangement of subshots [18]. In this study, according to the habit of film and television lens assembly, the information layout logic mode is divided into four types: forward type, backward type, jumping type, and ring type. According to the distance of the lens, the distance of the sight distance is adjusted to switch the audience's line of sight between the whole and the part, so that the audience's visual experience can be affected accordingly.

(1) Forward: the scope of the work is from large to small, and the lens is from far to near.

(2) Step back: the scope of the work is from small to large, and the lens is from near to far.

(3) Jump: the scope of the exhibition of the work from the largest directly to the smallest, and gradually increase the lens from the farthest to the nearest, gradually pulling away.

(4) Ring type: the scope of work display is from large to small, then from small to large, and the lens is from far and near and far.

2.2.4. Sublens Assembly Modeling. The ultimate goal of video editing is to combine a certain length of subshot sequence from a group of candidate subshots, and the final length of the sequence can be specified by the user. From the total video time T and the average subshot time t, it is needed to obtain the video sequence composed of $L = T/t$ subshots. Assuming that the number of effective subshots in the raw material is N, then N^L sequence spaces are obtained. Our task is to select the best sequence to meet the needs of effective information expression [19]. In this study, the group join problem is modeled as a hidden Markov model, and the sequence with the highest probability is obtained by the Viterbi algorithm.

(1) Assembly Model. In this study, video assembly is transformed into the problem of selecting the best subshot sequence from the subshot set, which is modeled as a hidden Markov model. The subshot set represents the hidden space, the shot of the ith node of the sequence represents the hidden

state at time i, and the information type of the shot of the ith node of the sequence represents the observation state at time i. The initial probability matrix π controls the selection of the initial shot, the transition probability matrix A controls the selection of the next shot when the current shot has been determined, and the output observation matrix B controls the selection of the information type.

The transition probability matrix A is determined by the coherent clips and "avoid skipping clip." The initial probability matrix π is determined by facial emotion constraints, and the observation matrix B is determined by the information layout constraint. Therefore, the video assembly problem is transformed into one of the three basic problems of the hidden Markov model. Given the model $\lambda = [A, B, \pi]$ and the observation sequence $x = \{x_1, x_2, x_3, \ldots, x_n\}$, how to find the best matching state sequence $y = \{y_1, y_2, y_3, \ldots, y_n\}$ with the observation sequence. In other words, how to train the model to best describe the observed data.

(2) Viterbi Algorithm Based on Global Constraints. In this study, the connection relationship of shots in video clips is modeled as a hidden Markov model, the information sequence is regarded as observation sequence, $o = \{o_1, o_2, o_3, \ldots, o_n\}$, constrains the information arrangement of shots, and the final shot sequence is regarded as a hidden sequence. Given the a transition probability matrix, output observation matrix B, and initial state matrix π, the problem of group join is transformed into the problem of solving the optimal sequence because considering the problem of nonrepetition and difference of nodes [20], the Viterbi algorithm needs to be improved to solve the problem of global constraint optimization, and the global constraint optimization is applied to solve the problem.

The classical Viterbi algorithm uses the idea of dynamic programming to find the optimal path. In the process of solving, the probability of passing each node is recorded, and the maximum value is selected as the next node. Finally, the optimal path from the starting point to the end point is found through backtracking. $\delta_t(i)$ is defined as the best path to the node when the hidden state is i at time t, that is, the maximum probability of all possible transition paths from node to node. $\psi_t(i)$ is defined as the hidden state at time t, when the hidden state is i in the best path to the node, the hidden state at time $t-1$ is the hidden state of the previous node of the current node in the best path, and i_t^* is defined as the hidden state of the final sequence at time t.

The state of each phase of the algorithm is as follows:

$$\delta_1(i) = \pi_i b_i(o_1), \quad i = 1, 2, \ldots, N. \tag{11}$$

$$\psi_t(i) = 0, \quad i = 1, 2, \ldots, N. \tag{12}$$

The dynamic programming recursive $t = 2, 3, \ldots, T$ is carried out.

$$\delta_1(i) = \max_{1 \ll j \ll N} \left[\delta_{t-1}(j) a_{ji}\right] b_i(o_t), \quad i = 1, 2, \ldots, N, \tag{13}$$

$$\psi_t(i) = \arg \max_{1 \ll j \ll N} \left[\delta_{t-1}(j) a_{ji}\right], \quad i = 1, 2, \ldots, N. \tag{14}$$

The maximum $\delta_T(i)$ at time T is calculated, which is the most likely the hidden state of T.

$$i_T^* = \arg \max_{1 \ll j \ll N} [\delta_T(i)], \quad \max_{1 \ll j \ll N} [\delta t - 1]. \tag{15}$$

Using local state to make backtracking:

$$i_{t-1}^* = \psi_t(i_t^*). \tag{16}$$

From the maximum value of each node $\delta_T(i)$ at the last time T, it can know that the hidden state at time T is i, and the transition path is the best path among all paths, which can determine the node at the end of the sequence. According to $\psi_t(i)$, the last node of the optimal path can be obtained, and a complete path can be obtained.

In this process, $\psi_t(i)$ represents the hidden state of the node before the current node in the best path of node i at time t. The determination of the last hidden state is determined by traversing the product of the maximum probability $\delta_{t-1}(j)$ of all nodes in the possible transition path at the previous time and the transition probability of the current node i. In all states, the one with the highest probability is selected, namely, $\max_{1 \ll j \ll N} [\delta_{t-1}(j) a_{ji}]$. In this process, the previous hidden state selection does not consider all nodes before it, so it is necessary to add the constraint that the nodes are not repeatable and have certain differences. When selecting the last hidden state j, it needs to traverse all hidden states before time $t-1$ in the transition path. The hidden state before time $t-1$ cannot be j, and the picture cannot be too similar to the j state. If the condition is negative, it needs to discard j, select the suboptimal node k, and continue to traverse the path of node k to ensure that there is no repetition and difference between nodes until the condition is met. This can fully ensure the richness of video information.

3. Results

To test the effect of the method in this study, a high school's teaching practice class is selected as the experimental object. There are 80 students in the class, including 35 girls. This class's academic performance is average, and learning attitudes and classroom discipline are similar to other urban classes, more representative. The experiment of mechanics, as a part of the teaching content of physics in high school, has been mentioned in the previous new teaching. At that time, the teacher also made a classroom demonstration. Therefore, before playing the video produced by this method, it is necessary to arouse students' memory, and the two teaching methods form a contrast. Then, through the form of a questionnaire survey, from the students' attitude to this method of making video, this method is used to make the video learning effect, and two aspects of investigation and analysis are carried out.

80 questionnaires are sent out, and 75 valid questionnaires are collected. First, the survey data are collected, and then, the statistical data are processed by Excel. The survey results are analyzed as follows:

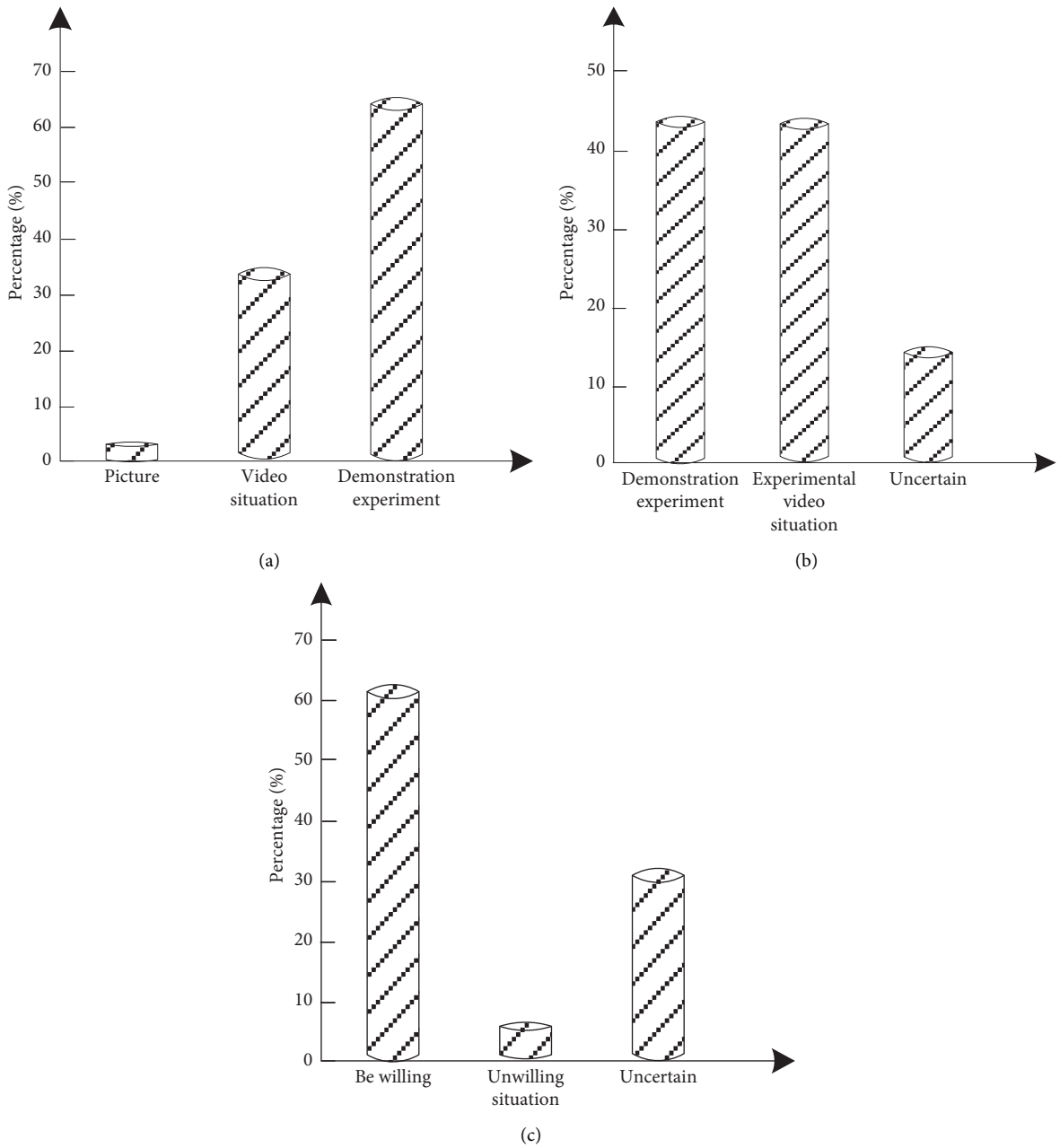


Figure 4: Survey results. (a) Question 1 survey results. (b) Question 2 survey results. (c) Question 3 survey results.

The students' attitude towards this method of making a video is set as follows.

Question 1: among the following teaching resources, you are most interested in the following situations: pictures, videos, and classroom demonstrations

Question 2: when you watch the classroom demonstration experiment and the experimental video, you think that the teacher's demonstration experiment, experimental video, and uncertainty are more conducive to learning

Question 3: are you willing to learn through the teaching video in the future? It can be divided into willing, unwilling, and uncertain. The experimental results are described in Figures 4(a)–4(c).

The data show that among the three teaching resources of pictures, videos, and demonstration experiments, the

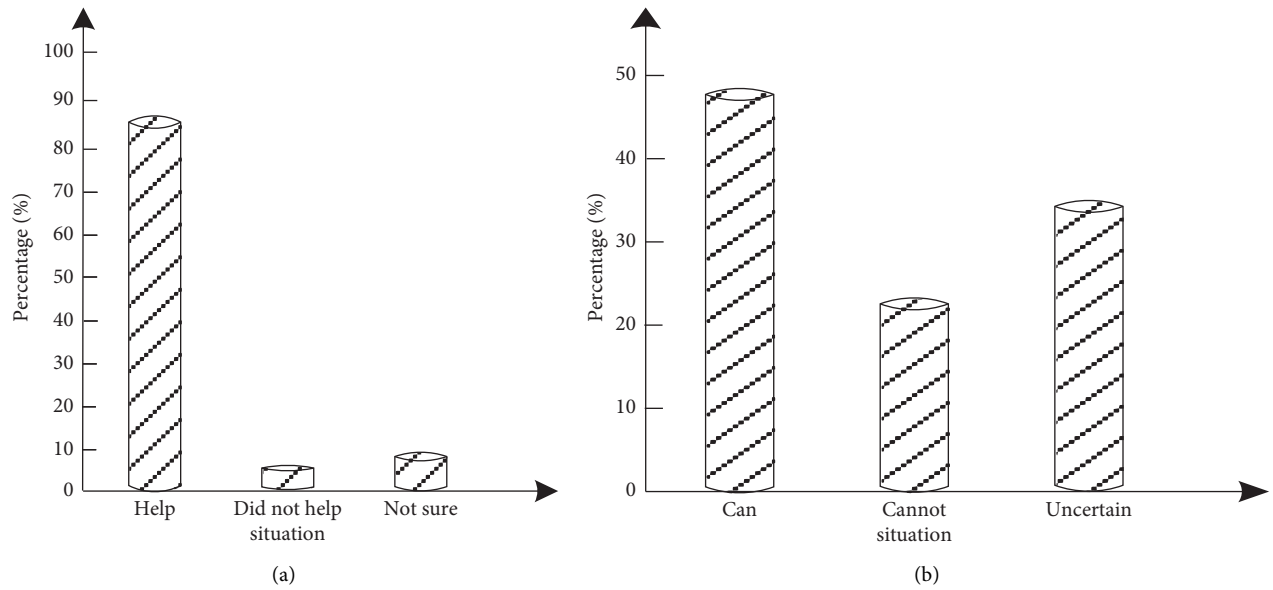


Figure 5: Learning effect survey results. (a) Question 4 survey results. (b) Question 5 survey results.

Table 1: Statistics of the application effect of the method in this study.

Test project		The sound is accurate and smooth	The video is clear and coherent	The video content is short and vivid
Attitude (%)	Very satisfied with	92.6	90.2	91.3
	Basic satisfaction	5.3	6.7	6.2
	General	2.1	3.1	2.5
	Not satisfied with	0	0	0

demonstration experiment is the most popular among students, accounting for 67.1%, far ahead of the second video, while the pictures are the least popular. It shows that the students are eager to experience the experimental operation and witness the experimental phenomenon. However, this does not mean that the demonstration experiment is the most beneficial way to learn, which can be seen from the survey results of question 2; the number of students who choose the experimental video is almost the same as the number of students who choose the demonstration experiment. This is because although the demonstration experiment can increase the sense of on-site experience, it is time-consuming and inconvenient to observe the phenomenon. The students in the back row can hardly observe the phenomenon, while the experimental video is clear and time-saving. Although the experimental video itself has some shortcomings, most students are still full of expectations for it, and 63.2% of them are willing to learn through the video in the future.

Using this method to make a video learning effect, the following two questions are set.

Question 4: does watching the video help you understand the mechanics: helpful, not helpful, and uncertain

Question 5: compared with listening to the teacher's explanation in the classroom, you can get the same learning effect by watching the "Mechanics" video after class: yes, no, and uncertain. The results are described in Figures 5(a) and 5(b).

Data show that by watching the video, as high as 85.2% of the students think that it is helpful to understand the broken key situation, while only 5% of the students think it is not helpful. For the fifth question, 49.8% of the students think that the effect of the self-study by watching video is the same as that of classroom listening, 24.1% think that the effect of classroom listening will be better, while other students have no agreement. This result is understandable. In the face of a new way of learning, some students will certainly show their maladjustment and be faithful to the way they have been used to. It is gratifying that nearly 50% of the students agree with it.

According to the questionnaire, the application effect investigated for the sound is more accurate and fluent, the video is clear and coherent, and the video content is short and vivid. The statistical results are given in Table 1.

From the above table, we can see that learners are generally very satisfied with the evaluation of the video produced this time. The average satisfaction is 91.36% in

terms of more accurate and fluent voice, clear and coherent video, short and vivid video content, and other options, which show that the application effect of the proposed method is better.

4. Conclusion

At present, driven by the concept of global open education resource sharing, massive MOOC are developing rapidly with the passage of time. Current, in the background of continuing to promote large-scale open online curriculum movement, we still have many tasks to do and many problems to solve in the digitization and automation of educational resources. The development of network video is only the tip of the iceberg in the digitization of educational resources. However, it is a more effective and efficient medium and approach to spread teaching contents in the new information technology education environment. The development of an online video course is a complex, professional, and high investment work which requires large resources. From the early classroom video to the excellent course video open class and now to the MOOC video and microclass video, people have never stopped exploring how to develop the high-quality video course under the network learning environment. The traditional video production methods are not achieving the requirements of today's era, and the video production approach is backward. This study focusses on the MOOC-based video production approach of microclass integrated with MOOC to improve the effect of video production for the best utilization. The experimental results of the study show that the video produced by this method has good and efficient results and a good application effect. The study will help researchers to devise novel solutions in the area of research.

Data Availability

The datasets used and/or analyzed during the current study are available from the corresponding author upon request.

Conflicts of Interest

The author declares that there are no conflicts of interest.

References

[1] W. Li and C. Wang, "Composition effect on elastic properties of model nico-based superalloysproject supported by the national key research and development program of China (grant no. 2017yfb0701502)," *Chinese Physics B*, vol. 29, no. 2, Article ID 026102, 2020.

[2] D. Zhang, C. Wu, and J. Liu, "Ranking products with online reviews: a novel method based on hesitant fuzzy set and sentiment word framework," *Journal of the Operational Research Society*, vol. 71, no. 3, pp. 528–542, 2020.

[3] W. Jindong, G. Junjie, L. Haitao, D. Yufen, and L. Junfeng, "Research of omnidirectional measurement method to detect the errors of multi-axis nc machine tool based on laser tracker," *Iet Science Measurement & Technology*, vol. 12, no. 3, pp. 418–425, 2018.

[4] Z. Zhuang, L. Wei, L. Xie, H. Ai, and Q. Tian, "Camera-based batch normalization: an effective distribution alignment method for person re-identification," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 99, p. 1, 2021.

[5] Q. Kong, G. Jiang, Y. Liu, and J. Sun, "Location of the leakage from a simulated water-cooling wall tube based on acoustic method and an artificial neural network," *IEEE Transactions on Instrumentation and Measurement*, vol. 70, no. 99, pp. 1–18, 2021.

[6] J. J. Ren, J. W. Wang, and L. VR. Cao, "Panoramic video production software object visualization 3D virtual simulation," *Computer Simulation*, vol. 37, no. 09, pp. 299–302+311, 2020.

[7] C. Cabo, E. Sanz-Ablanedo, J. Roca-Pardinas, and C. Ordonez, "Influence of the number and spatial distribution of ground control points in the accuracy of uav-sfm dems: an approach based on generalized additive models," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 1, no. 99, pp. 1–10, 2021.

[8] S. Zhu and Y. Zhou, "On the optimal voronoi partitions for ahlfors-david measures with respect to the geometric mean error," *Journal of Mathematical Analysis and Applications*, vol. 498, no. 2, Article ID 124897, 2021.

[9] V. A. Ezhov, "Concept of autostereoscopic full-screen resolution imaging based on a dynamic amplitude parallax barrier with two antiphase liquid-crystal layers," *Optical Engineering*, vol. 57, no. 11, p. 1, 2018.

[10] K. Shinha, W. Nihei, T. Ono, R. Nakazato, and H. Kimura, "A pharmacokinetic–pharmacodynamic model based on multi-organ-on-a-chip for drug–drug interaction studies," *Biomicrofluidics*, vol. 14, no. 4, Article ID 044108, 2020.

[11] Z. Shao, "Examining the impact mechanism of social psychological motivations on individuals' continuance intention of MOOCs," *Internet Research*, vol. 28, no. 1, pp. 232–250, 2018.

[12] F. Zhang, D. Liu, and C. Liu, "Mooc video personalized classification based on cluster analysis and process mining," *Sustainability*, vol. 12, no. 7, p. 3066, 2020.

[13] W. Ling, H. Gongliang, and Z. Tiehua, "Semantic analysis of learners' emotional tendencies on online mooc education," *Sustainability*, vol. 10, no. 6, p. 1921, 2018.

[14] J. A. Ruipérez-Valiente, S. Martin, J. Reich, and M. Castro, "The unmoocing process: extending the impact of mooc educational resources as oers," *Sustainability*, vol. 12, no. 18, p. 7346, 2020.

[15] A. Goriaev, T. Wauters, S. Mller, R. Brakel, and M. V. Schoor, "The upgraded tomas device: a toroidal plasma facility for wall conditioning, plasma production, and plasma–surface interaction studies," *Review of Scientific Instruments*, vol. 92, no. 2, Article ID 023506, 2021.

[16] L. H. Cherri, A. C. Cherri, M. A. Carravilla, J. F. Oliveira, F. M. Bragion Toledo, and A. C. Goncalves Vianna, "An innovative data structure to handle the geometry of nesting problems," *International Journal of Production Research*, vol. 56, no. 23-24, pp. 7085–7102, 2018.

[17] M. Hasni, M. S. Aguir, M. Z. Babai, and Z. Jemai, "Spare parts demand forecasting: a review on bootstrapping methods," *International Journal of Production Research*, vol. 57, no. 15-16, pp. 4791–4804, 2019.

[18] R. Pellerin and N. Perrier, "A review of methods, techniques and tools for project planning and control," *International Journal of Production Research*, vol. 57, no. 7-8, pp. 2160–2178, 2019.

[19] L. Deng, Y. Wang, Z. Han, and R. Yu, "Research on insect pest image detection and recognition based on bio-inspired methods," *Biosystems Engineering*, vol. 169, pp. 139–148, 2018.

[20] K. Ma, R. Pal, and E. Gustafsson, "What modelling research on supply chain collaboration informs us? identifying key themes and future directions through a literature review," *International Journal of Production Research*, vol. 57, no. 7-8, pp. 2203–2225, 2019.

Hindawi
Scientific Programming
Volume 2021, Article ID 5564326, 9 pages
https://doi.org/10.1155/2021/5564326

Research Article

Research on the Service Mode of the University Library Based on Data Mining

Sha Duan and Ziwei Wang

Hebei Women's Vocational College, Shijiazhuang 050091, China

Correspondence should be addressed to Ziwei Wang; wangziwei001@cumt.edu.cn

Received 1 March 2021; Revised 29 March 2021; Accepted 8 April 2021; Published 26 April 2021

Academic Editor: Shah Nazir

In the digital information age, data mining technology is becoming more widely used in libraries for its useful impact. In the context of big data, how to efficiently mine big data, extract features, and provide users with high-quality personalized service is one of the important issues that needs to be solved in the current university library big data application. Brain computing is a kind of comprehensive processing behavior of the human brain simulated by the computer, which can comprehensively analyze a variety of information and play a very good guiding role in processing library service behavior. This paper briefly introduces the related concepts and algorithms of data mining technology and deeply studies the classical algorithm of association rules, namely, Apriori algorithm, which analyzes the necessity and feasibility of applying data mining technology to university library management. The design idea and functional goal of the college book intelligent recommendation system are based on the decision tree method and association rule analysis method. Through the application research of data mining technology in the personalized service of the university library, combined with the actual work, this paper proposes data mining of association rules in the university library system. The research further elaborates on the system architecture, data processing, mining implementation algorithms, and application of mining results. The experimental results of the research have certain significance for the university library to explore personalized services, provide book recommendation services, and make corresponding decisions to optimize the library's collection layout.

1. Introduction

The concept of data mining originated from the 11th International Conference on Artificial Intelligence held in Detroit, USA, in August 1989. At that time, the concept of knowledge discovery (KDD) was proposed, which refers to the extraction or mining of hidden information from a large amount of data. Data mining technology uses statistical and artificial intelligence technology applications to integrate various types of information data, extract a large amount of useful information from massive data, and explore the rules, thereby improving the efficiency of production and service [1]. According to the comprehensive data, the data mining analysis methods include description and visualization, that is, using visualization tools to display, analyze, and drill data, so that the data mining analysis results are more vivid and profound; classification, that is, through the preset data classification model; screening the classification data; estimating, that is, taking the collected data to obtain the value of the continuous variable through the estimation, then classifying according to a preset threshold such as 0–9, and predicting, that is, by classifying or estimating the model, thereby unknown variable prediction; correlation grouping or association rules, that is, using association rules and sequence analysis to discover the law of what is going to happen; clustering, that is, grouping records, and recording similar records in a cluster so that each group has predictive or implied features; and complex data types (text, Web, graphics, video, audio, etc.) mining. Data mining technology requires database systems to provide efficient storage, indexing, and query processing support and to use high-performance (parallel) computing techniques when dealing

with massive datasets, such as distributed technology and crawling technology for rapid crawling of network information.

Compared with library development, data mining technology has developed from computer science research for more than a decade. In the middle and late 20th century, foreign scholars began to study the application of data mining technology in libraries. Domestically, with the development of the information age and the gradual accumulation of digital resources, digital libraries came into being [2]. University libraries began to introduce automated database-based management systems, and the number of databases increased dramatically. The application has gradually broadened and gradually infiltrated into the business fields of university library management and information services.

The specific contributions of this paper include the following:

This paper introduces the related concepts and algorithms of data mining technology and deeply studies the Apriori algorithm, a classical algorithm of association rules

This paper analyzes the necessity and feasibility of applying data mining technology to university library management

The system architecture, data processing, mining algorithm, and application of mining results are described

Performance analysis of the proposed algorithm and an evaluation of the algorithm with respect to other existing algorithms are given

The rest of this paper is organized as follows. Section 2 discusses the basic algorithm of data mining, followed by university library personalized services discussed in Section 3. The analysis of experimental results is discussed in Section 4. Section 5 concludes the paper with summary and future research directions.

2. Basic Algorithm of Data Mining

2.1. Definition of Data Mining. At present, there are many definitions of data mining. In short, data mining is to extract or "dig" knowledge from massive data. Currently, the broad definition of data mining is as follows: data mining is the process of mining useful content from a large amount of data placed in a database, data warehouse, or other information bases. A typical data mining system generally has the following components, as shown in Figure 1.

Data mining is the integration of multidisciplinary technologies, including database technology, statistics, machine learning, pattern recognition, artificial neural networks, data visualization, knowledge extraction, image and signal processing, and spatial data analysis. Data mining systems can also integrate techniques for spatial data analysis, information extraction, image analysis, signal processing, computer graphics, economics, or psychology.

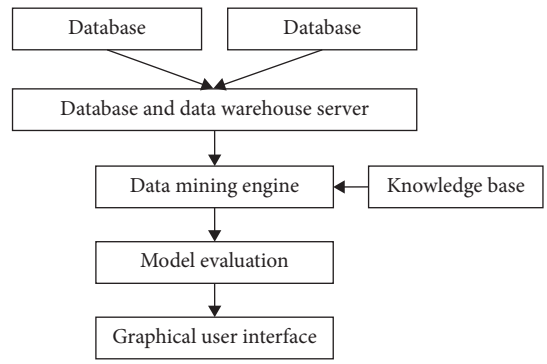


Figure 1: Composition of data mining.

Through data mining, interesting knowledge and laws implicit in massive data can be found from the database. These laws or knowledge can be applied in business areas such as guiding decision-making, process control, sales promotion, and medical diagnosis. The data mining system can also browse and store knowledge quickly, and at the same time can facilitate our research and study. Therefore, data mining is considered to be one of the most important frontier disciplines in the information industry and the most promising interdisciplinary subject in the information industry [3].

2.2. Data Mining Process. Data mining can be understood as a process of human-computer interaction through computer processing, manual analysis, and other methods [3]. The process is complete but iterative, mainly including data preparation, data selection, data preprocessing, data mining, and transformation model and mode. The five stages of data mining are shown in Figure 2.

2.3. Common Algorithms for Data Mining

2.3.1. Decision Tree Classification Algorithm. Decision trees have simple and efficient classification results. They mainly reflect the influence of different attributes on the instance by constructing a tree-like form, and its leaf nodes represent the categories to which it belongs to. For a tree branch from the root to the corresponding different leaf nodes, it can be equivalent to a conjunction rule, so the decision tree is equivalent to a collection of multiple rules.

Decision trees can be divided into two different types, classification tree and regression tree, each with its own strengths. The classification tree mainly constructs a tree structure for discrete attribute variables. The main function is to mark and classify the data. The regression tree mainly constructs a tree structure for continuous attribute variables. The main function is the value of the target variable. In general, the decision tree is for a given new data record, through its construction form to predict the category to which the record belongs to. The advantage of the decision tree is that the structure is simple, easy to understand, high in classification accuracy, and easy to optimize the overfitting of the data. The disadvantage is that the data are

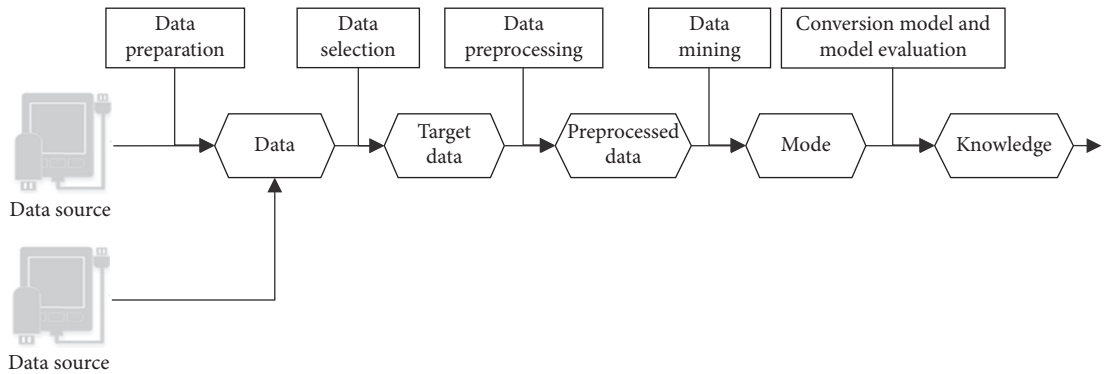


Figure 2: Data mining process.

relatively easy to handle, and it is difficult to process for complex data [4].

2.3.2. Artificial Neural Network. Artificial neural network originated from the characteristics of the animal neural network in biology, which is simply referred to as the neural network or connection model, and is a parallel distributed processing model. Compared with traditional artificial intelligence and information processing technology, the mechanism of the neural network is completely different, and it has the characteristics of adaptability, controllability, and multilayer training and learning.

At present, neural networks are mainly used in a wide range of fields such as image processing, predictive classification, pattern recognition, automatic control, machine learning, and medical diagnosis. Artificial neural networks have predictions of the results of complex relationships, but due to the complexity of their internal structure, the results of the predictions cannot be analyzed in detail. In addition, when there are too many input neuron nodes in the input layer of the artificial neural network, after the data training, the possible prediction results are not perfect. Therefore, in practical applications, a combination of decision trees and artificial neural networks can be adopted [5].

2.3.3. Association Rules. The association rule mainly refers to the rule characteristics of correlation in the values of two or more variables. There is generally an association between the data in the database, not in a single form. Correlation analysis is to discover the correlation characteristics between data through analysis, so as to obtain the dependence between data, which is convenient for future data design and analysis. The association rules are mainly composed of two stages: first, analyzing the data and obtaining the high-frequency names appearing in the dataset; secondly, performing the high-frequency names obtained in the previous step.

Applying the association rules to the personalized library management system can effectively help the library to quickly lock down the problems associated with its related issues when a problem occurs and can obtain the content of the current reader users based on the analysis of the reader's retrieval information. The process of mining information is to push the corresponding information to the reader more effectively.

2.4. Apriori Algorithm. The Apriori algorithm was proposed by R. Agrawal et al. in 1993. This algorithm is a classical algorithm for association rule mining. Many of the later algorithms are based on the idea of this algorithm [3]. The name of the algorithm is derived from the application in the algorithm. Any nonempty subset of frequent itemsets must meet the requirements, so as long as an itemset is infrequent, its superset does not need to be tested [6]. The flowchart of the first stage of the Apriori algorithm is shown in Figure 3 [7].

The Apriori algorithm uses a recursive search idea, which uses a candidate set to find frequent itemsets layer by layer, mainly through two steps of connection and pruning. The algorithm scans the database for the first time, finds all the frequent 1-itemsets, composes the frequent 2-sets by simple merging (joining) of the frequent 1-items, and then scans the database, which will support less than the minimum support. The itemset of degrees is deleted (pruned) from the candidate frequent 2-items, and the frequent 2-items are obtained. Then, the connection and pruning are used for the frequent 2-episodes, the frequent 3-episodes are found, and then iteration is performed until there are no frequent episodes higher than the minimum support [7]. The algorithm for mining frequent episodes ends, and find frequent itemsets to explore the content of the items. The Algorithm 1 is described as follows.

The next step in the algorithm is to mine association rules based on frequent itemsets. A rule with a confidence greater than the minimum confidence is called a frequent association rule. The algorithm mines all the association rules. These association rules may be frequent or infrequent. Then, based on the minimum confidence, the association rules greater than the minimum confidence are mined out to obtain the required frequent association rules.

3. University Library Personalized Service

3.1. Library Personalized Service Model. The library personalized service system based on association rule mining (as shown in Figure 4) mainly implements two functions:

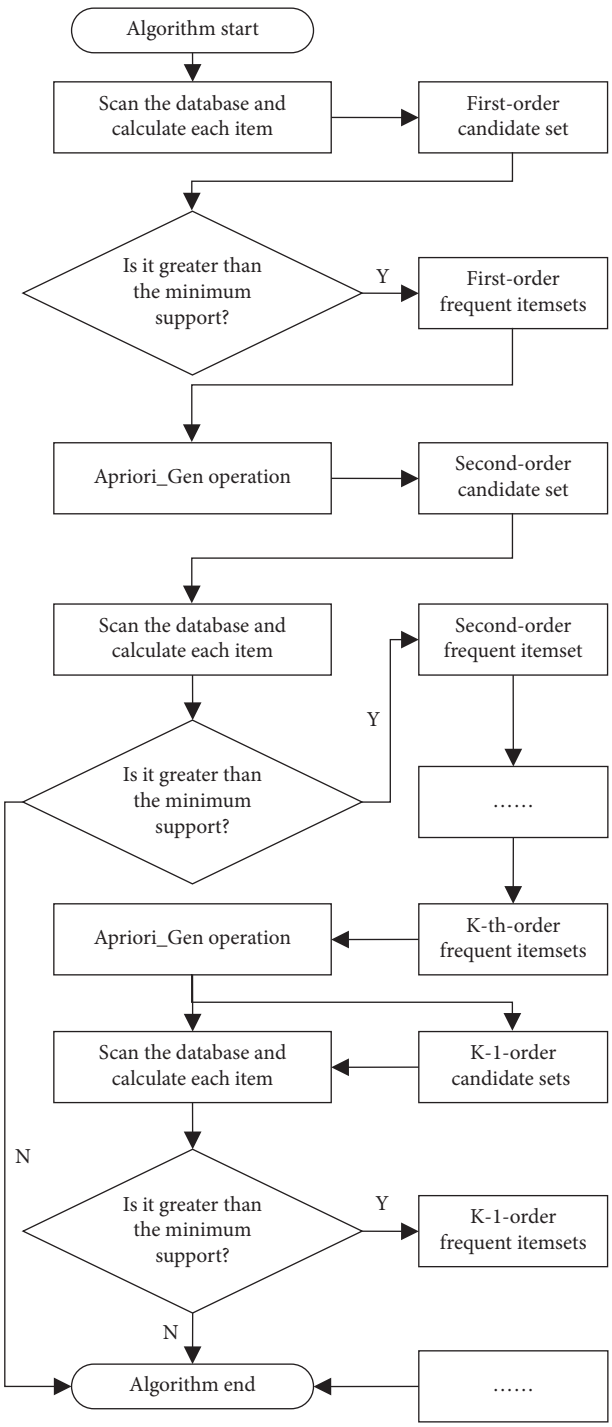


Figure 3: Flowchart of the first stage of the Apriori algorithm.

first, the association rule mining function, that is, the library readers borrow data to realize association rule mining, and find potential rules; personalized service function, which is to apply the generated association rules to the library personalized service [8]. The platform running on the system is based on the Windows Server 2003 operating system and adopts the B/S mode. The foreground uses Visual Studio 2005 integrated environment, and Visual C is the development tool; the server uses the background SQL Server 2005 database to save the user data; the data mining algorithm uses the Microsoft association algorithm [9]. The library personalized service system is shown in Figure 4.

The library personalized service system mainly includes three functional modules: data processing, association rule mining, and personalized service (not implemented). The system first performs data processing. The main functions include data import, data integration, data cleansing, data filtering, data conversion, and data reduction. This is a very important process that directly affects the efficiency of subsequent association rule mining. Then, according to the two tasks of association mining, the association model between the reader feature and the borrowed book and the association model borrowed by the reader are established. Finally, the mining association rules are applied to the reader personalized service.

3.2. The Way the Library Is Personalized. The library personalized service is a service that provides users with information resources and functions that meet their individual requirements according to the user's information usage behaviors, habits, hobbies, characteristics, and specific needs. It is a comprehensive consideration of the reader's individual. Features and special information need to provide readers with a personalized information environment [10].

According to whether the user actively provides the demand information [11], the library personalized information service mainly has two kinds of explicit feedback methods and implicit feedback methods. The explicit feedback and the implicit feedback are mainly based on whether the user needs to provide the demand. The difference and composition are shown in Figure 5.

3.3. Application of Association Rules in Book-Borrowing Data. The library management system is an indispensable part of the library management work. Its function is very important for the library administrators and users. Therefore, the library management system should be able to provide sufficient information and quickness for managers or readers. It is generally divided into the following subsystems: the book management subsystem, the book circulation subsystem, the reader management subsystem, and the reader query subsystem. Each subsystem contains several relational tables. Among them, the book circulation subsystem is one of the most important tasks of the library. It directly deals with the readers and deals with the readers' borrowing, book return, and renewal. Data mining in this section is the development of this part of the data [12].

The task of mining the circulation data of books by using association rules is mainly to find the regularity of the two aspects by analyzing the historical data of the readers [13]:

3.3.1. Discover the Characteristics of Readers and the Regularity between Them Borrowing Books. Investigate readers' different characteristics such as gender, age (grade), professional, and other aspects of their impact on borrowing books, and finally find out which features readers tend to

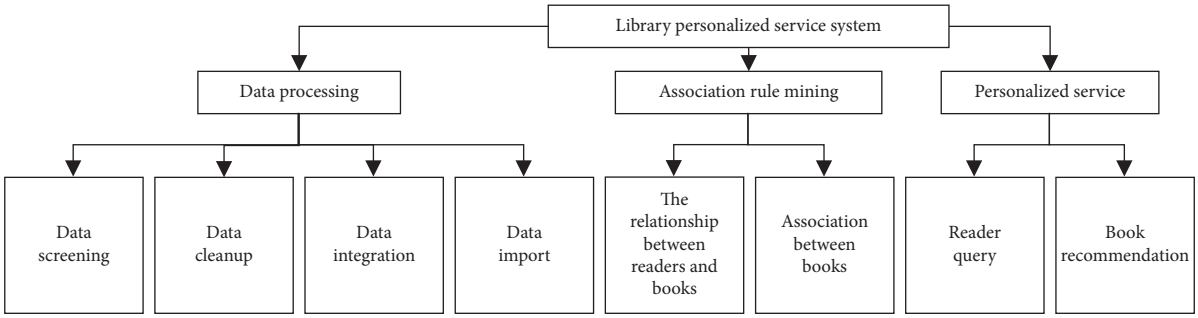


Figure 4: Library personalized service system.

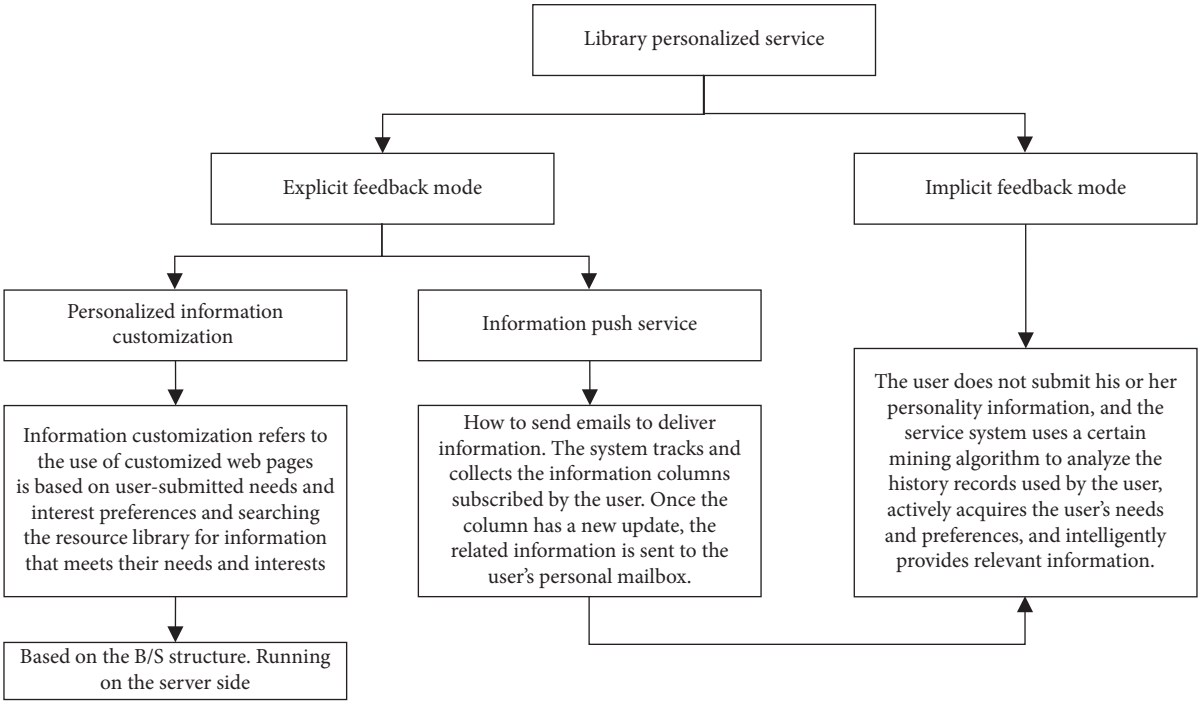


Figure 5: The browsing process of the library.

borrow books. This has a good guiding significance for future readers to borrow.

3.3.2. Discover the Association between Different Items in the Transaction Database, Reflecting the Reader's Borrowing Mode. For example, if 60% of readers borrowed book *A*, they would usually borrow book *B*. If they found the loan relationship between book *A* and book *B*, they could recommend book *B* to the reader who borrowed book *A*. Proper placement of classroom books can increase the number of loans or purchases.

The KDD process is shown in Figure 6 [14]. The KDD process can be summarized into three parts: data preprocessing, data mining, and interpretation and evaluation of results (interpretation and evaluation).

3.4. Microsoft Association Algorithm. The Microsoft correlation algorithm is very sensitive to the setting of parameters. If the parameters are not set properly, too many or too few

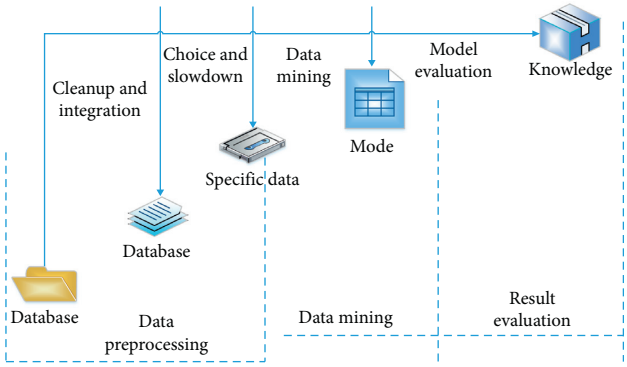


Figure 6: KDD process.

rules will be generated. It mainly involves the following three parameters.

3.4.1. Support. Support is used to describe the frequency of occurrence of an itemset [15], and its size affects the

generation of itemsets without affecting the generation of rules. The support for itemset $\{A, B\}$ is the total number of transactions that contain both A and B which is

$$\text{support}(A \Longrightarrow B) = p(A \cup B) = \text{number of transactions}(A, B). \tag{1}$$

3.4.2. Probability. It is also known in some literature studies as confidence or credibility. The probability that there will be B (i.e., A=>B) under the condition that there is A in a rule means

$$\begin{aligned}\text{probability}(A \Rightarrow B) &= \text{probability}(B|A)\\ &= \frac{p(A \cup B)}{p(A)}\\ &= \frac{\text{number of transactions}(A, B)}{\text{total number of transactions}}.\end{aligned} \tag{2}$$

Minimum probability means that the user is only interested in certain rules that reach the specified frequency. The setting of its value is the same as that of the minimum support (minimum support). Probability has no effect on the itemset, but has an effect on the formation of the rule. Specifying a certain minimum probability value limits the number of rules generated [12].

3.4.3. Importance. It is also referred to as interest or gain in some literature studies. It has an impact on the generation of itemsets and rules, the importance of itemsets and the importance of rules. The importance of an itemset is defined using the following formula:

$$\begin{aligned}\text{importance}(A \Rightarrow B) &= \frac{p(B|A)}{p(B)}\\ &= \left[\frac{\text{probability}(A, B)}{\text{probability}(A) * \text{probability}(B)}\right].\end{aligned} \tag{3}$$

It describes the magnitude of the influence of itemset A on itemset B. Its value range is $[0, \infty]$. If importance = I [16], it means that A and B are independent items, that is, purchase A and purchase B are two independent events; if import<1, it means that A and B are negatively correlated, that is, if a customer purchases A, then he purchases B which is unlikely to occur; if import> > 1, it means that A and B are positively related, that is, if a customer purchases A, he may also purchase B. The importance of the Bo rule is calculated using the following formula:

$$\text{importance}(A \Rightarrow B) = \frac{\log\ p(B|A)}{p(B|not\,A)}. \tag{4}$$

From the definition of equation (4), if the value is 0, it means that A and B have no relevance; positive values mean that when A is true, the probability of B will increase; negative value means that when A is true, the probability of B will decrease [17].

4. Analysis of Experimental Results

4.1. Mining the Association between Reader Characteristics and Borrowed Books. The above experiment is to mine the association rules based on the relationship between the reader's characteristics and the borrowed book class [18]. When the support degree is 0.1 and the confidence = 0.4, 186 rules are obtained. The experimental results are shown in Table 1. The association rule comparison diagram is shown in Figure 7.

By analyzing the above association rules, one can find the following rules:

(1) First-year (08) computer majors borrowed 10.3% of all computer-based readers for web design books, and 12.6% of all computer-based readers borrowed multimedia directions, in fact, in their second year. At the same time, the computer system reorganized the original three computer application classes into later computer application classes, web page orientation classes, and graphic image processing classes according to the student's interest orientation, which is consistent with the results obtained by this association mining.

(2) 56.2% of male students borrowed online books, and male students accounted for 68.5% of the entire computer department. The proportion is quite high, so it is possible to recommend online books to male students.

(3) Female students borrowed 14.8% and 15.6% of all computer science students in the web design and graphic image processing categories, so they can be considered when personalizing services.

The following minimum support degrees are 0.05, 0.1, 0.15, and 0.2, and the minimum confidence is 0.2, 0.3, 0.4, 0.5, and 0.6, which are Steps 5 and 6 in the above experiment, and the degree of confidence between the support and the rule number is obtained. The relationship table is shown in Table 2. The relationship between minimum support, minimum confidence, and rule number is shown in Figure 8.

Through the above experiments, it is found that choosing the appropriate minimum support and minimum confidence is the key to mining effective association rules [19]. The value will affect the number of export rules and the level of the concept layer. Library readers have a large amount of data to borrow, and it is impossible to predict how much support can filter out the appropriate data. Therefore, the minimum support and the minimum confidence threshold can be appropriately adjusted according to the actual number generated by the rule and the predetermined target to avoid excessive or too few rules. In addition, through the mining of association rules, it is found that the system is sensitive to support. When the support value is > 0.2, the rules cannot be mined [5].

```
 Step 1: L1=find_frequent_l-itemsets (D) // Mining frequently 1 set, scan the transaction database
   Step 2;) for (k=2;Lk-1=0;k++)
{
   Ck=apriori_gen (Lk-1,min-sup) //Call the apriori_gen method, generate
//Candidate frequent k itemsets
Step 3:
   for each transaction
      Ct=subset (Ck,t)
      for each candidate c
      c.count++
} //Scan the transaction database D
//Statistics count the number of candidate frequent k items
Step 4:
   Lk={c|c.count>+min-sup} //The k-item set that satisfies the minimum support is the frequency
}
return l=UkLk //Merge frequent k items set (k> 0)
```

Algorithm 1

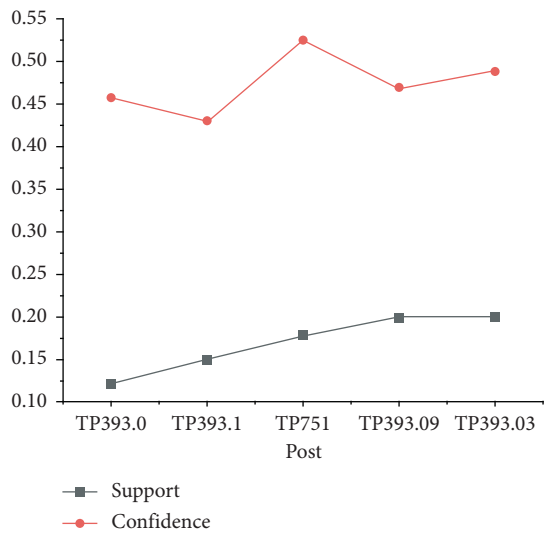


Figure 7: Association rule comparison diagram.

4.2. Mining the Association between Books. In the above experiment, one only needs to set the input column and predictable column to the book classification number and then adjust the algorithm parameters (support degree = 0.15; confidence level = 0.45). The above experiment process is repeated to get the reader to borrow books. There are 125 association rules between them. The experimental results are shown in Table 3:

The above rules are explained as follows.

The first rule: 15.2% of readers borrowed database theory and system and program language, algorithm language books. At the same time, borrowing database theory and system readers has a 48.7% chance of borrowing programming language and algorithm language books; the second rule: 4% of readers borrowed image processing software and text information for book processing at the same time, and readers of image processing software have a 56% chance to borrow text message processing books; the third rule: 15.8% of readers borrowed related machine-aided design (CAD), aided graphics, and image processing books [20].

The readers of the assisted graphics class have a 67.2% chance of borrowing image processing classes. The fourth rule: 16.2% of readers borrow computer security and network operating system books at the same time, while 52.4% of readers borrow computer network security. The fifth rule: there are 18.5% of the readers who also borrowed software maintenance and programming languages and a language-speaking book. At the same time, readers who borrowed the software maintenance class had a 47.3% chance of borrowing programming language and algorithmic language books.

Finally, comparing the rules derived from association mining with the actual work of the college library and the readers' book-borrowing survey, the results are relatively close, indicating that the data mining results of this system are effective. However, because the number of students in the computer department is relatively small compared to the students in the whole school, most students borrow books

Table 1: Association rules between reader characteristics and borrowed books.

Previous	Post	Support (%)	Confidence (%)
PROCESSIONAL = 1 and GRADE = 2	TP393.0	12.4	45.6
PROCESSIONAL = 1 and GRADE = 3	Tp393.1	15.2	42.9
PROCESSIONAL = 1 and R SEX = 1	TP751	18.1	52.2
R SEX = 0 |TP393.09|20.1 |46.8|	TP393.09	20.1	46.8
PROCESSIONAL = 2 and GRADE = 2	TP393.03	20.2	48.7
...	...	...	...

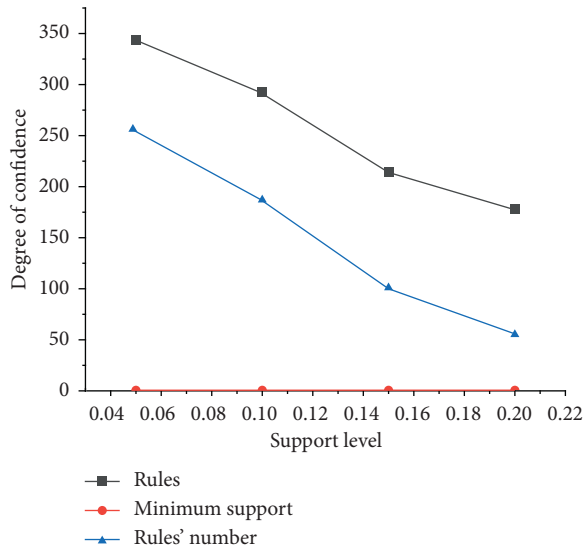


Figure 8: Relationship between minimum support, minimum confidence, and rule number.

Table 2: Relationship between minimum support, minimum confidence, and number of rules.

Minimum support	Minimum confidence	Rules' number	Minimum confidence	Rules' number
0.05	0.2	342	0.3	256
0.1	0.2	293	0.4	186
0.15	0.2	215	0.5	102
0.2	0.2	178	0.6	56

Table 3: Association rules between borrowed books.

Previous	Post	Support (%)	Confidence (%)
TP311.13	TP312	15.2	48.7
TP317.4	TP391.1	15.4	56.0
TP391.7	TP751	15.8	67.2
I'P393.08	TP316.8	16.2	52.4
CP311.53	TP312	18.5	47.3
...	...	...	...

according to their own majors, the number of books in the library is limited, and the book renewal period is relatively long. It also has some influence, which leads to some limitations of the excavated association rules.

5. Conclusion

Based on the data mining technology in the literature, this research studies how to use the data in the library management information database, uses the Apriori algorithm to mine data such as borrowing records, and finds the reader's relevance to the borrowing of documents. Different types of readers exist which are presented in the literature. The regularity of borrowing exists, and there is a certain connection between different disciplines. Excavating the relationship between these data, the librarian can purchase the book to provide service information, which is conducive to rational allocation of the library's literature resources and improves the utilization of resources, and promote a virtuous circle of book management. Taking the book management system as an example, we introduce the system structure and business process of the university library

management system and study how to build the data warehouse on this basis. Finally, we use the Apriori algorithm and the improved Apriori algorithm to mine the data such as borrowing records. There are a large number of borrowing records in the database of the library. We can mine the borrowing and reading data of the readers. It can be found that the readers have certain relevance to the borrowing of the documents. Different types of readers have certain rules for borrowing documents. There is also a certain kind of connection between different disciplines. We can analyze the relationship between readers and books in the borrowing record and discover the relationship between these data, which can provide the library administrator with the service information. It is conducive to rational allocation of library literature resources and improves the utilization of resources. At the same time, it provides some ideas for the application research of others in this aspect. The results of the proposed study show the effectiveness of the proposed study [20].

Data Availability

The datasets used and/or analyzed during the current study are available from the corresponding author upon reasonable request.

Disclosure

All authors agreed to submit this version and claimed that no part of this manuscript has been published or submitted elsewhere.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

[1] Y. Lie, "Data mining of university library management based on improved collaborative filtering association rules," *Wireless Personal Communications*, vol. 4, no. 11, 2018.

[2] M. Hahsler, "A model-based frequency constraint for mining associations from transaction data," *Data Mining and Knowledge Discovery*, vol. 13, pp. 137–166, 2006.

[3] A. Felty and S. Matwin, "Privacy-oriented data mining by proof checking," *Principles of Data Mining and Knowledge Discovery*, vol. 8, no. 2, pp. 138–149, 2002.

[4] P. Wang, "The application of data mining algorithm based on association rules in the analysis of football tactics," in *Proceedings of the International Conference on Robots & Intelligent System IEEE*, pp. 418–421, Deajeon, October 2016.

[5] L. Wang and H. Xu, "The application research of CRM in e-commerce based on association rule mining," *Advances in Intelligent and Soft Computing*, vol. 11, pp. 51–56, 2012.

[6] S. Song and Y. Ma, "The research and application of technology in the diagnosis of lung cancer warning association rule mining," *International Conference on Information Technology in Medicine and Education IEEE*, vol. 2, pp. 343–346, 2017.

[7] Y. Liu and X. Li, "Application of formal concept analysis in association rule mining," *International Conference on Information Science and Control Engineering IEEE Computer Society*, vol. 27, pp. 203–207, 2017.

[8] J. Wang, "Application of association rule mining algorithm in logistics information system design," *Journal of Software Engineering*, vol. 11, no. 2, pp. 217–223, 2017.

[9] Z. X. Wang, G. F. Ma, S. L. Ma, S. Y. Li, and Y. S. Cheng, "Mining association rules algorithm and its application based on key transactions," *Advanced Materials Research*, vol. 949, pp. 3387–3390, 2014.

[10] Z. Wang, "Research ON distributed parallel association rule mining," *Computer Applications & Software*, vol. 23, no. 1, p. 834, 2016.

[11] M. H. A. Haq, R. Saptono, and S. Palgunadi, "Association rule mining on library books lending data using apriori and jaccard similarity," *International Journal of Modern Trends in Engineering & Research*, vol. 5, no. 2, pp. 122–133, 2017.

[12] C. Ye, "The research on the application of association rules mining algorithm in network intrusion detection," *International Workshop on Education Technology and Computer Science IEEE*, vol. 5, pp. 849–852, 2009.

[13] J. Xue and S. Jarvis, "Mining association rules for admission control and service differentiation in e-commerce applications," *Wiley Interdisciplinary Reviews Data Mining & Knowledge Discovery*, vol. 8, no. 11, p. 1241, 2018.

[14] M. Li, "Investigation on application of association rule algorithm in English teaching logistics information," *Cluster Computing*, vol. 4, pp. 1–7, 2018.

[15] Y. Liu, "Data mining of university library management based on improved collaborative filtering association rules algorithm," *Wireless Personal Communications*, vol. 4, no. 11, pp. 1–10, 2019.

[16] J. Wang, "Application of association rule mining algorithm in logistics information system design," *Journal of Software Engineering*, vol. 11, no. 2, pp. 217–223, 2017.

[17] M. Meng, "Swarm oriented data organization and model analysis for libraries information service," in *Proceedings of the Eighth International Conference on Measuring Technology and Mechatronics Automation*, pp. 663–666, IEEE, Macau, China, March 2016.

[18] S. Zhang, J. Sun, and P. Wu, "Research on the fuzzy quantitative association rules," *Mining Algorithm and Its Simulation*, vol. 7, no. 1, pp. 99–118, 2013.

[19] X. Peng and Y. Wu, "Research and application of algorithm for mining positive and negative association rules," in *Proceedings of the International Conference on Electronic & Mechanical Engineering and Information Technology IEEE*, pp. 4429–4431, Harbin, Heilongjiang, China, August 2011.

[20] J. Chen, "Association rules mining algorithm based on interest measure and its application in medical audit," *Iet International Conference on Smart and Sustainable City IET*, vol. 4, pp. 278–281, 2014.

Hindawi
Scientific Programming
Volume 2021, Article ID 9913468, 9 pages
https://doi.org/10.1155/2021/9913468

Research Article

Design of a Cultural Tourism Passenger Flow Prediction Model in the Yangtze River Delta Based on Regression Analysis

Jian Xu

School of Economics and Management, Hefei University, Hefei 230601, Anhui, China

Correspondence should be addressed to Jian Xu; xujian@hfuu.edu.cn

Received 15 March 2021; Revised 5 April 2021; Accepted 15 April 2021; Published 24 April 2021

Academic Editor: Shah Nazir

Cultural tourism has gained much attention in the last decade and has promoted the preservation of a variety of tangible and intangible assets of culture. In order to accurately predict the cultural tourism passenger flow in the Yangtze River Delta and improve its economic benefits, this paper designs the prediction model of cultural tourism passenger flow in the Yangtze River Delta based on regression analysis. Taking the competitiveness of passenger flow as the core, this paper selects 28 indexes from four aspects of cultural tourism brand resources, cultural tourism support and protection, and urban tourism market income to build the evaluation index system of influencing factors of passenger flow. The principal component analysis method is used to simplify many related factors into a few uncorrelated factors to eliminate the multicollinearity caused by too many dependent variables; on this basis, the principal component regression model is constructed, and the determination coefficient is used to test the model fitting. Taking 15 cultural tourism cities in the Yangtze River Delta as the research object, the results show that the designed model has a good fitting degree, and the average error is only 0.41%, which can meet the needs of the prediction of cultural tourism passenger flow in the Yangtze River Delta. After the application of the prediction model, the foreign exchange earning amount of each cultural tourism city can be increased by more than 12%. The study has revealed good results.

1. Introduction

The improvement of living standards makes the tourism industry gradually rise. Generally speaking, tourism cities have developed commercial economy, and there are many tourist attractions and classic snacks, which have attracted a large number of tourists. The development and prosperity of cultural tourism in the near future will be due to the dual role of industry consciousness and system design [1]. Promoting the consciousness, self-confidence, and self-improvement of regional cultural construction is not only the main tone of current cultural construction but also the fundamental driving force for the rapid development of cultural tourism [2]. Different cities may have the same mode and path of tourism development, but the connotation of urban cultural tourism is unique and competitive. The tourism development mode dominated by cultural tourism often has more vitality and durability.

The Yangtze River Delta region is an important growth pole of China's economic development and also one of the regions with the highest level of urbanization in China [3, 4]. As an important strategic base of China's tourism industry, the Yangtze River Delta region is not only the main force and vanguard of China's tourism development but also the most successful region of China's cross regional tourism cooperation and the region with great potential to become a world-class tourism destination. In the process of implementing the national strategy of regional integration in the Yangtze River Delta, cultural tourism needs to play a greater role. The Yangtze River delta needs a complete image in the process of representing China to participate in global competition and cooperation. Cultural tourism is the advance force and core backbone of regional international image publicity. At present, regional integration has become a consensus of China's future development. Through resource sharing, complementary advantages, and market

interaction, we can break the regional, spatial, and institutional barriers, create barrier-free tourism areas, and realize the all-round development of social economy. At the same time, the Yangtze River Delta region is rich in tourism resources, with a complete range of natural and cultural landscapes. It is an important concentrated distribution area of tourism resources in China and has the inherent advantages of developing regional tourism integration. Therefore, it has attracted a large number of tourists to travel all the time.

Tourist flow refers to the number of people and the flow pattern of tourists from the source to the destination [5]. According to the different travel time, it can be divided into four types: daily passenger flow, monthly passenger flow, quarterly passenger flow, and annual passenger flow. Tourist flow is an important indicator of the development level of tourism industry, an important part of tourism planning by national or regional tourism authorities, an effective guarantee to improve the quality of tourism products, and an important basis for the development of tourism resources and the construction of reception facilities such as hotels. Accurate prediction of tourist flow is related to the successful operation of an international and regional tourism project, which will directly affect the scientific decision making of the tourism project and is an important part of urban tourism development planning [6, 7].

There are many factors that affect the cultural tourism passenger flow in the Yangtze River Delta, among which economy, politics, education level, resources, and transportation will have an impact on it [8]. Taking transportation as an example, the places with convenient transportation with the Yangtze River Delta will have a large tourist flow to the Yangtze River Delta. In addition, psychological factors and personal preferences will also have an impact on the tourist flow of the Yangtze River Delta. Therefore, many factors must be fully considered in the process of analyzing the tourist flow in the Yangtze River Delta.

Scientific and reasonable prediction of tourist flow is of great guiding significance for the efficient use of tourism resources and local economic development. It can also help the government to formulate tourism development planning and tourism emergency plan, improve the quality and level of tourism service, and then, improve the tourist satisfaction and sense of experience [9]. The interdisciplinary research on "tourism passenger flow" and "passenger flow prediction" has developed rapidly. It has gone deep into systems science, computer science and technology, and other disciplines and has derived a number of interdisciplinary themes.

At present, the methods of predicting tourist flow are mainly divided into quantitative prediction and qualitative prediction. Qualitative prediction is generally based on qualitative analysis combined with empirical judgment [10], with low prediction accuracy. Quantitative prediction is mainly to establish a quantitative prediction model through mathematical methods. It is a widely used prediction method with high prediction accuracy. There are three kinds of models for tourism passenger flow prediction by using quantitative prediction methods, namely, the permeability model, gravity model, and GM (1, 1) model. The penetration model is an intuitive model with strong subjectivity [11], which is based on the interviewees' willingness to visit as the main data, combined with the population base and carrying coefficient to make an intuitive inference of the passenger flow of tourist attractions. However, the model has its own regional bias due to the willingness to visit, so the permeability model is only used for interval estimation of the willingness to visit and the passenger flow. The gravity model is a commonly used international method to predict passenger flow under normal conditions [12]. However, in the actual prediction process, this method only considers a single or a small number of factors affecting passenger volume, which is lack of comprehensiveness, resulting in biased passenger volume prediction. In the practical application of the GM (1, 1) model, too many factors may be considered, resulting in multicollinearity among factors [13], and the regression coefficient cannot pass the significance test, and even the sign of some regression coefficients is inconsistent with the actual economic significance.

In view of the problems existing in the abovementioned models, a prediction model of cultural tourism passenger flow in the Yangtze River Delta based on regression analysis is designed on the basis of the gravity model and permeability model. In the parameter selection of the multiple regression model, the model comprehensively considers many factors that affect the cultural tourism passenger flow in the Yangtze River Delta, so compared with the gravity model and GM (1, 1) model, it considers the factors affecting passenger flow more comprehensively, overcomes the influence of multicollinearity, and improves the accuracy of passenger volume prediction.

The organization of the paper is as follows: the materials and methods of the paper are presented in Section 2 with details. Section 3 of this paper shows the results of the paper. The conclusion of this paper is given in Section 4.

2. Materials and Methods

The following sections briefly present the materials and methods used in this study.

2.1. Index Selection. Compared with other traditional types of industries or industries, cultural tourism has its own unique characteristics in the development connotation, development context, and social effects [14, 15]. First of all, in the world, the government vigorously promotes the cultural tourism industry, making it an increasingly important way to cultivate national cultural identity and national identity. Cultural tourism shoulders the sacred function of nurturing citizens. Secondly, the system structure and operation mechanism of cultural tourism embody the characteristics of the binary compound system of culture and tourism integration. The correct understanding of the concept of culture and the reasonable mining of tourism elements in the cultural field are the primary problems of cultural tourism transformation from resources to products, which provides an important reference for identifying the

elements of cultural tourism competitiveness. Thirdly, cultural tourism reflects the close interaction between the host and the guest. Cultural tourism space is the combination of geographical space, cultural space, and social space of cultural tourism destination. It is a field of interaction between the host and the guest with clear geographical space, which also defines a clear spatial scope for clarifying the system of cultural tourism competitiveness. Therefore, the development of theory and practice shows that the competitiveness of urban cultural tourism passenger flow in the Yangtze River Delta can carry out more scientific index screening and system construction.

Based on the theory of cultural tourism passenger flow competitiveness in the Yangtze River Delta region, the following three principles are considered and followed in the design process of the specific factor index:

(i) The combination of comprehensiveness and operability: this paper selects and analyzes the connotation of cultural tourism development, the structure of tourist flow competitiveness, and the influencing factors in the Yangtze River Delta, so as to present the overall situation of tourist flow competitiveness of cultural tourism in the Yangtze River Delta as far as possible. At the same time, the feasibility and reliability of the index data sources are fully considered. According to the different grades of the Yangtze River Delta, in the design of the index system, besides quoting some necessary total indexes reflecting the scale effect, the strength indexes are considered as much as possible.

(ii) Systematic and hierarchical: by fully combining the competitiveness of cultural tourism passenger flow in the Yangtze River Delta with the regional reality, the differences and emphases of cultural and tourism industry in different regions of the Yangtze River Delta are very different. In the evaluation, we should consider the problems from the actual situation of different regions as far as possible, so that the differences can be reflected in the evaluation index system. The hierarchy is reflected in that the competitiveness of cultural tourism passenger flow is mainly composed of three-level indexes. Besides the general goal and target decomposition level, it is explained and evaluated by element indexes [16].

(iii) The evaluation objectives and methods are consistent. This paper focuses on the theme of the evaluation of cultural tourism passenger flow competitiveness in the Yangtze River Delta region, designs the indexes around the center as far as possible, and always defines the direction and goal of the evaluation system. In terms of evaluation methods, after solving the basic problems through qualitative methods, we use scientific mathematical methods to calculate and screen the indexes needed by the research according to the quantitative relationship of each indicator and then carry out comprehensive measurement and evaluation.

Based on the abovementioned competitiveness theory and index system design principles, this paper decomposes the evaluation objectives of cultural tourism passenger flow competitiveness in the Yangtze River Delta region into four subobjectives: cultural tourism brand resources, cultural performance and creativity, cultural tourism support and protection, and urban tourism market income. At the same time, 28 quantitative indexes are designed to form the Yangtze River Delta region's cultural tourism. The evaluation index system of influencing factors of passenger flow is shown in Table 1.

2.2. Construction of the Prediction Model. Before the establishment of the model, the passenger volume is y, the sample size is n, and the observation value of the i-th index is $x_i\,(i = 1, 2, \ldots, 27)$.

2.2.1. Multicollinearity Diagnosis. The dependent variable y and the independent variable $x_1, x_2, \ldots, x_{27}$ are used to establish the regression model. SPSS statistical software is used to select the variables by the backward regression method, and p variables are set to enter the regression analysis model, which are $x_{s1}, x_{s2}, \ldots, x_{sp}\,(p \le n)$, called initialization variables.

From the regression results to find the variance expansion factor (VIF), if $VIF_i \ge 10$, it means that there is a serious multicollinearity between the independent variables.

Multicollinearity has a great influence on the regression coefficient [17], which can be processed by principal component analysis.

2.2.2. Principal Component Analysis.

(1) Data standardization

In order to eliminate the influence of different orders of magnitude and dimensions, it is necessary to standardize the original data. The standardized formula is as follows:

$$x_{ij}' = \frac{x_{ij} - \overline{x}_j}{s_j}, \tag{1}$$

where x_{ij}' is the standardized data and $\overline{x}_j$ and $(j = 1, 2, \ldots, n)$ represent the mean value and standard deviation of the j-th index sample, respectively.

(2) Calculation of the correlation coefficient matrix: after processing the original data, the standardized data matrix $(x_{ij}')_{p\times n}$ is obtained, and the corresponding correlation coefficient matrix is calculated.

$$R = \left(r_{ij}\right)_{n\times n}. \tag{2}$$

In formula (2), R is a symmetric matrix of order n.

(3) The eigenvalues and eigenvectors of correlation coefficient matrix R are calculated.

The eigenvalue $\lambda_i\,(i = 1, 2, \ldots, n)$ of R and its corresponding eigenvector $u_i\,(i = 1, 2, \ldots, n)$ are solved,

Table 1: Evaluation index system of influencing factors of cultural tourism passenger flow in the Yangtze River Delta.

General objective	Target decomposition	Element description
Influencing factors of cultural tourism passenger flow in the Yangtze River Delta	Brand resources of cultural tourism	Number of world cultural heritages
		Number/item of abnormal intangible culture at the national level or above
		Number of national key cultural relics protection units
		Number of famous historical and cultural cities in China
		Number of famous historical and cultural towns and villages in China
		Number of 4A tourist attractions
		Number of red tourist attractions
	Cultural performance and creative tourism	Number of local folk festivals and special activities
		Creative tourism performance project
		Number of provincial and national cultural industry demonstration bases
		Number of art performing groups
		Number of cultural centers, art performance centers, and mass art centers
		Number of libraries
		Number of open museums
	Cultural tourism support	Number of accommodation facilities
		Number of beds
		Number of rooms
		Number of travel agencies
		Number of food blocks
		Number of passengers sent by civil aviation
		Scale of highway passenger transport
		The scale of railway passengers
	Regional tourism market revenue	Number of inbound tourists
		Domestic tourists
		Foreign exchange earned by tourism
		Total tourism revenue
		The proportion of total tourism revenue in total economic output

and $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$, where λ_i is the variance of the main component F_i, and the greater the variance is, the greater the contribution to the total variance is [18].

(4) The contribution rate is calculated, and the principal component is determined.

Formula (3) is defined as the contribution rate of main component F_i.

$$e_i = \frac{\lambda_i}{\sum_{i=1}^{n} \lambda_i} \times 100\%. \tag{3}$$

In formula (3), $\sum_{i=1}^{n} e_i$ is the contribution rate of cumulative variance. Generally, m principal components with $\sum_{i=1}^{n} e_i$ greater than or equal to 85% are selected for comprehensive analysis. Therefore, n factors are reduced to m principal components, and the main factors are selected [19].

2.2.3. Principal Component Regression Model. Multiple regression analysis is conducted between F_r and dependent variable Y (standardized value of y) to obtain the standardized regression equation.

$$Y = \sum_{r=1}^{m} B_r F_r, \tag{4}$$

In formula (4), B_r is the standardized partial regression coefficient of the r-th principal component F_r.

By synthesizing formula (2), (3), and (4), the standardized regression equation is obtained.

$$Y = \sum_{j=1}^{p} b_j' X_{s_j}. \tag{5}$$

In formula (5), b_j' is the j-th standardized partial regression coefficient of the standardized regression equation.

Through the abovementioned analysis, we can get the principal component regression model as follows:

$$y = \sum_{j=1}^{p} b_j x_{s_j} + b_0, \tag{6}$$

where

$$b_j = b'_j \sqrt{\frac{L_{yy}}{L_{x_{ij}x_{ij}}}}, \tag{7}$$

$$b_0 = \overline{y} - \sum_{j=1}^{p} b_j \overline{x}_{s_j}. \tag{8}$$

In the abovementioned formula, b_j is the j-th partial regression coefficient of the general linear regression equation; L_{yy} is the sum of squares of deviation of y; $L_{x_{ij}}$ is the sum of squares of deviation of x_{sj}; $\overline{y}$ is the mean value of y; $\overline{x}_{s_j}$ is the mean value of x_{sj}; and b_0 is the constant of the general linear regression equation.

By substituting the observed values of each index in the forecast year into formula (6), the cultural tourism passenger flow of the Yangtze River Delta in the forecast year can be obtained.

2.3. Model Test. The determination coefficient R^2 is used to test the model fitting. The formula for determining coefficient R^2 is as follows:

$$R^2 = 1 - \frac{\sum e_t^2}{\sum (y_t - \overline{y})^2}. \tag{9}$$

In formula (9), $\sum e_t^2$ is the sum of squares of residuals.

In the multiple linear regression model, the number of variables in each regression model may not be the same [20]. It is not appropriate to use the size of R^2 as a measure of fitting quality. Therefore, the coefficient of determination R^2 of modified degrees of freedom is often used. The calculation formula is as follows:

$$\overline{R}^2 = 1 - \frac{\left(\left(\sum e_t^2\right)/n - p\right)}{\left(\left(\sum (y_t - \overline{y})^2\right)/n - 1\right)} = 1 - \frac{n-1}{n-p}\left(1 - R^2\right). \tag{10}$$

In formula (10), n is the sample size, and p is the number of regression coefficients.

3. Results

3.1. Case Profile and Data Sources. The Yangtze River Delta is an urban sprawling area with a high level of tourism development in China. It has rich cultural resources and profound cultural heritage. In view of the different regional coverage of the Yangtze River Delta, this paper takes the core area of the Yangtze River Delta as the research object in the reply of the State Council on the regional planning of the Yangtze River Delta (Guo Han (2010), No. 38), that is, taking Shanghai as the leader, and Nanjing, Suzhou, Wuxi, Changzhou, Zhenjiang, Nantong, Yangzhou, and Taizhou in Jiangsu Province, Hangzhou, Ningbo, and Huzhou in Zhejiang Province, and Jiaxing, Shaoxing, Zhoushan, and Taizhou, a total of 16 cities as both wings. This paper selects the 10-year statistical data from 2008 to 2018 as an example, and the data come from the statistical yearbooks and statistical bulletins of various cities, including the China Urban Statistical Yearbook, China Regional Economic Statistical Yearbook, China Tourism Statistical Yearbook, China Tourism Yearbook, List of Top 100 Travel Agencies, and List of National Star Hotels. Data collection adheres to the principle of combining scientificity, authority, standardization, and data availability. When the data from different channels are not unified, the higher-level government department shall prevail.

3.2. Data Processing. Since the data collected are from the statistical yearbooks and statistical bulletins of various cities, the time span is from 2008 to 2018. Due to the snow disaster, earthquake, financial crisis, and other major events in 2008, the cultural tourism reception in the Yangtze River Delta region is seriously affected, and the number of tourists decreased significantly. In order to improve the accuracy of the prediction model, the data of cultural tourism passenger flow in the Yangtze River Delta in 2008 and 2009 are revised.

Linear interpolation is used to correct the data.

Firstly, the starting year a_1 and the ending year a_2 which are suitable for linear interpolation are selected, and the passenger flow in the starting year and the ending year is expressed by y_1 and y_2, respectively. The tolerance d is determined by formula (11).

$$d = \frac{y_2 - y_1}{a_2 - a_1}. \tag{11}$$

The correction value is calculated by the interpolation equation, and the calculation formula is as follows:

$$y_n = y_1 + (a - a_1)d. \tag{12}$$

In formula (12), n is the year to be corrected and y_n is the correction value of the n-th year.

According to the abovementioned method, the starting year of linear interpolation is 2010, and the ending year is 2013. By substituting the data of these two years into formula (11) and formula (12), we can get the revised value of cultural tourism passenger flow in the Yangtze River Delta in 2008 and 2009. Table 2 shows the revised tourist flow of cultural tourism in the Yangtze River Delta from 2008 to 2018.

3.3. Multicollinearity Diagnosis. Taking the statistical data of the Yangtze River Delta from 2008 to 2018 as an example, this paper makes regression analysis on the cultural tourism passenger flow and 28 influencing factors in the Yangtze River Delta and uses SPSS statistical software to select variables by the backward regression method. Finally, there are 9 variables X_7, X_8, X_{15}, X_{16}, X_{17}, X_{18}, X_{20}, X_{22}, and X_{23} to enter into the regression model, and the VIF values are greater than 10, which indicates that there is still a serious collinear relationship between variables. Therefore, it is necessary to use principal component regression to simplify the analysis. On the premise of retaining all or most of the original information, the abovementioned interrelated variables are transformed into a few independent or

Table 2: Revised tourist flow of cultural tourism in the Yangtze River Delta from 2008 to 2018.

Year/year	Passenger flow/10000 person times
2008	1090.09
2009	1110.62
2010	1128.60
2011	1175.27
2012	1241.94
2013	1268.52
2014	1364.39
2015	1392.41
2016	1379.66
2017	1485.59
2018	1590.33

unrelated variables, and then, these variables are integrated to establish a regression model.

Principal component analysis is performed on X_7, X_8, X_{15}, X_{16}, X_{17}, X_{18}, X_{20}, X_{22}, and X_{23} variables. Using SPSS statistical software, the eigenvalues and eigenvectors are obtained. The cumulative variance contribution rate of the first two eigenvalues has reached 91.041%. It is generally believed that the effective information can be retained when the cumulative contribution rate of principal components reaches 85%. Therefore, this paper only needs to take the first two principal components to reflect most of the information of all indexes:

$$F_1 = 0.149X_7 + 0.180X_8 + 0.227X_{15} - 0.217X_{16} - 0.204X_{17} + 0.199X_{18} + 0.179X_{20} - 0.179X_{22} - 0.210X_{23}, \tag{13}$$

$$F_2 = 0.455X_7 - 0.359X_8 + 0.060X_{15} + 0.137X_{16} + 0.245X_{17} + 0.219X_{18} + 0.342X_{20} + 0.289X_{22} - 0.059X_{23}. \tag{14}$$

The standardized regression equation is obtained by multiple regression analysis between the evaluation values of F_1 and F_2 and the dependent variable Y.

$$Y = 0.973F_1 + 0.113F_2\left(R^2 = 0.972\right). \tag{15}$$

By substituting formula (10) and (13) into (14), the standardized regression equations for the standardized variables X_7, X_8, X_{15}, X_{16}, $X_{17}X_{18}$, X_{20}, X_{22}, and X_{23} are obtained.

$$Y = 0.204X_7 + 0.138X_8 + 0.230X_{15} - 0.200X_{16} - 0.159X_{17} + 0.219X_{18} + 0.220X_{20} - 0.142X_{22} - 0.215X_{23}. \tag{16}$$

Using formula (7), the principal component regression model is obtained as follows:

$$y = 1596.733x_7 + 964.461x_8 + 5.013x_{15} - 576.898x_{16} \quad 529.266x_{17} + 0.164x_{18} + 7.999x_{20} - 392158.897x_{22} - 1478326.246x_{23} - 158927.732. \tag{17}$$

3.4. Model Validation. Through the test formula of formula (8) and (9), it can get $R^2 = 0.984$; $\overline{R}^2 = 0.859$, showing that formula (17) has a high degree of fit and can make a reasonable forecast of passenger volume. By substituting the values of X_7, X_8, $X_{15}X_{16}$, X_{17}, X_{18}, X_{20}, X_{22}, and X_{23} from 2008 to 2018 into formula (17), the predicted passenger volume of each year can be obtained. The comparison between the predicted value and the actual value (Table 2) is shown in Figure 1.

It can also be seen from Figure 1 that the predicted value of the principal component regression model has a good fit with the actual value, with the highest error of 1.23%, the lowest error of 0.01%, and the average error of 10 years is only 0.41%, which can basically meet the needs of the prediction of cultural tourism passenger flow in the Yangtze River Delta.

3.5. Model Performance Analysis. In order to analyze the application performance of the model, this paper uses the method to predict the cultural tourism passenger flow in the Yangtze River Delta (Shanghai as the leader, Nanjing, Suzhou, Wuxi, Changzhou, Zhenjiang, Yangzhou, and Taizhou in Jiangsu Province and Hangzhou, Ningbo, Huzhou, Jiaxing, Shaoxing, Zhoushan, and Taizhou in Zhejiang Province, a total of 15 cities as both wings) in 2020. MAPE (average absolute percentage error, which can directly reflect the pros and cons of the prediction effect), MAE (average absolute error, the smaller the value is, the smaller the error is), RMSE (root mean square error, the smaller the value is, the smaller the error is), and EC (equalization coefficient, the higher the value is, the higher the fitting degree is) are taken as evaluation indexes to verify the performance of the model. The results are shown in Table 3. The calculation formula of each evaluation index is as follows:

$$\mathrm{MAPE} = \frac{1}{V}\sum_t \left|\frac{C_p(t) - C_r(t)}{C_r(t)}\right| \times 100\%, \tag{18}$$

$$\mathrm{MAE} = \frac{1}{V}\sum_t \left|C_p(t) - C_r(t)\right|, \tag{19}$$

$$\mathrm{RMSE} = \sqrt{\frac{\sum_t \left(C_p(t) - C_r(t)\right)^2}{V}}, \tag{20}$$

$$\mathrm{EC} = 1 - \frac{\sqrt{\sum_t \left(C_p(t) - C_r(t)\right)^2}}{\sqrt{\sum_t \left(C_p(t)\right)^2} + \sqrt{\sum_t \left(C_r(t)\right)^2}}. \tag{21}$$

In the abovementioned formula, $C_p(t)$, $C_r(t)$, and V are the predicted output value, the measured value of cultural tourism passenger flow of each city in 2020, and the number of predicted samples, respectively. The prediction accuracy of the proposed model and the comparison model (the prediction model based on the gravity model and the prediction model based on GM (1, 1)) are compared fairly to the maximum extent by using the abovementioned indexes. Prediction results of this model are shown in Table 3.

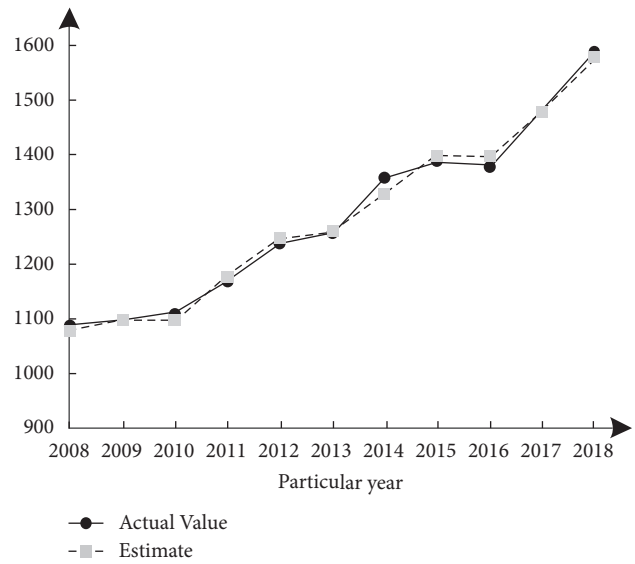


Figure 1: Comparison between the predicted value and actual value.

Table 3: Prediction results of this model.

Cultural tourism cities in the Yangtze River Delta	MAPE	MAE	RMSE	EC
Shanghai	0.57	1.26	2.05	0.97
Nanjing	0.60	1.14	2.05	0.98
Suzhou	0.62	1.30	2.01	0.98
Wuxi	0.35	1.06	1.65	0.98
Changzhou	0.51	1.10	1.79	0.98
Zhenjiang	0.50	1.05	1.81	0.99
Yangzhou	0.40	1.02	1.69	0.98
Taizhou	0.37	1.05	1.77	0.98
Hanzghou	0.69	1.14	1.96	0.97
Ningbo	0.43	1.12	1.72	0.98
Huzhou	0.49	1.11	1.82	0.99
Jiaxing	0.39	1.09	1.93	0.97
Shaoxing	0.44	1.13	2.00	0.98
Zhoushan	0.51	1.06	1.91	0.98
Taizhou	0.56	1.11	2.02	0.98
Average value	0.495	1.116	1.879	0.979

The average values of each index obtained by the algorithm in Table 3 are compared with those obtained by the two comparative prediction models, and the results are shown in Table 4.

After analyzing the prediction performance of the three prediction models in Table 4, it is found that the evaluation indexes of the proposed model are significantly better than those of the two comparison models, and the MAPE is reduced by more than 50 compared with the comparison model, which indicates that this model has higher prediction accuracy compared with the comparison model.

The running time (the larger the value is, the higher the complexity of the method is) is used as the evaluation index of prediction performance. The running time of this model and the two comparison models in the evaluation process is compared, and the results are shown in Table 5.

Analysis of Table 5 shows that the running time of the proposed model and the GM (1, 1) model is significantly better than that of the gravity model. The average running time of the proposed model is 0.71 s, and the average running time of the GM (1, 1) model is 0.74 s. There is no significant difference between the two models. Combined with the data in Table 4, it can be concluded that the proposed model has a significant performance advantage compared with the comparison model.

3.6. Application Test. This paper compares the amount of foreign exchange earned by cultural tourism cities in the Yangtze River Delta after using this model to predict the passenger flow with the amount of foreign exchange earned before using this model (the average of the previous three years is the standard value). The results are shown in Figure 2, in which the numbers 1–15 represent the cultural tourism cities in the Yangtze River Delta.

According to the analysis of Figure 2, after using the proposed model to predict the passenger flow, the amount of

Table 4: Comparison results of prediction performance of different prediction models.

Prediction model	MAPE	MAE	RMSE	EC
Model of this paper	0.495	1.116	1.879	0.979
Prediction model based on the gravity model	1.326	2.735	2.789	0.962
Prediction model based on the GM (1, 1) model	0.866	2.003	2.104	0.941

Table 5: Running time of different prediction models.

Cultural tourism cities in the Yangtze River Delta	Model of this paper	Prediction model based on the gravity model	Prediction model based on the GM (1,1) model
Shanghai	0.71	1.05	0.58
Nanjing	0.74	1.16	0.92
Suzhou	0.77	1.13	0.84
Wuxi	0.70	1.08	0.73
Changzhou	0.66	1.15	0.81
Zhenjiang	0.69	1.02	0.69
Yangzhou	0.72	0.99	0.75
Taizhou	0.73	1.15	0.68
Hanzghou	0.69	1.03	0.75
Ningbo	0.71	1.01	0.72
Huzhou	0.70	0.97	0.67
Jiaxing	0.68	1.08	0.73
Shaoxing	0.68	1.04	0.64
Zhoushan	0.71	1.03	0.79
Taizhou	0.72	1.10	0.80

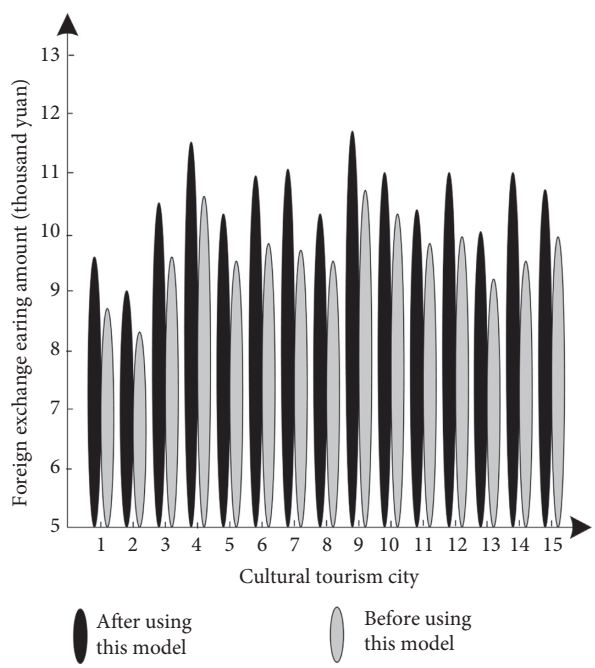


Figure 2: System sales comparison.

foreign exchange earnings of each city shows different increases, and the increase rate is maintained at more than 12%, which shows that the economic benefits of cultural tourism in the Yangtze River Delta can be significantly improved by using this model.

4. Conclusions

Cultural tourism has gained much attention in the last decade and has promoted the preservation of a variety of tangible and intangible assets of culture. The applications of cultural assets for development of tourism has generated various debates such as the matter of whether the intangible values of cultural assets including those of their education, aesthetics, and history can properly be carried out in order to attract tourist. In order to accurately predict the cultural tourism passenger flow in the Yangtze River Delta and improve its economic benefits, this paper designs the prediction model of cultural tourism passenger flow in the Yangtze River Delta based on regression analysis. This paper considers many factors that affect the cultural tourism passenger flow and designs a prediction model of cultural tourism passenger flow in the Yangtze River Delta based on regression analysis. Through the experimental analysis, this model can accurately predict the cultural tourism passenger flow in the Yangtze River Delta in the process of practical application, and significantly improve the cultural tourism amount of foreign exchange in the Yangtze River Delta.

Data Availability

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

Conflicts of Interest

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

Acknowledgments

This work was supported by University Humanities and Social Sciences Research Project of Anhui Province "Research on the high quality development mode of culture and tourism in the process of Yangtze River Delta integration," Project No. SK2020A0431..

References

[1] X. Wang, E. Yao, and S. Liu, "Urban rail transit passenger flow forecastingfor large special event based on afc data," *Beijing Jiaotong Daxue Xuebao*, vol. 42, no. 1, pp. 87–93, 2018.

[2] C. Lin, K. Wang, D. Wu, and B. Gong, "Passenger flow prediction based on land use around metro stations: a case study," *Sustainability*, vol. 12, no. 17, p. 6844, 2020.

[3] X. Liu, L. Li, X. Liu, and T. Zhang, "Analysis of passenger flow and its influences on hvac systems: an agent based simulation in a Chinese hub airport terminal," *Building and Environment*, vol. 154, pp. 55–67, 2019.

[4] K. Mbbchir, R. Statistics, C. G. Hons, and M. Alain Vuylstekec, "Prediction of patient length of stay on the intensive care unit following cardiac surgery: a logistic regression analysis based on the cardiac operative mortality risk calculator, euroscore," *Journal of Cardiothoracic and Vascular Anesthesia*, vol. 32, no. 6, pp. 2676–2682, 2018.

[5] Y. Liu, C. Lyu, X. Liu, and Z. Liu, "Automatic feature engineering for bus passenger flow prediction based on modular convolutional neural network," *IEEE Transactions on Intelligent Transportation Systems*, vol. 99, pp. 1–10, 2020.

[6] W. Zhou, W. Wang, and D. Zhao, "Passenger flow forecasting in metro transfer station based on the combination of singular spectrum analysis and adaboost-weighted extreme learning machine," *Sensors*, vol. 20, no. 12, p. 3555, 2020.

[7] S. Y. Song, "Analysis and prediction of passenger flow differences in shanghai 3a scenic spots and above based on arima model," *Statistics and Applications*, vol. 08, no. 3, pp. 537–552, 2019.

[8] H. Chen, B. Wang, W. He, and J. Zheng, "Research on passenger flow early warning of urban rail transit station based on system dynamics," *MATEC Web of Conferences*, vol. 308, no. 2, p. 01003, 2020.

[9] C. Li, "Combined forecasting of civil aviation passenger volume based on arima-regression," *International Journal of System Assurance Engineering and Management*, vol. 10, no. 5, pp. 945–952, 2019.

[10] F. Wang, H. Y. Liu, C. F. Shao, and J. J. Zhang, "Prediction model of public transport vehicle allocation based on multiple regression analysis," *Advances in Transportation Studies*, vol. 3, pp. 69–78, 2018.

[11] Y. Liu, C. Liu, and Z. Zheng, "Traffic congestion and duration prediction model based on regression analysis and survival analysis," *Open Journal of Business and Management*, vol. 08, no. 02, pp. 943–959, 2020.

[12] L. Z. Jing, Q. S. Li, J. L. Xu, X. L. Jia, and Y. J. Han, "Average travel time prediction model in basic expressway sections based on v/c ratio and truck percentage," *Chang'an Daxue Xuebao (Ziran Kexue Ban)/Journal of Chang'an University (Natural Science Edition)*, vol. 38, no. 5, pp. 106–113, 2018.

[13] T. Wu, P. Zhang, J. Qin, D. Wu, and Y. Wan, "A flood-discharge-based spatio-temporal diffusion method for multi-target traffic hotness construction from trajectory data," *IEEE Access*, vol. 99, pp. 1-2, 2020.

[14] L. Liu, R.-C. Chen, and S. Zhu, "Impacts of weather on short-term metro passenger flow forecasting using a deep lstm neural network," *Applied Sciences*, vol. 10, no. 8, p. 2962, 2020.

[15] T.-Z. Chen, Y.-Y. Chen, and J.-H. Lai, "Estimating bus cross-sectional flow based on machine learning algorithm combined with wi-fi probe technology," *Sensors*, vol. 21, no. 3, p. 844, 2021.

[16] X. Chang, J. Wu, H. Sun, G. Wang, and X. Bao, "Understanding and predicting short-term passenger flow of station-free shared bike: a spatiotemporal deep learning approach," *IEEE Intelligent Transportation Systems Magazine*, vol. 1, no. 99, pp. 1–3, 2021.

[17] D. Li, L. Deng, and Z. Cai, "Statistical analysis of tourist flow in tourist spots based on big data platform and da-hkrvm algorithms," *Personal and Ubiquitous Computing*, vol. 24, no. 1, pp. 87–101, 2020.

[18] H. J. Heng and P. Ren, "Short-Term passenger flow forecasting for the check-in process in the airport based on time series," *Computer Simulation*, vol. 37, no. 02, pp. 31–37, 2020.

[19] W. Wang, Y. Wang, G. Correia, and Y. Chen, "A network-based model of passenger transfer flow between bus and metro: an application to the public transport system of Beijing," *Journal of Advanced Transportation*, vol. 2020, no. 15, 4 pages, Article ID 6659931, 2020.

[20] L. Qiu, H. Zhou, Z. Wang, S. Zhang, L. Zhang, and W. Lou, "High-speed elevator car air pressure compensation method based on coupling analysis of internal and external flow fields," *Applied Sciences*, vol. 11, no. 4, p. 1700, 2021.

Hindawi
Scientific Programming
Volume 2021, Article ID 9994965, 10 pages
https://doi.org/10.1155/2021/9994965

Research Article

Research on E-Commerce Purchasing Model in Crude Oil Trade

Yunfei Li

Xijing University, Xi'an 710123, China

Correspondence should be addressed to Yunfei Li; 20130044@xijing.edu.cn

Received 24 March 2021; Revised 9 April 2021; Accepted 20 April 2021; Published 24 April 2021

Academic Editor: Shah Nazir

Crude oil is the main raw material for the petrochemical industry and a major component of total production costs. It is also an indispensable resource for petrochemical enterprises. The implementation of crude oil e-commerce procurement by petrochemical enterprises can not only reduce the procurement cost of crude oil but also create good economic benefits in the field. It can also establish strategic partnerships with oil field enterprises to lay the foundation for the scientific development of enterprises. Based on the production of Company A and the "10 Million Tons Refining" project, this paper introduces the concept of e-commerce procurement for the problems in all aspects of crude oil procurement management. This paper deeply studies the e-commerce procurement theory, constructs the e-commerce procurement management implementation framework, and proposes that the scientific evaluation of crude oil suppliers is the key idea for petrochemical enterprises to implement enterprise e-commerce procurement management. At the same time, it also conducted an in-depth analysis of the internal and external environment of a company's crude oil e-commerce procurement, conducted a comprehensive evaluation of existing crude oil suppliers and potential suppliers, and formulated corresponding procurement strategies. Results of the study show the success of the proposed study, and on the basis of the proposed study, researchers can devise new solutions in the field.

1. Introduction

The development of computer networks and communication technologies has made the processing and transmission of information technology break through the limitations of time and region. The network and globalization of the economy have become an inevitable trend in the development of the world economy. The new economic model has challenged the traditional business model, and the emergence of e-commerce has adapted to the rapid development of the requirements of the new economic development [1]. E-commerce not only provides enterprises with new development space but also provides enterprises with new ways to enhance their competitive advantage. E-commerce will become the main form of trade in the era of the knowledge economy. If enterprises do not seize the opportunity to develop e-commerce, they will be eliminated by the market, and it is imperative for enterprises to develop an e-commerce [2].

International research on e-commerce began in the late 2000s and was only booming at the end of the century. As for the concept of e-commerce, people have different opinions depending on the angle and field of the station, and there is no exact definition [3]. Generally speaking, e-commerce is an electronic transaction, that is, conducting business activities online, and is a commercial trade activity conducted by people through the Internet. When an enterprise directly connects its main business through intranets, extranets, and employees, customers, vendors, and partners, the various activities that take place are e-commerce.

E-commerce in China began in 1997 and developed rapidly. According to the monitoring data of China Electronic Commerce Research Center, the transaction scale of China's electronic commerce market reached 1.02 trillion yuan in 2013, accounting for 17.3% of GDP. China's e-commerce has been developing at a rapid speed. More and more industries and enterprises realize that only by speeding up the development of e-commerce can they keep up with the pace of industry development. E-commerce has become an indispensable business channel for modern enterprises, even the commanding heights of enterprise competition. In the world petroleum and petrochemical

market competition, more and more new technologies such as e-commerce platforms have been used to reduce transaction costs and improve customer service. The main competitors of foreign petroleum and petrochemical enterprises actively penetrate the domestic market and directly participate in the domestic competition, which will bring pressure to the petroleum and petrochemical enterprises of China. If the petroleum and petrochemical enterprises of China do not adopt the advanced e-commerce management mode in time, it will be difficult for them to participate in the world competition. However, the e-commerce system currently used by petroleum and petrochemical enterprises is still not perfect: poor security of online payment, imperfect logistics system, relative lack of required compound talents, weak marketing system, etc., will restrict the application of e-commerce in petroleum and petrochemical industry to play its greatest strength. Therefore, it is necessary to constantly improve the electronic commerce system in order to ensure the steady development of the petroleum and petrochemical industry.

According to the definition of business model, you can participate in the main business and individual consumers as two dimensions according to business activities and divide the e-commerce business model into the following four categories: B2B, B2C, C2B, and C2C. In a traditional supply chain, a product manufacturer often connects several levels of supplier nodes. These suppliers mainly include raw material suppliers, outsourced manufacturers, and in-house manufacturers, which are more complex supply chain networks downstream of the supply chain, as shown on the left in Figure 1.

Based on this, this article takes Company A as the research object and conducts a series of researches around e-commerce. Company A is located in Jilin City, Jilin Province. It is the country's first large-scale chemical industrial base built during the "First Five-Year Plan" period. This paper first analyzes the supply and demand situation of crude oil procurement resources of Company A. Then, it deeply analyzes the various drawbacks of the original traditional procurement model of Company A, compares the economic benefits brought by the use of e-commerce procurement, and discusses the inevitability of Company A's e-commerce model for crude oil procurement. Finally, combined with the characteristics of crude oil procurement of petrochemical enterprises, this paper constructs an evaluation index system for crude oil strategic procurement plan of petrochemical enterprises according to the principle established by the evaluation index system.

The contribution of the paper is to deeply study the e-commerce procurement theory, construct the e-commerce procurement management implementation framework, and propose that the scientific evaluation of crude oil suppliers is a vital idea for petrochemical enterprises to implement enterprise e-commerce procurement management. It also conducted an in-depth analysis of the internal and external environment of a company's crude oil e-commerce procurement, conducted a comprehensive assessment of prevailing crude oil suppliers and possible suppliers, and formulated corresponding procurement strategies.

2. Related Work

In the past, in many organizations, procurement was considered to be an unimportant function, and its role and influence in the enterprise were relatively low as a supporting role. Until the 21st century, people have a preliminary understanding of the strategic importance of procurement and its importance for reducing organizational costs. E-commerce procurement is regarded as the source of the enterprise value chain and plays a vital role in the construction of enterprise competitiveness. The focus of enterprise management becomes the further integration of procurement management and the entire business process of the enterprise. The development of world-class concepts, lean supply, value chains, etc. requires a strategic and dynamic procurement function. Lu outlines the definition of e-commerce procurement: e-commerce procurement is the process of planning, implementing, and controlling strategic and operational procurement decisions [4]. The purpose is to guide all activities of the procurement department around improving the company's capabilities to achieve the company vision plan. E-commerce procurement emphasizes the process of establishing a service supply channel at the lowest total cost. One of its core issues is "total cost". Several other issues are supplier selection, information sharing, and quality issues. Wang pointed out that the company's e-commerce management level determines the level of e-commerce procurement of enterprises to a certain extent [5].

With regard to research on supplier partnerships and strategic alliances, research in the literature shows that supplier alliances bring better market expansion, new technologies and knowledge, and higher return on investment, but the rate of failure is also high; probably, it is 60% to 70%. This shows that a successful strategic alliance requires the two sides to fulfill their commitments learn from each other and strengthen information communication on the issue of coordinated development, rather than simply providing assistance or convening a seminar to solve a problem. An H refers to partnerships in the supply and demand chain as supply alliances, stating that partnerships must be beneficial to both parties [6]. Al-Somali et al. pointed out that partnerships are a long-term commitment and agreement between buyers and suppliers, including information sharing and sharing and sharing of benefits and risks due to partnerships [1]. In other words, the concept of a partner must be based on cooperation and trust. They looked at an example of how to use reverse marketing methods to find qualified suppliers and considered this to be the basis for successful partnerships. Guan pointed out that in order to improve the efficiency of the supply chain, it is more important to build partnerships than information technology, because partnerships are a prerequisite for the successful use of IT [7]. Valmohammadi analyzed the factors that hindered the internationalization of procurement, such as distance, time zone differences, consumer behavior differences, and monetary and financial differences [8]. At present, e-procurement has become an indispensable means of establishing a competitive advantage for enterprises. The development of e-procurement has a huge impact on the global economy.

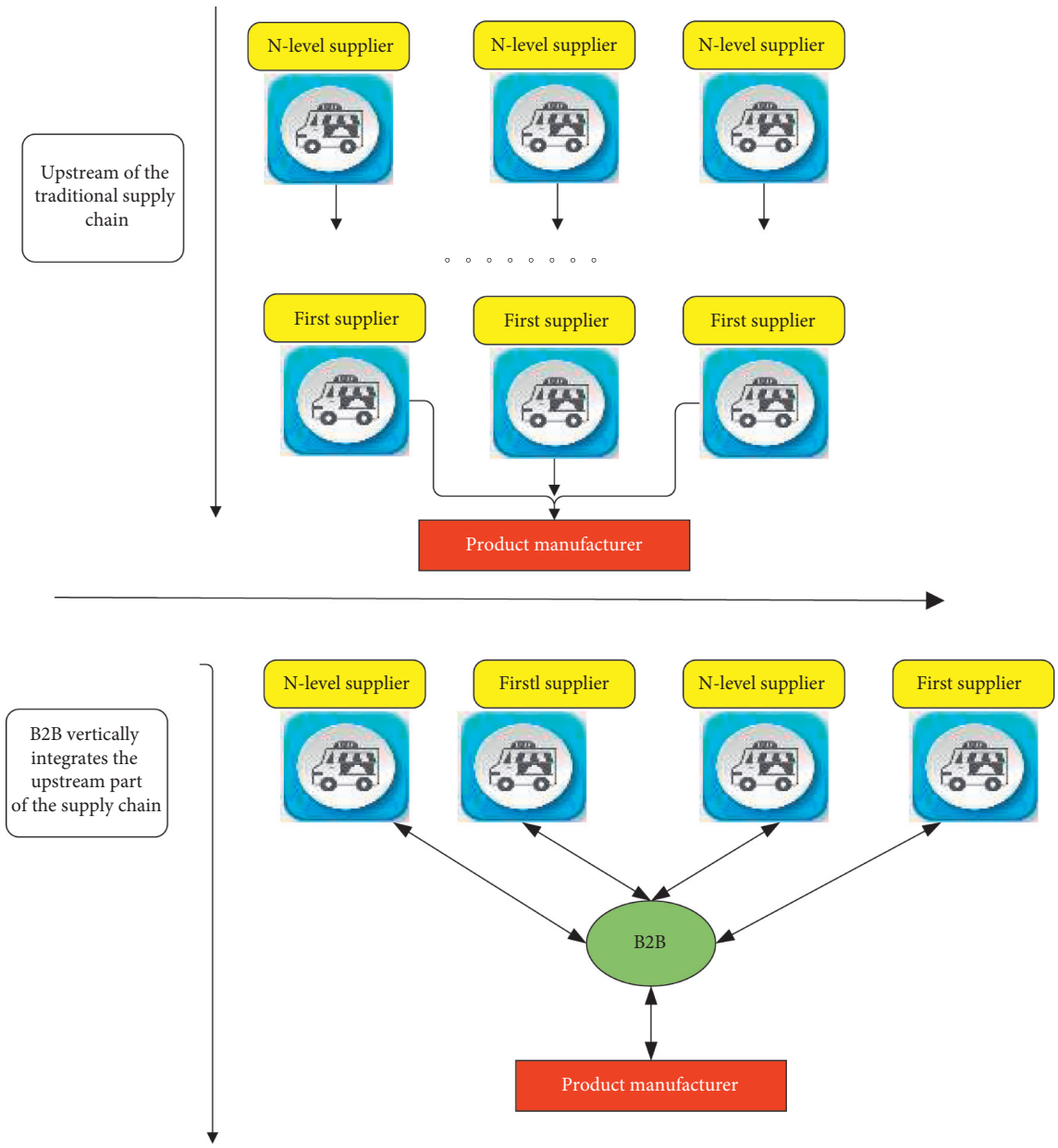


Figure 1: B2B impact on the upstream part of the supply chain.

3. Analysis of Supply and Demand Situation of Crude Oil Procurement Resources of Company A

In a traditional supply chain, a product manufacturer often connects several levels of supplier nodes. These suppliers mainly include raw material suppliers, outsourced manufacturers, and in-house manufacturers, which are more complex supply chain networks downstream of the supply chain, as shown on the left in Figure 1 [9]. However, in the e-commerce environment, through the procurement platform, product manufacturers can achieve not only the vertical integration of several upstream suppliers but also the best way to select the best products and the best in the world. The partner is shown on the right in Figure 1. The B2B model transforms the "bargaining" relationship between product manufacturers and suppliers into a win-win "partnership" relationship, eliminating unnecessary waste and operations on the supply chain network, and promoting the supply chain to be dynamic, virtual, and global. The direction of networking is developing. Compared to the B2C model, B2B achieves a deeper change in the enterprise supply chain.

3.1. International Crude Oil Resource Supply Situation. The global demand for oil is increasing, and the price of crude oil continues to reach a high level, and the core position of oil as an important energy source cannot be shaken in a short period of time. Although the current crude oil price has fallen to a low point, with the economic

recovery and extreme concerns about inflation, a rise in crude oil prices is expected. The international oil price is obviously affected by the trend of the US dollar and the US economy. The reduction of oil companies has added more variables to the world economy, and it is also directly linked to the fluctuation of the US dollar exchange rate. Oil production cuts are undoubtedly factors that cannot be underestimated in the economic recovery. The International Energy Agency, the source and state regulator, predicts that by current prices, by 2020, OPEC will need to increase its daily output from 30 million barrels to 50 million barrels to meet rising demand. But Saudi senior energy officials warned the US and European counterparts that it is extremely difficult for OPEC to meet this level of demand. According to Saudi calculations, there will be a gap of 4.5 million barrels between the global demand for oil produced by Saudi Arabia and the capacity that Saudi officials believe can be achieved. As shown in Table 1, the world's top ten crude oil producers rank first in the world in terms of Saudi oil reserves and will shoulder half of OPEC's production growth in the next 10 to 20 years. The rest will be mainly from Kuwait and the United Arab Emirates. Although global oil demand continues to rise, OPEC's oil production is nearing its limit, and the space for oil production is gradually reaching its peak, while global oil giants are facing the test of oil production cuts. Although the world has made certain breakthroughs in new energy development, the key to the problem is that the cost of use is too high, making it difficult to gain popularity in the short term. At present, the competition for petroleum resources in the world is becoming increasingly fierce. The strategic layout of the United States in global energy competition has taken shape, and the fast-growing economies such as China, Japan, and India are all in the energy industry [10]. We expect the EU to make strategic adjustments in energy in the coming period. Various economies around the world are strategically deploying around energy issues, and this energy competition is expected to intensify as global oil production capacity approaches its limits. The world's top ten original producing countries are shown in Table 1.

3.2. Domestic Crude Oil Resource Supply Situation. China's oil is mainly consumed in the industrial sector, followed by transportation, agriculture, commerce, and consumer spending. Among them, the proportion of industrial oil consumption in the total oil consumption of the country has been maintained at more than 50%. Transportation oil consumption is second only to industry, accounting for about 25%. The proportion of oil consumption in agriculture, forestry, animal husbandry, fishery, and water conservancy accounts for less than 10% of total oil consumption. The proportion of domestic consumption of oil to total oil consumption is only about 6%. According to data released by the customs, the top five countries in China's crude oil imports in 2017 are Saudi Arabia, Angola, Iran, Russia, and Oman. The total import growth of the five countries to China's crude oil suppliers is between 3.7% and 22.4% [11].

Oil is an important strategic material that plays a pivotal role in the national economy. At present, in China's one-time energy consumption, oil consumption accounts for 17-18%. With the rapid development of China's national economy and the continuous improvement of the level of industrialization, the demand for high-quality energy such as petroleum is increasing, and the petroleum industry has gradually become a bottleneck restricting the development of the national economy. Therefore, measures must be taken to promote the healthy and stable development of China's oil industry and ensure the balance of supply and demand of China's oil. In summary, the rise in crude oil prices is an inevitable trend unless humans find alternative energy sources. International crude oil prices have an important impact on the world economy, production efficiency, balance of payments, social price levels, and international relations [4]. Countries are paying close attention to their trends in order to facilitate the healthy development of their economies and establish a good diplomatic environment. As a large-scale enterprise in China's petrochemical industry, Company A's demand for crude oil is increasing, and the strategic management of crude oil procurement will be extremely important.

4. Comparative Analysis of the Advantages of E-Commerce Procurement Model

4.1. Traditional Procurement Model. Company A is a large state-owned enterprise established more than 50 years ago. For a long time, the company mainly exists as a raw material producer of the petrochemical group, and the material management method has been following the model of the planned economy period. There are many problems such as long cycles, low efficiency, slow response, high inventory, and high capital occupation. Its traditional procurement process is shown in Figure 2.

However, Company A's traditional procurement model has the following drawbacks:

(1) There are too many organizational levels, and people are overstaffed. It can be seen from the traditional material procurement flow chart that in order to complete the material procurement task, the corresponding organizational structure and personnel must be set up to complete the corresponding procurement business. In particular, in the search for suppliers and consulting prices, there must be a large number of personnel and complicated procedures to complete. Buyers are running around the world to find and evaluate suppliers. Sometimes, in order to find a reasonable price, the buyer has to run many places. By looking for suppliers in a variety of ways, such as telephone and fax, it is possible to find materials that you can accept and then to determine the price and variety [12]

(2) The channel is single and the supply is unstable. The channel mentioned here is not only a channel for material procurement but also a channel for information acquisition. In the traditional material

Table 1: The world's top ten original producing countries.

Ranking	Country	Daily average crude oil production (10,000 barrels)
1	Saudi Arabia	1085.9
2	Russia	976.9
3	United States	687.1
4	Iran	434.3
5	China	368.4
6	Mexico	368.1
7	Canada	296.6
8	United Arab Emirates	282.2
9	Venezuela	279.8
10	Norway	277.3

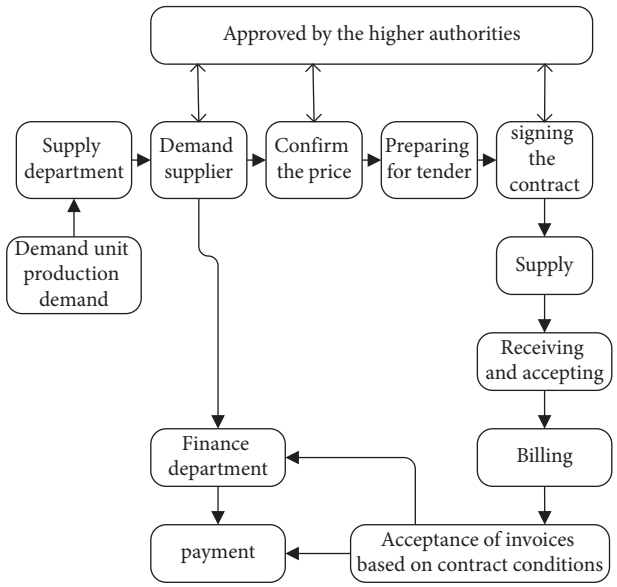


Figure 2: A company's traditional procurement model diagram.

procurement process, the first thing the factory has to do is to prepare before the transaction. For the trading process, pretransaction preparation is the process of how the supply and demand sides can promote or obtain effective product information. The marketing strategy of the supplier of the goods is to promote their own product information through various forms of advertising such as newspapers, television, and outdoor media. For the needs of the goods-enterprises and consumers, you should get the information you need as much as possible to enrich your purchase channels. Therefore, the preparation before the transaction is actually a process of publishing, inquiring, and matching the product information. For buyers, they mainly obtain information about the items they need and the companies that produce them through newspapers, magazines, television, outdoor media, online catalog flyers, and so on. However, the information obtained in this way is always unreliable, and a large number of manpower and material resources have to be dispatched to find out the authenticity of the product and the enterprise. Even so, it is often difficult to get the real information you need. In addition, the choice of suppliers is also very random, and there is no long-term partnership, which may lead to the purchaser's purchase, supply, and production disjoint

(3) The procurement cost is too high. As mentioned above, obtaining information requires a certain amount of cost, and the information of both parties is not shared because the transaction between the buyer and the seller is based on a zero-sum strategy and is a hostile relationship. The profit of such a party is based on the loss of the other party, so the information of both parties is kept secret to the other party

(4) High inventory costs: In the traditional production process, in order to ensure the production of the required materials, the material procurement department always builds a large number of warehouses to gather a large number of regular materials in production; otherwise, it will not be able to meet production. This will take up a lot of money and directly affect the normal flow of corporate funds

(5) Long procurement cycle: The actual operation in the traditional material procurement process consists of pretransaction preparation, contract negotiation, contract execution, and contract settlement. The material procurement department starts from the information of the material demand plan, to find information, business negotiation, material transportation, and other aspects. In the end, the materials that arrive at the demand department (such as various production workshops) take a long time, which is one or two weeks, and several months. After such a long period of time, some materials were needed at that time. After waiting for a long time, due to changes in production processes and technologies, the purchased materials could no longer meet the production needs. These problems are often encountered [13]

Through the above analysis of the shortcomings of the traditional procurement model, it can be seen that the model is no longer suitable for the survival and development of enterprises under the market economy. So what can be done to solve or improve? With the rapid development of the Internet, e-commerce work and e-commerce procurement

have become indispensable technical means to promote continuous restructuring, strengthen management, reduce production and operation costs, and enhance market competitiveness.

4.2. Crude Oil Procurement Model under E-Commerce. E-commerce procurement is a procurement model in the e-commerce environment, also known as online procurement. E-commerce procurement is a new way of purchasing along with the development of the information society. E-commerce purchases are companies that publish supply and demand information or find suitable suppliers and items via the Internet. E-commerce procurement activities are realized through online product understanding, online negotiation, online comparison, online bidding, online ordering, and online payment. E-commerce procurement is a new procurement model that enables all-day, worldwide, and transparent procurement activities. It is beneficial to shorten the procurement time, reduce the cost of purchasing, reduce the purchase inventory, and achieve the efficiency of enterprise procurement [14].

E-commerce procurement can search for information of products from the Internet through the Internet, which is beneficial for enterprises to collect all product information when purchasing and to do the comparison when there is no need to travel and go to the market to see the goods. This is a great aspect of the progress of corporate procurement and the efficiency of corporate procurement. The process of e-commerce mining from the process of product understanding, product negotiation, product price comparison, product bidding, product ordering, payment, etc. is carried out online without the need to negotiate. This is conducive to the transparency of the procurement process and can effectively reduce the problem of corruption in the procurement process. At the same time, after e-commerce procurement conducts product transactions online, logistics can quickly deliver products to the enterprise, which helps to shorten the procurement cycle.

4.2.1. Characteristics of E-Commerce Procurement. E-commerce procurement is a new procurement model, which can achieve an all-day, full-efficiency world and full transparency of purchasing activities. It is beneficial to shorten the procurement time, reduce the cost of purchasing, reduce the purchase inventory, and achieve the efficiency of enterprise procurement.

E-commerce procurement from the product understanding, product negotiation, product price comparison, product open bidding, product ordering, payment, and other processes is carried out online without the need to face the transaction, which is conducive to the open and transparent procurement process. There is no distance or border on the Internet, and any supplier can bid to the purchasing industry for a wide range of industries. The purchaser has also made a targeted selection based on the supplier's situation. In the process of e-commerce procurement, publishing supply and demand information or finding suitable suppliers and interactive items is done through the Internet. The information between the purchaser's enterprise and the supplier is also exchanged through the network, which is a convenient, fast, and low-cost network. E-commerce procurement can search for information on products from the Internet through the Internet, which is beneficial for enterprises to collect all product information when purchasing. In the case of no need to travel and go to the market, low-cost manufacturers to see the goods compared to the situation, to achieve a shop-by-shop, can be said to save money, save time, save people, and save the workload [15].

E-commerce procurement activities can be realized through online product understanding, online negotiation, online comparison, online bidding, online high-speed ordering, online payment, etc. After e-commerce purchases are conducted online, the products can be quickly sent to the products. The summary is shown in Figure 3.

4.2.2. Procurement Model in E-Commerce Environment. Material procurement and management play an important role in the enterprise and are the key to the integration of internal and external resources. At present, in the traditional material procurement process of Company A, the structure is bloated, huge, and people are overstaffed. The problems in the management of mining and management have seriously affected the development of the entire company and the improvement of market competitiveness. In the material procurement process in the modern e-commerce environment, how can the process be set up to complete the procurement task and adapt to the needs of the times? As an effective way to change the status quo of a company's procurement management, in the process of process reengineering, the flat procurement rules should be followed, and the material procurement process should be reengineered in the e-commerce environment [16]. The procurement flow chart is shown in Figure 4.

We set up a special supplier management department in the material supply department to increase the dynamic evaluation of suppliers' supply capacity, material quality, and supply speed, eliminate unqualified suppliers, and motivate excellent suppliers. We optimize the supply chain of Company A through the management of suppliers. Company A has reorganized the material procurement organization and strengthened the original material supply department. It is directly affiliated with Company A and has unified planning, unified procurement, and unified management of the materials required by the entire company. The original decentralized procurement mechanism was transformed into centralized procurement and management, which strengthened the company's control and management of material procurement. In addition, the Materials Supply Department has set up a special supplier management department, which is responsible for the management of suppliers, strengthens the ability of Company A to integrate external resources, and is conducive to the establishment of an effective supply chain. After the reorganization, the entire company's procurement philosophy and procurement operations have undergone great changes [17], as shown in Table 2.

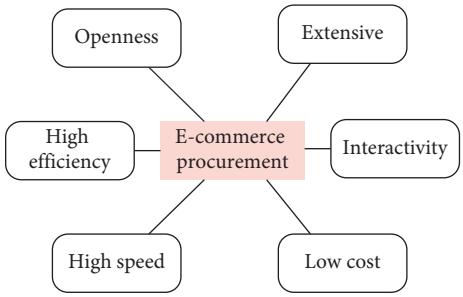


Figure 3: Characteristics of e-commerce procurement.

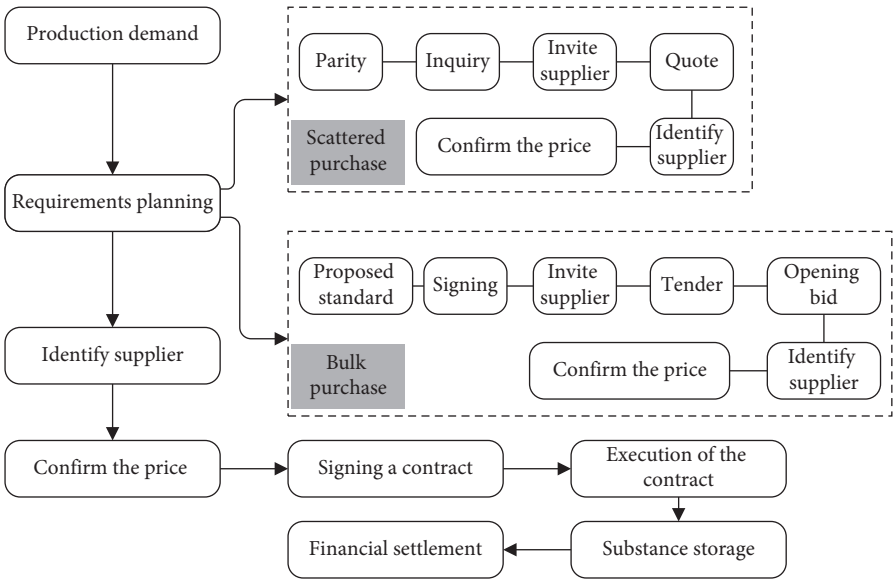


Figure 4: A company's overall e-commerce procurement process.

Table 2: Comparison of the company's traditional procurement model and e-commerce-based procurement model.

Project	Traditional procurement model	E-commerce procurement model
Drive mode	Inventory plan	Production demand plan
Departmental focus	Commodity trading activities with suppliers, such as price negotiations, signing contracts, arranging transportation, and goods acceptance	Inspection and evaluation of suppliers, operation of procurement plans, supplier information maintenance and substance acceptance, feedback
Departmental nature	An auxiliary department within the enterprise that manages the procurement activities of the enterprise	The outward-oriented pioneering department of the enterprise develops and manages the external resources of the enterprise
Work goals	Complete the purchase plan and reduce the purchase price	Ensure the material needs of the production sector, actively develop external resources, find or cultivate excellent suppliers, and guarantee competitive advantage
Purchasing behavior	Mostly one or a few	Teamwork
Purchase cost	Doron negotiations, consultation, high cost	Program simplification and cost reduction
Control of quality, quality, and delivery time	Ex post control	Strict selection and evaluation of suppliers, feedback on the use of prior control
Exchange of information with suppliers	Asymmetry	Information sharing
Number of suppliers	Multiple, encourage, and compete	A small number of partnerships that encourage cooperation

Table 3: General indicators of crude oil supplier evaluation.

Overall target layer	Primary indicator	Secondary indicators
General indicator of crude oil supplier evaluation	Crude oil quality	Light oil yield
		Density
		Moisture
		Sulfur content
	Purchase cost	Crude oil prices
		Transport fees
		Financial expenses
	Supplier potential	Future sustainable mining capabilities
		Policy impact
		Cost of mining

From this point of view, under the centralized procurement model, the organizational structure of enterprises has undergone tremendous changes compared to the decentralized procurement model. Each subsidiary and secondary accounting unit does not have its own independent procurement department. They focus on their procurement department (Material Supply Division) by presenting their own procurement needs to the company's only centralized purchasing department. Through the collection and unification of various needs, a large purchase order is formed, and the demand is satisfied through various procurement methods such as bidding procurement, price comparison procurement, advanced source procurement, and special procurement of shortage products [18], thereby reducing procurement costs and improving procurement quality and efficiency.

5. Construction of Supplier Evaluation Index System in Procurement Mode

5.1. Establishment of the Indicator System. The relevant research results of this study used the evaluation indicators to evaluate the crude oil suppliers of Company A, which are crude oil quality, procurement cost, and potential. The overall indicators are shown in Table 3.

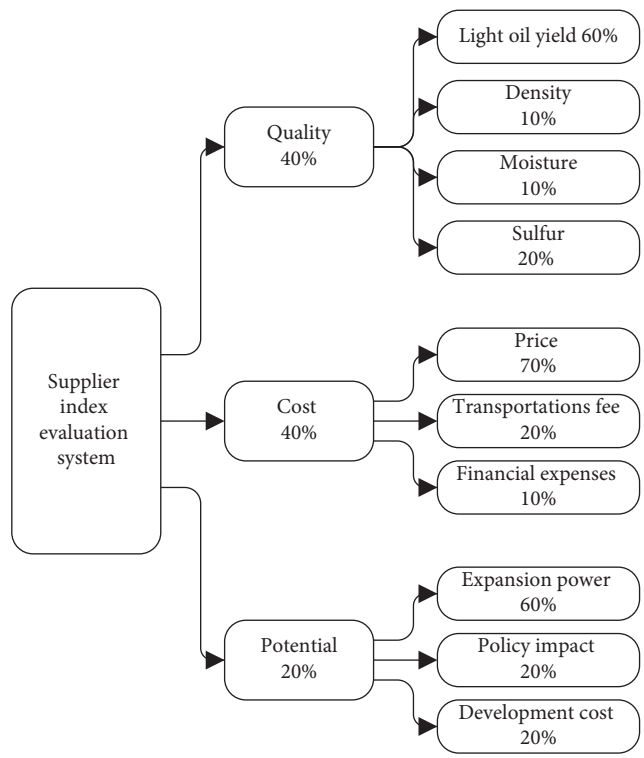


Figure 5: Hierarchical analysis of supplier indicators.

5.2. Construction of Supplier Indicator System. Crude oil is the main raw material of refinery equipment. The quality of crude oil, the cost of procurement, and the potential of suppliers are the key to the safe and stable production and operation efficiency of enterprises. It is also the key to the smooth production and optimization of chemical equipment. The above indicators must be considered in the selection of suppliers. Establishing a scientific indicator system for suppliers is conducive to optimizing the structure of crude oil procurement, reducing procurement costs, and ensuring smooth operation of the equipment. The data of the supplier indicator system is processed by the analytic hierarchy process, as shown in Figure 5 [21].

We establish a supplier database based on the supplier's relevant information and analyze and process the data. The supplier data of Supplier A's supplier index system is evaluated by the expert research method, and the data is classified and scored. Based on this, a comprehensive evaluation is carried out. The supplier's scoring data is shown in Tables 4–6.

5.3. Determination of Calculation Results and Allocation of Procurement Ratio. According to the above data, the supplier is evaluated. The results are shown in Table 7.

After calculation, each supplier can be divided into four files: A, B, C, and D. A is a Russian import crude oil supplier; the Russian crude oil supplier comprehensive evaluation score is the preferred supplier, although its procurement cost is high, its crude oil quality is high, and its future sustainable development ability is strong; B is Inner Mongolia Cooperative Oilfield and Jilin Oilfield; C is Huadian shale oil; D is a Shuangyang foreign-funded cooperative oilfield and Daqing Oilfield.

Company A's crude oil processing capacity will reach 10 million tons/year. According to the above data analysis

Table 4: Supplier's crude oil quality data.

Supplier	Daqing oilfield	Jilin oilfield	Russian imports	Inner Mongolia cooperation	Shuangyang foreign investment cooperation
Light oil yield	27.3	26.5	58.3	39.5	21.2
Score	7	7	10	9	6
Density (20 Cg/cm)	0.8356	0.8631	0.8221	0.8281	0.8812
Score	9	9	10	10	8
Moisture %	0.06	0.06	0.09	0.13	0.10
Score	10	10	8	6	7
Sulfur content %	0.16	0.16	0.68	0.10	0.18
Score	9	9	5	10	7

Table 5: Formula to purchase the cost of crude oil from various suppliers.

Supplier	Daqing oilfield	Jilin oilfield	Russian imports	Inner Mongolia cooperation	Shuangyang foreign investment cooperation
Price	5886	5886	6219	5055	5055
Score	7	7	6	8	8
Transport fees	70	51	70	89	32
Score	7	8	7	6	10
Financial expenses	45	45	75	45	35
Score	8	8	5	8	10

Table 6: Supplier's sustainability capabilities.

Supplier	Capacity expansion potential	Policy support	Development costs
Daqing oilfield	6	10	10
Jilin oilfield	8	10	7
Russian imports	9	9	9
Inner Mongolia cooperation	4	8	7
Shuangyang foreign investment cooperation	4	4	5

Table 7: Calculation results of supplier evaluation indicators.

Supplier	Comprehensive evaluation score
1. Daqing oilfield	6.87
2. Jilin oilfield	7.95
3. Russian imports	8.59
4. Inner Mongolia cooperation	8.38
5. Shuangyang foreign investment cooperation	6.94
Average value	7.746

Table 8: A company's crude oil procurement allocation table.

Supplier	Existing capacity constraints	Sulfur content of crude oil	Allocation amount
Daqing oilfield	5500	0.18	25
Jilin oilfield	560	0.17	500
Russian imports	2600	0.56	350
Inner Mongolia cooperation	80	0.15	100
Shuangyang foreign investment cooperation	55	0.21	25
Total	—	0.38	1000

results, crude oil procurement should first select Russia to import crude oil, followed by Inner Mongolia cooperative oilfield and Jilin oilfield. Other crude oil suppliers are trying to reduce the number of purchases while considering the impact of transportation capacity risks and unexpected risks on production. The final purchase allocation is shown in Table 8.

6. Conclusion

The developments in the field of information and communication technologies have made the processing and transmission of information technology break through the limitations of time and region. The network and globalization of the economy have become an inevitable trend in the development of the world economy. The implementation of crude oil e-commerce procurement can reduce the cost of crude oil procurement, create good economic benefits, improve the management level of enterprises, and improve the market competitiveness of enterprises. This lays a solid foundation for the sustainable development of petrochemical enterprises. Based on the characteristics of crude oil procurement of petrochemical enterprises and the principles established by the evaluation index system, this paper constructs an evaluation index system for the strategic procurement plan of crude oil in petrochemical enterprises. The internal and external environment of a company's crude oil e-commerce procurement was analyzed in depth, and the existing crude oil suppliers and potential suppliers were comprehensively evaluated and the corresponding procurement strategies were formulated. Results of the study show the success of the proposed study, and on the basis of the proposed study, researchers can devise new solutions in the field.

Data Availability

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Conflicts of Interest

The author declares that he has no conflicts of interest.

References

[1] S. A. Al-Somali, R. Gholami, and B. Clegg, "A stage-oriented model (SOM) for e-commerce adoption: a study of Saudi Arabian organisations," *Journal of Manufacturing Technology Management*, vol. 26, no. 1, pp. 2–35, 2015.

[2] L. Kilian and D. P. Murphy, "The role of inventories and speculative trading in the global market for crude oil," *Journal of Applied Econometrics*, vol. 29, no. 3, pp. 454–478, 2014.

[3] M. J. S. Chandrasekar Subramaniam, "A study of the value and impact of B2B e-commerce: the case of web-based procurement," *International Journal of Electronic Commerce*, vol. 6, no. 4, pp. 19–40, 2002.

[4] L. Lu and T. Reardon, "An economic model of the evolution of food retail and supply chains from traditional shops to supermarkets to E-commerce," *American Journal of Agricultural Economics*, vol. 100, no. 5, pp. 1320–1335, 2018.

[5] W.-T. Wang, Y.-S. Wang, and E.-R. Liu, "The stickiness intention of group-buying websites: the integration of the commitment-trust theory and e-commerce success model," *Information & Management*, vol. 53, no. 5, pp. 625–642, 2016.

[6] H. An, W. Zhong, Y. Chen, H. Li, and X. Gao, "Features and evolution of international crude oil trade relationships: a trading-based network analysis," *Energy*, vol. 74, no. 1, pp. 254–259, 2014.

[7] Q. Guan, H. An, X. Gao, S. Huang, and H. Li, "Estimating potential trade links in the international crude oil trade: a link prediction approach," *Energy*, vol. 102, no. 1, pp. 406–415, 2016.

[8] C. Valmohammadi and S. Dashti, "Using interpretive structural modeling and fuzzy analytical process to identify and prioritize the interactive barriers of e-commerce implementation," *Information & Management*, vol. 53, no. 2, pp. 157–168, 2016.

[9] L. Roselli, D. Carlucci, and B. C. De Gennaro, "What is the value of extrinsic olive oil cues in emerging markets? Empirical evidence from the U.S. E-Commerce retail market," *Agribusiness*, vol. 32, no. 3, pp. 329–342, 2016.

[10] A. Davila, M. Gupta, and R. Palmer, "Moving procurement systems to the Internet: the adoption and use of e-procurement technology models," *European Management Journal*, vol. 21, no. 1, pp. 11–23, 2003.

[11] S. X. Xu and G. Q. Huang, "Efficient multi-attribute multi-unit auctions for B2B E-commerce logistics," *Production and Operations Management*, vol. 26, no. 2, pp. 292–304, 2017.

[12] H. O. Awa, O. U. Ojiabo, and B. C. Emecheta, "Integrating TAM, TPB and TOE frameworks and expanding their characteristic constructs for e-commerce adoption by SMEs," *Journal of Science & Technology Policy Management*, vol. 6, no. 1, pp. 76–94, 2015.

[13] H. G. Huntington, "Crude oil trade and current account deficits," *Energy Economics*, vol. 50, pp. 70–79, 2015.

[14] S.-J. Eom, S.-C. Kim, and W.-S. Jang, "Paradigm shift in main contractor-subcontractor partnerships with an e-procurement framework," *KSCE Journal of Civil Engineering*, vol. 19, no. 7, pp. 1951–1961, 2015.

[15] R. Du, Y. Wang, G. Dong et al., "A complex network perspective on interrelations and evolution features of international oil trade, 2002-2013," *Applied Energy*, vol. 196, no. 6, pp. 142–151, 2017.

[16] R. Du, G. Dong, L. Tian et al., "A complex network perspective on features and evolution of world crude oil trade," *Energy Procedia*, vol. 104, no. 12, pp. 221–226, 2016.

[17] M. A. Rahman and L. Casanovas, "Strategies to predict E-commerce inventory and order planning," *International Journal of Technology Diffusion*, vol. 8, no. 4, pp. 17–30, 2017.

[18] P. Yan and Z. Aihui, "An analysis of purchasing model of furniture enterprises based on E-commerce platform," *Forest Engineering*, vol. 3, no. 1, p. 38, 2015.

Hindawi
Scientific Programming
Volume 2021, Article ID 9936337, 11 pages
https://doi.org/10.1155/2021/9936337

Research Article

Measurement Method of Distributed Nodes in Wireless Sensor Networks Based on Multiple Attributes

Bing Zheng and Jing Yang

Department of Information Engineering, Hainan Vocational University of Science and Technology, Haikou 571126, China

Correspondence should be addressed to Jing Yang; yj13876195734@163.com

Received 8 March 2021; Revised 28 March 2021; Accepted 2 April 2021; Published 22 April 2021

Academic Editor: Shah Nazir

Wireless sensor network (WSN) is usually organized in a particular area for achieving some specific tasks and functions. It has broad application prospects in military, environmental monitoring, disaster relief, and other many commercial areas. Measuring the significance of distributed nodes is the foundation of many applications of WSN. Therefore, this paper studies the distributed node measurement methods and approaches based on multiple attributes in wireless sensor networks to improve the accuracy of distributed node importance measurement. The entropy weight TOPSIS method is used to calculate the importance index of distributed nodes of degree centrality, eigenvector centrality, compactness centrality, betweenness centrality, K-kernel decomposition centrality, and aggregation coefficient; the PR value of each node is given by combining the PageRank algorithm; and then, the parameters α and β are introduced to calculate the comprehensive importance measurement results of each node, to finally obtain the importance measurement results of each node dynamically through the improved node deletion method. The experimental results of the research show that the proposed method can accurately measure the importance of distributed nodes in wireless sensor networks, and the accuracy is as high as 98%.

1. Introduction

Wireless sensor network is a hot research field in the world, which involves highly interdisciplinary and highly integrated knowledge [1]. It integrates sensor technology, embedded computing technology, modern network and wireless communication technology, distributed information processing technology, and so forth. It can monitor, perceive, and collect information of various environments or monitoring objects in real time through various integrated microsensors. The information is sent by the wireless way and transmitted to the user terminal by a self-organizing multihop network [2], so as to realize the connection of the physical world, the computing world, and the human society. Wireless sensor network has a very broad application prospect. It has potential practical value in many important fields, such as military defense, industry and agriculture, urban management, and biomedicine. It has attracted great attention of academic and industrial circles in many countries [3]. It is considered to be one of the technologies that have a great influence on the 21st century. As a new type of information acquisition and processing mode, wireless sensor network (WSN) has randomly distributed nodes integrated with sensors, data processing units, and communication modules [4].

With the help of built-in various sensors, it can measure the thermal, infrared, sonar, radar, and seismic signals in its surrounding environment, so as to detect many networks including temperature, humidity, noise, and light intensity. With the rapid development of Internet technology, wireless communication technology, and electronic technology, wireless sensor network has attracted wide attention in the world [5]. Although the wireless sensor network is similar to the ad hoc network, there are also differences. The wireless sensor network is a network system integrated with monitoring, control, and wireless communication. The number of nodes is larger (thousands or even tens of thousands), and the distribution of nodes is more intensive. Due to

environmental impact and limited energy, nodes are more prone to failure [6]. Environment interference and node failure are easy to cause the change of network topology. Usually, most sensor nodes are fixed. The processing capacity, storage capacity, and communication capacity of sensor nodes are very limited. In addition, sensor nodes are small in size and usually carry batteries with limited energy. Due to the large number of nodes, low-cost requirements, wide distribution area, and complex deployment area, some areas cannot even be reached by personnel. Therefore, compared with the traditional wireless network, the primary design goal is to provide high-quality service and bandwidth utilization. Wireless sensor network takes data acquisition as the center, and the primary design goal is the efficient use of energy [7], which is also one of the most important differences between the wireless sensor network and traditional network. The typical application of the wireless sensor network is deployed in a specific area to complete the functions of environmental monitoring, target tracking, and positioning. In these applications, sensor nodes transfer the collected data to the server for processing. In order to ensure the validity of data, the importance ranking information of nodes is often essential.

Xiaoyang Sun et al. designed a distributed wireless sensor node monitoring system, which has been successfully applied to the evaluation of underground explosion damage effectiveness; Wang Lu et al. studied the node importance measurement based on the weighted method [8]. First, considering the global information of the network, the characteristic center vector value of each node was calculated, and the weights of the nodes at both ends and the edges formed a weighted network. The local information of weighted network is used to calculate the node importance of weighted network; and Gang Hu et al. [9] study the calculation of node importance of complex network based on interpretive structure model and apply interpretive structure model to matrix the relationship between nodes of the directed network to obtain the corresponding adjacency matrix and reachable matrix. The reachable matrix is divided by location and level; the matrix is reduced, and leapfrog and self-adaptive are deleted. The hierarchical digraph of the network is obtained, and the network matrix is weighted to simulate evolution, and the identification and division of network region importance and level importance are given. This method is not only suitable for the hierarchical division and identification of the directed network but also suitable for the node ranking calculation of the directed network. In order to solve the limitations of distributed node measurement, based on multiple attributes, a distributed node measurement method based on multiple attributes is studied, which effectively avoids the limitations of single attribute measurement and improves the accuracy of distributed node importance measurement. The entropy weight TOPSIS method is considered for calculating the importance index of distributed nodes of degree centrality, eigenvector centrality, compactness centrality, betweenness centrality, K-kernel decomposition centrality, and aggregation coefficient; the PR value of each node is given by combining the PageRank algorithm; then, the parameters α and β are introduced to calculate the comprehensive importance measurement results of each node, to finally obtain the importance measurement results of each node dynamically through the improved node deletion method.

2. Distributed Node Measurement Method Based on Multiple Attributes in Wireless Sensor Networks

The following sections briefly show the details of this section.

2.1. Distributed Node Importance Metrics in Wireless Sensor Networks. Let $G = (V, E)$ be a wireless sensor network, where $V = \{v_1, v_2, \ldots, v_N\}$is the set of all distributed nodes in the wireless sensor network, $|V| = N$; $E = \{e_1, e_2, \ldots, e_m\} \subseteq V \times V$ is the set of edges between distributed nodes, $|E| = m$, $A = a_{ij}$ of wireless sensor network, where $a_{ij} = 1$ means that distributed node i is connected to distributed node j; otherwise, $a_{ij} = 0$.

Due to the different evaluation criteria of the importance of distributed nodes, the definition of the importance index of distributed nodes is also different. The importance index of distributed nodes is defined as follows:

(1) Degree centrality:

The degree of distributed node i is defined as the number of neighbors of the node, specifically expressed as

$$K(i) = \sum_{j \in G} a_{ij}. \tag{1}$$

(2) Eigenvector centrality:

The eigenvector index considers the status and reputation of distributed nodes in wireless sensor networks and considers the reputation of a single node as a linear combination of the reputation of other nodes, so as to obtain a linear equation group. The eigenvector corresponding to the maximum eigenvalue of the system is the importance of each distributed node, which can be specifically expressed as

$$C_e(i) = \lambda^{-1} \sum_{j=1}^{N} a_{ij} e_j, \tag{2}$$

where λ is the maximum eigenvalue of the adjacency matrix A and $e = (e_1, e_2, \ldots, e_n)^T$ is the eigenvector corresponding to the maximum eigenvalue λ of the adjacency matrix A.

(3) Closeness centrality:

Compactness is used to measure the ability of distributed nodes in wireless sensor networks to exert influence on other nodes through wireless networks, which can be specifically expressed as

$$C_c(i) = \frac{(N-1)}{\sum_{j=1}^{N} d_{ij}}, \tag{3}$$

where d represents the shortest distance from node i to node j.

(4) Betweenness:

Betweenness is used to measure an individual's social status, which means the number of nodes passing through all the shortest paths in wireless sensor networks. It can be expressed as

$$C_b(i) = \sum_{s<t} \frac{\sigma_{st}^i}{\gamma_{st}}, \tag{4}$$

where γ represents the number of shortest paths between node s and node t and σ represents the number of shortest paths between node s and node t through node i.

(5) Index of location attributes in wireless sensor networks:

The index of location attribute of wireless sensor network is K-shell centrality. K-shell is a distributed node importance ranking index obtained by the K-kernel decomposition method, and the importance of nodes is measured according to the position of nodes in the whole wireless sensor network [10].

(6) Aggregation coefficient:

Aggregation coefficient is the local attribute of the wireless sensor network of distributed nodes, which reflects the frequency of communication between each node in the wireless sensor network. The aggregation coefficient is inversely proportional to the importance of the distributed node. The larger the calculated value, the weaker the importance of the node. The aggregation coefficient of distributed node i is defined as the actual number of edges E_i among μ_i neighboring nodes directly connected to the node divided by the total number of possible edges $C_{k_i}^2$, which can be expressed as

$$C_\tau(i) = \frac{E_i}{C_{\mu_i}^2} = \frac{2E_i}{\mu_i(\mu_i - 1)}. \tag{5}$$

2.2. Node Importance Measurement Method Based on Multiple Attributes and Node Deletion

2.2.1. Node Importance Measurement Attribute. A variety of distributed node importance measurement attributes, including degree centrality, intermediate centrality, and proximity centrality. Distributed node measurement attribute not only describes the basic characteristics of nodes but also can be easily calculated. Based on the above reasons, the six basic attributes of degree centrality, eigenvector centrality, compactness centrality, dielectric centrality, k-kernel decomposition centrality, and aggregation coefficient are selected as part of the multiple attributes in the node importance measurement method integrating multiple attributes and node deletion. After the measurement attribute is selected, it is necessary to consider the influence of attribute value on distributed node importance [11], whether it is positively correlated or negatively correlated, and then unify the attribute value with the node importance measurement relationship. The calculated values of the five attributes selected in this paper, namely, medium centrality, eigenvector centrality, compactness centrality, intermediate centrality, and k-kernel decomposition centrality, are positively correlated with the importance of nodes; that is, the larger the calculated result, the more important the node [12]. The results of the clustering coefficient attribute are negatively correlated with the importance of nodes. In order to unify the correlation of the selected attributes, the calculation formula of compactness centrality should be adjusted to change it into a positive correlation attribute. Given that the maximum value of compactness centrality is 1, the adjusted formula is

$$C'_\tau(i) = 1 - C_\tau(i) = 1 - \frac{2E_i}{\mu_i(\mu_i - 1)}. \tag{6}$$

2.2.2. PageRank Algorithm. PageRank algorithm is a link analysis algorithm. The idea of the algorithm is that when a high-quality distributed node points to another node, the quality of the other node is related to the quality of the node pointing to it. If the node pointing to it is of good quality and the node being pointed to is of good quality, the regression relationship is used to calculate the importance of the node [13]. The principle of the PageRank algorithm is as follows: firstly, PageRank assigns the same PR value to all nodes in the wireless sensor network; then, it updates the PR value of each node in real time through continuous iterative recursive operation, and so on until the PR value is stable. The result calculated by PageRank is the result of node importance measurement. Its expression is as follows:

$$\mathrm{PR}(P) = \frac{1-q}{n} + q\sum_{i=1}^{n} \frac{\mathrm{PR}(U_i)}{N_{U_i}}, \tag{7}$$

where PR(P) is the value PR of node P and q is the damping factor bounded by the interval (0,1) and represents the probability that the user continues to visit backward in the process of accessing a certain node at any time [14]. The general value is $q = 0.85$; n is the total number of nodes; $U_1, U_2, \ldots, U_n$ is the node pointing to node P; PR(U) is the PR value of node U pointing to node P; and N is the number of nodes U points to. According to the calculation formula of PageRank, the importance of P node is related to the amount of data transmission of this node. The more the amount of data transmission is [15], the more important P is. At the same time, the more important the node pointing to node P is, the more important node P is, and the greater the PR value of P is.

2.2.3. TOPSIS Algorithm Based on Entropy Weight. TOPSIS is a multiobjective decision-making method. The core idea of this method is to first select a positive ideal solution and a negative ideal solution in wireless sensor networks and then take the solution closest to the positive ideal solution and farthest from the negative ideal solution as

the optimal solution [16]. However, TOPSIS does not explain the rationality of attribute weight setting, so the calculation results are subjective. As an objective weighting method, the entropy weight method is often used to determine the weight of each attribute. Its basic idea is to give weight to each attribute according to the amount of information reflected by the variation degree of each evaluation attribute [17]. Therefore, the concept of "information entropy" is introduced, and the entropy weight method is used in the process of giving multiple attribute weights in TOPSIS; the specific algorithm is as follows:

In the multiattribute measurement problem, the metric matrix $Y = \{y_{iq}\}$ and the normalized metric matrix $Z = \{z_{iq}\}$ are set, and the normalized metric matrix is calculated by the vector specification method. The calculation formula is as follows:

$$z_{iq} = \frac{y_{iq}}{\sum_{i=1}^{n} y_{iq}}, \quad i = 1, 2, \ldots, n; q = 1, 2, \ldots, p, \tag{8}$$

where n represents the number of distributed nodes; p represents the number of attribute types that measure the importance of nodes; y_{iq} is the value of the i-th node on the g-th attribute index.

The entropy weight method is used to deal with the canonical metric matrix of multiple attributes, and the weight of each attribute w_q is obtained. The entropy weight of the g attribute is expressed as

$$w_q = \frac{1 - H_q}{p - \sum_{q=1}^{p} H_q}, \quad q = 1, 2, \ldots, p, \tag{9}$$

where H_q represents the information entropy value of the attribute of item q, and its expression is as follows:

$$H_q = -k \sum_{i=1}^{n} z_{iq} \ln z_{iq}, \quad i = 1, 2, \ldots, n; q = 1, 2, \ldots, p. \tag{10}$$

Set the weight of each attribute as $w = (w_1, w_2, \ldots, w_p)^T$, and use formula (11) to form a weighted canonical matrix $X = \{x_{iq}\}$:

$$x_{iq} = w_q \cdot z_{iq}. \tag{11}$$

Using formulae (12) and (13), the positive ideal solution $x^*$and negative ideal solutionx^o with each attribute are obtained. Let the q-th attribute value of ideal x^* be x_q^*. The q-th attribute value of the negative ideal solution $x^o$is x_q^o; then,

$$x_q^* = \begin{cases} \max\limits_i x_{iq}, \\ \max\limits_i x_{iq}, \end{cases} \tag{12}$$

$$x_q^o = \begin{cases} \max\limits_i x_{iq}, \\ \max\limits_i x_{iq}, \end{cases} \tag{13}$$

d_i^* is the distance between x_i and the positive ideal solution, and d_i^o is the distance between x_i and the negative ideal solution. Calculate the distance from each attribute to the positive and negative ideal solutions, respectively, according to equations (14) and (15):

$$d_i^* = \sqrt{\sum_{q=1}^{p} \left(x_{iq} - x_q^*\right)^2}, \quad i = 1, 2, \ldots, n, \tag{14}$$

$$d_i^o = \sqrt{\sum_{q=1}^{p} \left(x_{iq} - x_q^o\right)^2}, \quad i = 1, 2, \ldots, n. \tag{15}$$

According to formula (16), the comprehensive measurement value of each attribute is calculated, and the value is sorted from large to small, and finally, the importance ranking measurement result of nodes is obtained; the formula is

$$C_i^o = \frac{d_i^o}{d_i^o + d_i^*}, \quad i = 1, 2, \ldots, n. \tag{16}$$

2.2.4. Comprehensive Ranking Measurement Method. Degree centrality, eigenvector centrality, compactness centrality, betweenness centrality, K-kernel decomposition centrality, and aggregation coefficient are six kinds of centrality indexes that affect the ranking of distributed nodes [17]. Starting from the structural characteristics of wireless sensor networks, the importance of nodes is quantitatively analyzed; the PageRank algorithm theory is used to accurately locate nodes in wireless sensor networks importance, on the basis of the above, through the improved node removal method to investigate the influence of the change of wireless sensor network connectivity on the importance of nodes. Based on the above three main research methods of node importance in wireless sensor networks, a node importance measurement method combining multiple attributes and node deletion is proposed; the specific process of this method is shown in Figure 1.

Figure 1 shows the entropy-based multiattribute method and the node types in distributed wireless sensor networks. It is assumed that the node importance metric based on the PageRank algorithm is calculated as $\mathrm{PR}(i)$, introducing α, β to C_i^o and get linear regression node scheduling $\mathrm{PR}(i)$ measurement C_i, node deletion method based on the improved dynamic get distributed node importance ranking of four parts; the concrete steps of the algorithm are as follows:

Step 1: according to the acquired about the connection between the node and the node data to establish the corresponding wireless sensor network, computing degree of each node in wireless sensor network centricity, characteristic vector centricity, tightness of centricity, betweenness centrality, K nuclear decomposition of centricity, and six class attribute values and then by formula (8), the calculation results being normalized processing, establish a standardized measure of $n \times 4$ matrix node properties.

Step 2: the concept of information entropy is introduced into the determination of weight, and the normalized metric matrix obtained in step 1 is calculated

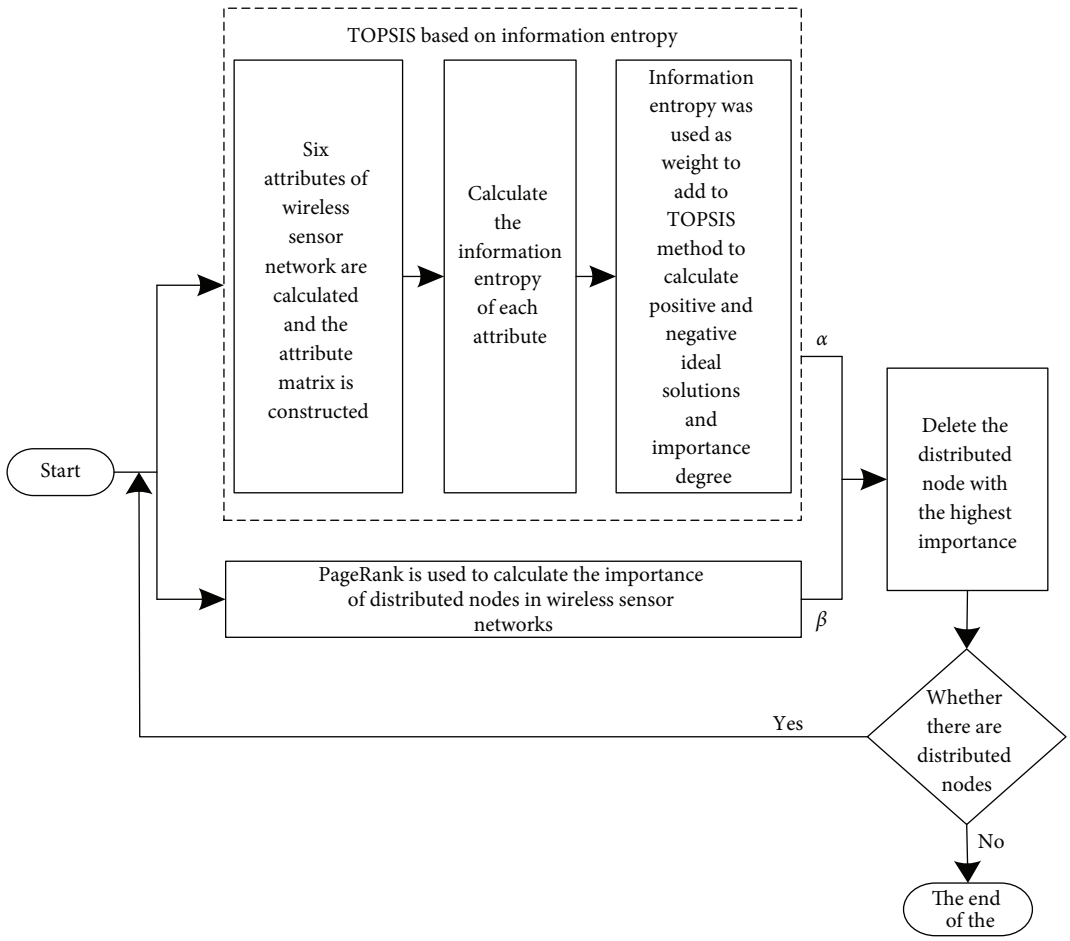


Figure 1: Flow chart of node importance measurement method integrating multiple attributes and node deletion.

by using the entropy weight method to obtain the entropy weights corresponding to various attributes [18], and then the weighted canonical matrix is obtained.

Step 3: the six kinds of attribute values of distributed nodes are taken as multiple attributes in the TOPSIS method, the entropy weight corresponding to each attribute is taken as the attribute weight in the TOPSIS method, and the positive and negative ideal solutions of each attribute are calculated.

Step 4: obtain the importance ranking C_i^o of distributed nodes.

Step 5: the PageRank algorithm is used for reference and its principle is introduced to calculate the PR value of distributed nodes $\mathrm{PR}(i)$. According to the PR value, the importance measurement results of nodes can be obtained and taken as a part of the importance measurement of wireless sensor network nodes.

Step 6: normalize the PR value of distributed nodes. Linear regression is conducted between the node ranking measurement result C_i^o calculated based on entropy weight TOPSIS method and the node ranking measurement value PR calculated based on PageRank algorithm to obtain the improved node importance comprehensive ranking measurement result C_i. The specific calculation formula is as follows:

$$C_i = \alpha \cdot C_i^o + \beta \cdot \mathrm{PR}(i), \tag{17}$$

α and β represent the weighting coefficients of node importance of TOPSIS method and PageRank node importance based on multiattribute entropy weight, respectively, $\alpha + \beta = 1$.

Step 7: on the basis of the above steps, first disconnect the C_i value maximum distributed nodes in wireless sensor network connection. Then connect the disconnected node to the node whose PR value is evenly distributed, and then according to the steps one to six to, recalculate the importance of remaining in the new wireless sensor network node measurement [19, 20], then disconnect the new C_i value maximum in the wireless sensor network node, and recalculate the importance of the residual node values. This cycle continues until only the last node is left in the wireless sensor network, thus obtaining the final node importance ranking metric.

3. Experimental Analysis

In order to verify the effectiveness and relative accuracy of this method, the simulation experiment of wireless sensor network distributed node measurement is carried out in MATLAB, and 20 and 100 randomly distributed nodes are set in the simulation program. The topology of the wireless sensor network in two cases is shown in Figure 2.

3.1. Effectiveness Analysis. Taking a wireless sensor network with 20 distributed nodes as an example, six kinds of attribute values of each distributed node in the wireless sensor network are calculated through simulation experiments, including degree centrality, eigenvector centrality, compactness centrality, betweenness centrality, K-kernel decomposition centrality, and aggregation coefficient. The calculation results of each attribute value are shown in Table 1; the result of the calculation is rounded to four decimal places.

Then, the normalization decision matrix is obtained by using the vector norm method, the entropy weight method is used to process the decision matrix, the information entropy value H_q and weight w_q of each attribute are obtained, and the result retains four decimal numbers where the information entropy values of each attribute are, respectively, $H_1 = 0.9979$, $H_2 = 0.9429$, $H_3 = 0.9987$, $H_4 = 0.9939$, $H_5 = 0.9915$, and $H_6 = 0.9865$. The attribute weights are, respectively, $w_1 = 0.1412$, $w_2 = 0.6863$, $w_3 = 0.2495$, $w_4 = 0.1819$, $w_5 = 0.1607$, and $w_6 = 0.5641$. Then, the weighted gauge matrix X is obtained. After that, the positive ideal solution x^* and negative ideal solution x^0 for each attribute are calculated, and the calculated result remains four decimal places, positive ideal solution $x^* = (0.0111, 0.0793, 0.0015, 0.0089)$, and negative ideal solution $x^o = (0.0051, 0, 0.0010, 0)$. Finally, the distance between each attribute in the wireless sensor network and the positive ideal solution and the distance between each attribute and the negative ideal solution are calculated, and the comprehensive measurement value of each attribute is calculated. Finally, the distributed node importance value C_i^o of TOPSIS method based on the entropy weight of multiple attributes is obtained. The importance value and ranking measurement results of each node are shown in Table 2.

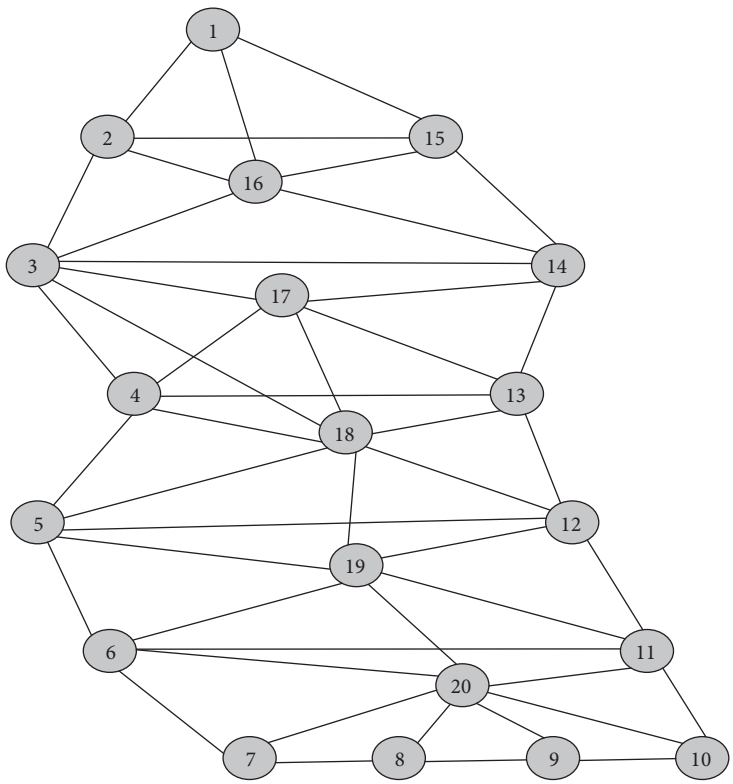
PageRank algorithm is used to calculate the PR value of each node in the wireless sensor network. The node PR value and ranking measurement results of the PageRank algorithm are shown in Table 3.

It can be seen from Tables 2 and 3 that the ranking of distributed nodes obtained by evaluating the importance of distributed nodes from different perspectives is not identical. In order to more comprehensively and more accurately get the importance of distributed node measurement, in the calculation method of TOPSIS attribute based on entropy, this method recalculates the importance ranking index of distributed nodes and the nodes after the PR value. Among them, the normalization processing and PR value are performed based on the Page Rank algorithm, and the proportion is introduced. The coefficients α and β are postprocessing linear regression calculated values. The results accord with the actual circumstances of the wireless sensor network and distributed nodes. The proportionality coefficient is $\alpha = 0.7$, $\beta = 0.3$, after determining the proportionality coefficient; the comprehensive ranking measurement result of node importance of wireless sensor network containing 20 distributed nodes is calculated.

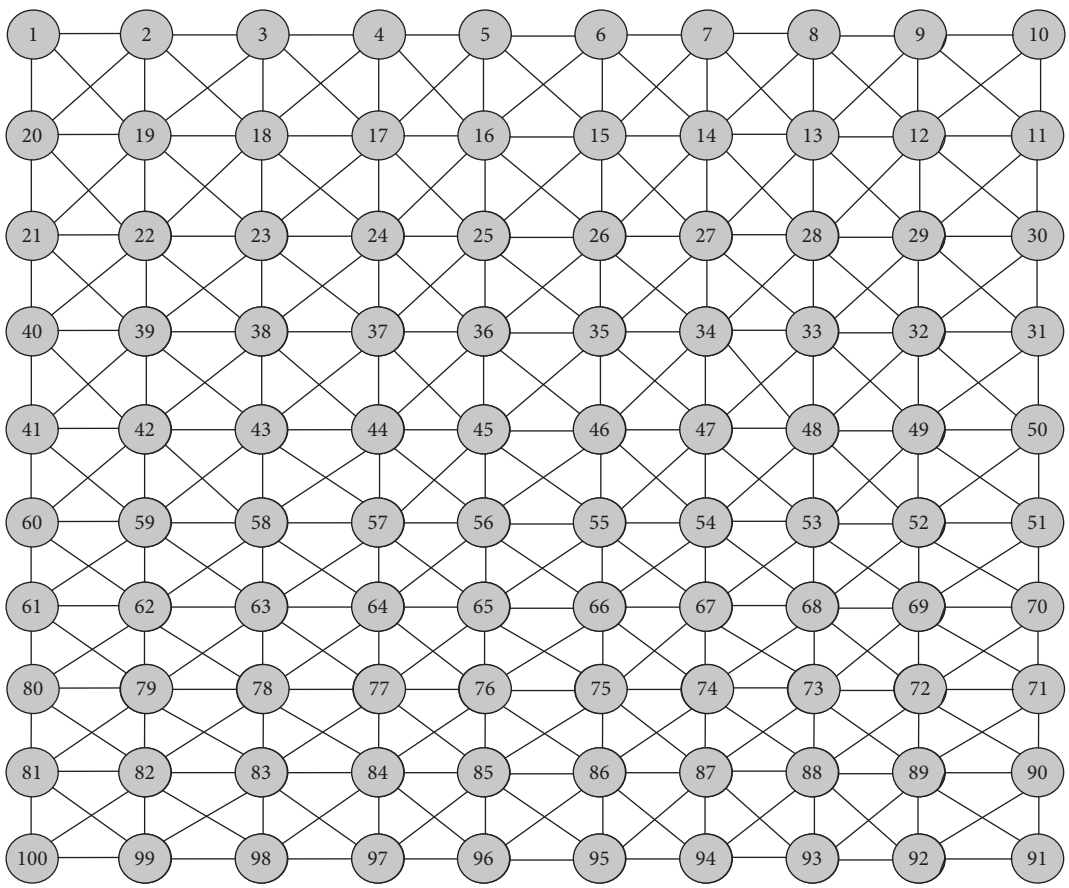
Then, the node ranked first in the measurement result is deleted, and its PR value is evenly distributed to the neighbor nodes connected with it, making it a part of the PR value of neighbor nodes. Then, the remaining deleted nodes were used as a new wireless sensor network. The above steps were repeated to recalculate the attribute value, attribute weight, C_i^o, PR value, and comprehensive ranking C_i of the remaining nodes to obtain the new distributed node ranking measurement result. Then, continue to delete the first ranked node, until there is only one node left in the wireless sensor network. The importance of each distributed node in the wireless sensor network is measured by the method in this paper, and the ranking result of the final measurement is as follows [1–19, 21, 22].

Network efficiency indicates the connectivity of the wireless sensor network. The higher the efficiency of the wireless sensor network is, the better the connectivity is. According to different importance metrics of distributed nodes, the same proportion of distributed nodes are deleted, and the larger the decline rate of network efficiency is, which indicates that the greater the impact of these distributed nodes on network efficiency is, the better the attack effect is, and the corresponding distributed nodes are selected. The point importance measure is more effective. In order to compare and verify the effectiveness of this method, which uses multiple attributes as metrics to measure the importance of distributed nodes in wireless sensor networks, six single attribute metrics, namely, degree centrality, eigenvector centrality, compactness centrality, betweenness centrality, K-kernel decomposition centrality, and aggregation coefficient, and multiple attribute metrics are used to measure the importance of distributed nodes in two wireless sensor networks. In this case, the network efficiency after the deletion of distributed nodes is compared and analyzed. The deletion of distributed nodes in the wireless sensor network is mainly to delete the top 10% sorted according to different metrics, and the proportion of deleted distributed nodes is from 0% to 20%. At the same time, the efficiency decline proportion of the wireless sensor network after the deletion of distributed nodes is calculated. The experimental results are shown in Figure 3.

According to the analysis of Figure 3, with the gradual increase of the proportion of deleted distributed nodes, the decline proportion of network efficiency using single attribute measurement index and multiple attribute measurement index in the two sizes of wireless sensor networks is increasing, and the decline proportion trend of network efficiency of the two types of wireless sensor networks is basically the same; for wireless sensor networks with fewer distributed nodes, the decline proportion of multiple attribute measurement index is increasing. When the proportion of deleting distributed nodes is less than 10% in the wireless sensor network with more distributed nodes, the gap between single attribute metrics and multiple attribute metrics is not obvious, but the overall network efficiency decline proportion of multiple attribute metrics is low. For each single attribute metric, when the



(a)



(b)

Figure 2: Topological structure of wireless sensor network. (a) Wireless sensor network topology with 20 distributed nodes. (b) Wireless sensor network topology with 100 distributed nodes.

Table 1: Attribute values of each distributed node.

Distributed node	Degree of centricity	Centrality of eigenvectors	Compactness centrality	Intermediate centrality	K nuclear decomposition centrality	Aggregation coefficient
1	0.1111	0.1532	0.2643	0.0000	0.1310	0.0000
2	0.2111	0.1885	0.2996	0.1592	0.1663	0.8444
3	0.2111	0.2571	0.3682	0.3269	0.2349	1.1111
4	0.1111	0.3175	0.3286	0.1895	0.2953	1.1111
5	0.1111	0.1843	0.2952	0.1677	0.1621	1.1111
6	0.1611	0.1843	0.2952	0.2124	0.1621	1.1111
7	0.1111	0.1532	0.2643	0.1177	0.1310	1.1111
8	0.1111	0.1288	0.2399	0.0699	0.1066	1.1111
9	0.1111	0.1326	0.2437	0.0729	0.1104	1.1111
10	0.1111	0.1532	0.2643	0.1186	0.1310	1.1111
11	0.1111	0.1857	0.2968	0.1668	0.1635	1.1111
12	0.1611	0.2380	0.3491	0.2817	0.2158	1.1111
13	0.1111	0.2175	0.3286	0.1699	0.1953	1.1111
14	0.2111	0.2175	0.3286	0.1857	0.1953	1.1111
15	0.1611	0.1778	0.2889	0.0572	0.1556	0.6778
16	0.1111	0.1740	0.2851	0.0229	0.1518	1.1111
17	0.1111	0.2077	0.3188	0.0585	0.1855	1.1111
18	0.1111	0.2333	0.3444	0.1298	0.2111	1.1111
19	0.1611	0.2448	0.3559	0.2265	0.2226	1.1111
20	0.1111	0.1885	0.2996	0.0985	0.1663	1.1111

Table 2: Importance values of each node and ranking measurement results.

Distributed node	The importance of value
3	1.1000
12	0.8661
19	0.6951
6	0.6515
4	0.2512
14	0.5694
13	0.5179
5	0.5112
11	0.4985
2	0.4891
18	0.3989
10	0.3619
7	0.3593
20	0.3319
9	0.2578
8	0.2492
17	0.1912
15	0.1744
16	0.1192
1	0.0012

Table 3: Node PR value and ranking measurement results.

Distributed node	PR value
3	0.0711
14	0.0698
2	0.0682
6	0.0619
12	0.0585
19	0.0582
15	0.0524
9	0.0459
8	0.0457
10	0.0452
7	0.0445
11	0.0433
5	0.0429
20	0.0423
4	0.0422
18	0.0396
13	0.0393
17	0.0377
16	0.0371
1	0.0369

proportion of deleted distributed nodes is more than 10%, the network efficiency decline proportion of multiattribute metric is significantly higher than that of single attribute metric; comprehensive analysis shows that the network efficiency decline proportion of multiattribute metric is the smallest. Experiments show that this method uses multiple attributes as the distributed node importance measurement index; compared with each single attribute measurement index, it has a better measurement effect on the wireless sensor network. Through the comparative analysis of the network efficiency decline ratio after the deletion of distributed nodes, the multiple attribute measurement index can better measure the importance of distributed nodes.

3.2. Comparative Analysis. In this paper, the node importance measurement method based on the weighted method (reference [8]) and the node importance measurement method of complex network based on interpretive structure model (reference [9]) are used to measure the importance of distributed nodes in wireless sensor networks. The two methods are recorded as method 1 and method 2. As a comparison method of the methods in this paper, the performance of the three methods under different network connectivity is tested. The test results are shown in Figure 4.

It can be seen from the analysis of Figure 4 that in two-scale wireless sensor networks, the measurement accuracy of the three methods improves with the increase of network

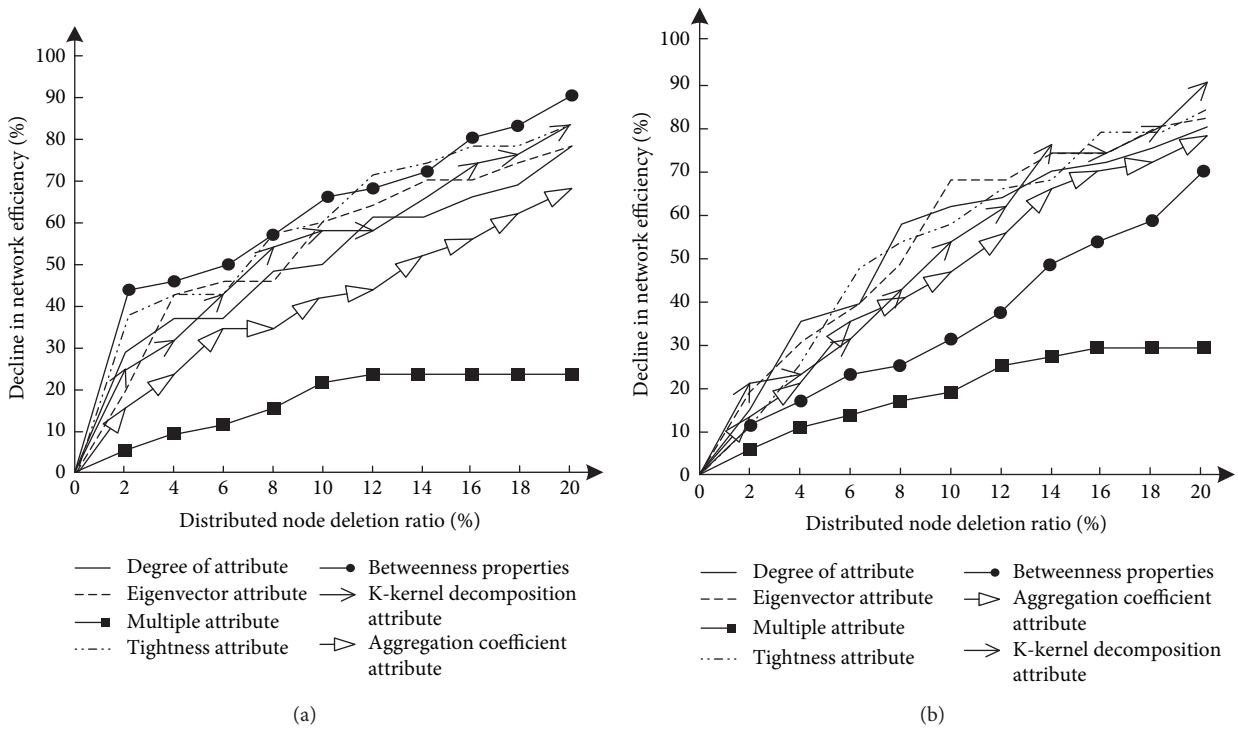


Figure 3: Ratio of efficiency decline of the wireless sensor network. (a) Decline ratio of network efficiency of wireless sensor network with 20 distributed nodes. (b) Decline ratio of network efficiency of wireless sensor network with 100 distributed nodes.

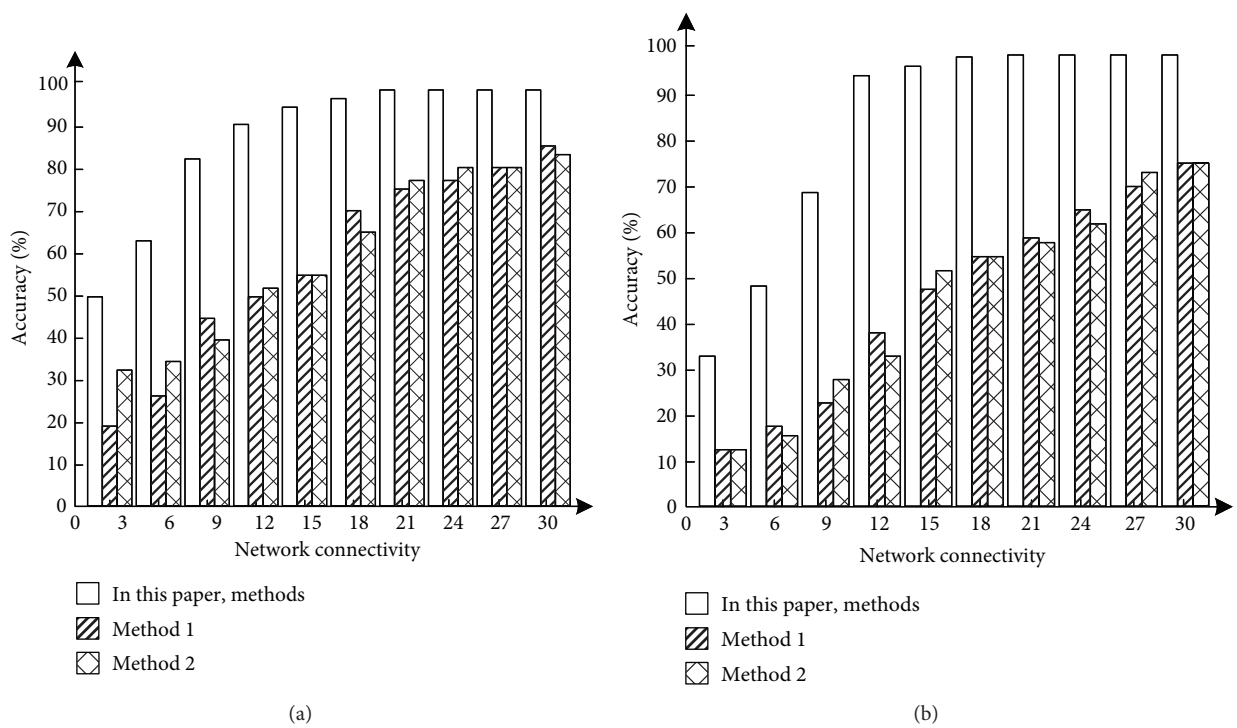


Figure 4: Three methods of the distributed node measurement accuracy. (a) 20 distributed nodes of wireless sensor network distributed node measurement accuracy. (b) 100 distributed nodes of wireless sensor network distributed node measurement accuracy.

connectivity. In two-scale wireless sensor networks, the measurement accuracy of this method is higher than that of the other two methods. In larger scale wireless sensor networks, the measurement accuracy of this method is higher than that of the other two methods. When the network connectivity reaches a certain degree, the

measurement accuracy of this method tends to be stable. In the small-scale wireless sensor network, the measurement accuracy of this method is as high as 96%, and the measurement accuracy of the other two methods is not as high as 90%; in the large-scale wireless sensor network, the measurement accuracy of this method is as high as 98%, and the measurement accuracy of the other two methods is not as high as 90%. The accuracy of the method is about 75%. Experimental results show that the proposed method has the highest accuracy in measuring the importance of distributed nodes in wireless sensor networks.

4. Conclusion

Wireless sensor network integrates sensor technology, microelectronics technology, wireless communication, and network technology. WSN has broad application prospects in the fields of data acquisition, environmental monitoring, target tracking, and many other areas. However, due to the characteristics of self-organization, large number of nodes, limited resources, and constantly changing network topology, node measurement has become a crucial problem to be solved. At present, a variety of measurement methods have been proposed, but these methods often have the disadvantages of low measurement accuracy, high complexity, and weak fault tolerance. In view of the existing importance ranking methods of distributed nodes in wireless sensor networks, such as the one sidedness of evaluation index, the objectivity of attribute weight setting, the neglect of network structure changes, and the influence of important neighbor nodes, a distributed node importance measurement method based on multiple attributes for WSN is proposed. This method uses various attributes to effectively and efficiently avoid the limitations of single attribute measurement; introduces the concept of information entropy into TOPSIS algorithm to weight the multiple attributes of nodes; eliminates the error caused by artificial weight; and uses page for reference. Rank algorithm is improved to make it a part of the importance measurement of distributed nodes; the improved node deletion method and repeated calculation are used to comprehensively measure the importance of nodes, so as to reduce the influence of topology changes and important neighbor nodes and improve the accuracy of the importance measurement of distributed nodes. The experimental results achieved the effectiveness of the proposed research.

Data Availability

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

Conflicts of Interest

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

Acknowledgments

The authors acknowledge Beichuang teaching assistant fund, University Industry Research Innovation Fund of Science and Technology Development Center of Ministry of Education (Grant no. 2018A05031).

References

[1] Y. Guangyou, X. Chenbo, G. Xiong, and Z. Tuo, "Energy efficient node deployment optimization for cts based on cooperative beamforming in single-hop wireless sensor networks," *Automatic Control and Computer Sciences*, vol. 54, no. 2, pp. 147–155, 2020.

[2] D. Peng, J. Zhongliang, S. Kai, and L. Minzhe, "A distributed consensus filter for sensor networks with heavy-tailed measurement noise. Sciece China," *Information Sciences*, vol. 61, no. 11, Article ID 119201, 2018.

[3] S. Dolha, P. Negirla, F. Alexa, and I. Silea, "Considerations about the signal level measurement in wireless sensor networks for node position estimation," *Sensors*, vol. 19, no. 19, pp. 41–79, 2019.

[4] M. Zhang, X. Zhang, and Y. Huang, "Data collection optimization method for wireless sensor networks based on linear regression," *IOP Conference Series Materials Science and Engineering*, vol. 569, Article ID 032064, 2019.

[5] E. H. Houssein, M. R. Saad, K. Hussain, W. Zhu, H. Shaban, and M. Hassaballah, "Optimal sink node placement in large scale wireless sensor networks based on harris' hawk optimization algorithm," *IEEE Access*, vol. 8, no. 99, pp. 19381–19397, 2020.

[6] S. B. Chandanapalli, E. S. Reddy, and D. R. Lakshmi, "Dftdt: distributed functional tangent decision tree for aqua status prediction in wireless sensor networks," *International Journal of Machine Learning and Cybernetics*, vol. 9, no. 9, pp. 1419–1434, 2018.

[7] W. Xuqi, H. Fangyuan, G. Pengfei, and W. Wenqing, "Localization model of wireless sensor network node integrated with constrained convergence distance," *IPPTA: Quarterly Journal of Indian Pulp and Paper Technical Association*, vol. 30, no. 8, pp. 840–845, 2018.

[8] L. Wang, Q. Guo, and J. G. Liu, "Measuring node importance based on weighted nonlinear method," *Application Research of Computers*, vol. 35, no. 5, pp. 1426–1428, 2018.

[9] G. Hu, X. X. Xu, and X. C. Guo, "Importance calculation of complex network nodes based on interpretive structural modeling method," *Journal of Zhejiang University(Engineering Science)*, vol. 52, no. 10, pp. 1989–1997+2022, 2018.

[10] S. S. R. Krishnan and A. Thangavelu, "An early prevention method for node failure in wireless sensor networks," *International Journal of Internet Technology and Secured Transactions*, vol. 10, no. 5, p. 507, 2020.

[11] M. A. Jan, P. Nanda, X. He, and R. P. Liu, "A Sybil attack detection scheme for a forest wildfire monitoring application," *Future Generation Computer Systems*, vol. 80, pp. 613–626, 2018.

[12] Y.-R. Lee, X. Jin, and H.-N. Kim, "Node localization based on neural network using semi-supervised learning in wireless sensor networks," *The Journal of Korean Institute of Communications and Information Sciences*, vol. 44, no. 3, pp. 517–527, 2019.

[13] Y. Huang, H. Wang, and K. Li, "An anchor node selection mechanism-based node localisation for mines using wireless sensor networks," *International Journal of Computational Science and Engineering*, vol. 1, no. 1, p. 1, 2019.

[14] M. A. Jan, M. Usman, X. He, and A. U. Rehman, "SAMS: a seamless and authorized multimedia streaming framework for

WMSN-based IoMT," *IEEE Internet of Things Journal*, vol. 6, no. 2, pp. 1576–1583, 2018.

[15] Z. J. Jiang, T. Zhou, and Y. H. Yang, "Privacy intensity policy for optimizing source node location in wireless sensor networks," *Laser & Optoelectronics Progress*, vol. 57, no. 24, Article ID 241017, 2020.

[16] A. Pang, F. Chao, H. Zhou, and J. Zhang, "The method of data collection based on multiple mobile nodes for wireless sensor network," *IEEE Access*, vol. 8, pp. 14704–14713, 2020.

[17] Y. Cheng, Q. Liu, J. Wang, S. Wan, and T. Umer, "Distributed fault detection for wireless sensor networks based on support vector regression," *Wireless Communications and Mobile Computing*, vol. 2018, Article ID 4349795, 8 pages, 2018.

[18] Y. Tang and Y. Dai, "Wireless strain synchronization acquisition method based on kalman filter," *Journal of Physics: Conference Series*, vol. 1754, no. 1, Article ID 012064, 2021.

[19] F. Shi, "Simulation of information segment fusion point extraction method for wireless sensor network," *Computer Simulation*, vol. 35, no. 4, pp. 262–265, 2018.

[20] X. Yu, Y. Chu, F. Jiang, Y. Guo, and D. Gong, "SVMs classification based two-side cross domain collaborative filtering by inferring intrinsic user and item features," *Knowledge-Based Systems*, vol. 141, pp. 80–91, 2018.

[21] H. Khan, M. A. Jan, M. Alam, and W. Dghais, "A channel borrowing approach for cluster-based hierarchical wireless sensor networks," *Mobile Networks and Applications*, vol. 24, no. 4, pp. 1306–1316, 2019.

[22] X. Yu, F. Jiang, J. Du, and D. Gong, "A cross-domain collaborative filtering algorithm with expanding user and item features via the latent factor space of auxiliary domains," *Pattern Recognition*, vol. 94, pp. 96–109, 2019.

Hindawi
Scientific Programming
Volume 2021, Article ID 5587188, 10 pages
https://doi.org/10.1155/2021/5587188

Research Article

Machine Learning-Based Model to Predict the Disease Severity and Outcome in COVID-19 Patients

Sumayh S. Aljameel,[1] Irfan Ullah Khan,[1] Nida Aslam,[1] Malak Aljabri,[1] and Eman S. Alsulmi[2]

[1]*College of Computer Science and Information Technology, Imam Abdulrahman Bin Faisal University, Dammam, Saudi Arabia*
[2]*Department of Obstetrics and Gynecology, College of Medicine, Imam Abdulrahman Bin Faisal University, Dammam, Saudi Arabia*

Correspondence should be addressed to Sumayh S. Aljameel; saljameel@iau.edu.sa

Received 1 February 2021; Revised 6 March 2021; Accepted 10 April 2021; Published 20 April 2021

Academic Editor: Shah Nazir

The novel coronavirus (COVID-19) outbreak produced devastating effects on the global economy and the health of entire communities. Although the COVID-19 survival rate is high, the number of severe cases that result in death is increasing daily. A timely prediction of at-risk patients of COVID-19 with precautionary measures is expected to increase the survival rate of patients and reduce the fatality rate. This research provides a prediction method for the early identification of COVID-19 patient's outcome based on patients' characteristics monitored at home, while in quarantine. The study was performed using 287 COVID-19 samples of patients from the King Fahad University Hospital, Saudi Arabia. The data were analyzed using three classification algorithms, namely, logistic regression (LR), random forest (RF), and extreme gradient boosting (XGB). Initially, the data were preprocessed using several preprocessing techniques. Furthermore, 10-k cross-validation was applied for data partitioning and SMOTE for alleviating the data imbalance. Experiments were performed using twenty clinical features, identified as significant for predicting the survival versus the deceased COVID-19 patients. The results showed that RF outperformed the other classifiers with an accuracy of 0.95 and area under curve (AUC) of 0.99. The proposed model can assist the decision-making and health care professional by early identification of at-risk COVID-19 patients effectively.

1. Introduction

Coronavirus (COVID-19) started in China in December 2019. As of January 2021, over 95 million cases have been reported around the world, with a mortality rate of 2% of the total closed cases [1]. This rapid pandemic expansion represents a global concern and a serious threat to the public health and economy worldwide. To prevent the infection from spreading, most countries restricted social interaction through precautionary measures such as isolation and quarantine. However, many infected patients did not benefit from the proper treatment due to late diagnosis and the novel and unknown nature of the virus. Recently, many researchers focused on developing new methodologies to screen infected patients in different stages to find notable associations between the patient's clinical features and the chances to succumb to the disease [2, 3]. Current investigation studies determined that artificial intelligence (AI) and machine learning (ML) techniques can play a key role in reducing the effect of the virus spread [4–6]. ML application technologies on patients' data fall under a range of different research directions [7]. One of the most important research directions is predicting the infection rate and mortality rate and building a model to classify patients based on their clinical findings [8, 9]. These research investigations are extremely important and would greatly assist people in the health sectors to be well prepared and take all necessary precautions to minimize the pandemic spread.

The aim of this research is to develop a prediction model to calculate the severity of the disease in COVID-19 patients, using risk factors that can be monitored remotely, with the patient being at home. Moreover, the study explores the impact of vital signs, chronic diseases, preliminary clinical

investigations, and demographic features to predict the survival versus the mortality of COVID-19 patients. The study used COVID-19 patients' data from the King Fahad University Hospital containing the clinical findings and demographic information to validate the model performance and effectiveness. All the risk factors or vital signs that can be measured through widely used sensors were included in the study such as oxygen level in the blood, temperature, pulse rate, and blood pressure. The model will serve as an early warning system to timely identify at-risk patients.

1.1. Related Work. Early detection and diagnosis using AI techniques help to prevent the spread and to combat the COVID-19 pandemic using different data such as CT scans, X-ray, clinical data, and blood sample data.

Yan et al. [10] predicted the criticality and survival chances of patients with severe COVID-19 infection based on different risk factors and demographic information. The dataset used consists of 375 records from patients admitted to Tongji Hospital from January 10th to February 18th, 2020, including 201 survivors and 174 deceased within the same period. They used an XGBoost (XGB) model and identified only three main clinical features as significant, i.e., lactic dehydrogenase (LDH), lymphocyte, and high-sensitivity C-reactive protein (Hs-CRP), selected from more than 300 features. The proposed model was validated using data from 29 patients. The key findings of the research were the model's ability to predict the risk of death with 0.95 precision and 0.90 prediction accuracy. Such models will equip physicians with a tool for identifying critical conditions, thereby helping to reduce the mortality rate. Even though these findings are of great importance, the research has some limitations, which affect the accuracy of the reported results. These limitations were due to the small size of the dataset, namely, 29 records of patients only.

Similarly, Wong and So [11] also used XGB with another dataset to predict the severe and the death cases and identify the risk factors associated with COVID-19. The dataset was retrieved from United Kingdom Biobank (UKBB) and includes 93 different variables collected between 16 March 2020 and 19 July 2020. Two different studies have been conducted based on the sample's groups. For the first study, the data were clinical prediagnostic data of 1747 COVID-19 infected patient records containing both severe and death cases. For the severity class, the accuracy achieved was 0.668, and for the fatality class, the accuracy was 0.712. For the second study, the data were taken from the negative cases, the general population with no COVID-19 infection, consisting of 489987 records. The same model was applied, and the accuracy achieved was similar to the first study, with an accuracy of 0.669 for the severity class and 0.749 for the fatality class, respectively. It is worth mentioning that the researchers identified the five most significant risk factors for severe cases and death cases, with age being the top factor for both cases. Other factors include obesity, impaired renal function, multiple comorbidities, and cardiometabolic abnormalities.

Sun et al. [12] developed a prediction model using the support vector machine (SVM) to predict the severe cases of COVID-19 patients. In the study, they used the clinical and laboratory features that are significantly associated with these cases. Using 336 cases of COVID-19 patients, 26 severe/critical cases and 310 noncritical, they found that the main features to discriminate the mild and severe cases are age, growth hormone secretagogues (GHSs), immune feature cluster of differentiation 3 (CD3) percentage, and total protein. They found that the proposed model was effective and robust in predicting patients in severe conditions with up to 0.775 accuracy.

Another research conducted by Yao et al. [13] also applied the SVM model to classify the COVID-19 patients according to the severity of the symptoms. They applied SVM for the binary class label on a total of 137 records including urine and blood test results and combining both severely ill patients and patients with mild symptoms. The results showed that around 32 factors have high correlations with severe COVID-19, with an accuracy of 0.815. It is worth mentioning that, amongst all factors, age and gender had mostly affected the classification of cases between severe and mild. Patients aged around 65 had more severe cases than others. Moreover, male patients were at a higher risk of developing severe COVID-19 symptoms. In terms of the urine and blood test samples, blood test result features show more significant differences between severe and mild cases than urine test result features.

Hu et al. [14] used the logistic regression (LR) model to identify the COVID-19 patients' severity. They used a dataset containing demographic and clinical data for 115 COVID-19 patients under the nonsevere condition and 68 COVID-19 patients under the severe condition. Four features have been selected as the most significant features to discriminate the mild and severe cases: age, high-sensitivity C-reactive protein level, lymphocyte count, and d-dimer level. This model was evaluated, and the results showed that the prediction was effective with area under the receiver operating characteristic (AUROC) of 0.881, sensitivity of 0.839, and specificity of 0.794, respectively. Bertsimas et al. [15] used 3927 COVID-19 patients' sample for predicting the mortality risk using XGB. The study used demographic and the clinical features of the patients from 33 hospital data. The model achieved the accuracy of 0.85 and AUC of 0.90. Moreover, Sánchez-Montañés et al. [16] developed LR-based mortality prediction using 1969 COVID-19-positive patients. The study found age and O_2 as the significant features and achieved an AUC of 0.89, sensitivity of 0.82, and specificity of 0.81, respectively.

In [5], supervised machine learning techniques have been investigated to predict the COVID-19 outbreak. In [5], SVM has been used for prediction over the dataset obtained from the WHO with 303 patients. The proposed scheme exhibits an accuracy of 0.967 during the testing phase. Similarly, An et al. [17] developed the model to predict the mortality of COVID-19 patients using several machine learning algorithms such as LASSO, SVM (linear and RBF), RF, and KNN. The models were trained to identify three cases, i.e., mortality and survived and mortality and survived

within 14 and 30 days after the initial diagnosis. Linear SVM achieved the highest performance with an AUC of 0.962, sensitivity of 0.92, and specificity of 0.91, respectively. The study found age, diabetes mellitus, and cancer as a significant factor in the mortality prediction for COVID-19 patients.

In conclusion, the importance of machine learning specifically, on predictive analysis, has been proven from several studies. Some of the studies have been conducted to perform the prediction and forecasting, yet there is still a need for further exploration and to extend the findings associated with COVID-19 using a real dataset of clinical records. The summary of the related studies is shown in Table 1. The proposed model in this study attempts to predict and forecast the patients that are at risk along with identifying the main risk factors associated with COVID-19. Targeted patients are isolated at home. The dataset (clinical findings) has been retrieved from King Fahad University Hospital in the Kingdom of Saudi Arabia. The main aim of the study is to develop a preemptive warning model that can identify at-risk COVID-19 patients that are monitored in quarantine at home.

This paper is organized as follows: Section 2 introduces the materials and methods, and Section 3 shows the experimental setup and results. Finally, the conclusion and future work are identified in Section 4.

2. Methodology

The following section covers the dataset description and the methodology used. Due to the class imbalance in the dataset, the synthetic minority oversampling technique (SMOTE) was used.

2.1. Dataset Description. The study was conducted in the Department of Computer Science of Imam Abdulrahman bin Faisal University (IAU) and approved by the Deanship of Scientific Research of IAU under the research grant IRB-2020-09-160. The data were collected from King Fahad University Hospital, Dammam, Kingdom of Saudi Arabia (KSA). The dataset contains the demographic and clinical data of COVID-19-positive patients in the period from 30 April 2020 to 24 July 2020. The dataset contains all the positive patients that were admitted in King Fahad University Hospital during the specified data collection period. There are 287 COVID-19 patient records in the dataset with a binary class label, namely, "survived" and "deceased," respectively. The number of survived patients is 243, and 44 patients deceased. The distribution of instances per class label is shown in Figure 1, while the description of the dataset is mentioned in Table 2. The field BodyTemp 1 in the table indicates the first body temperature taken at the time of the patient's admission to the hospital. However, BodyTemp 2 indicates the last body temperature reading taken before the patient's discharge. Similarly, SOB indicates shortness of breath, chr_dm indicates chronic disease diabetes mellitus, chr_htn indicates hypertension, chr_cardiac represents cardiovascular diseases, chr_dlp represents dyslipidemia, and chr_ckd indicates chronic kidney disease.

The baseline characteristics of the numeric attributes of the dataset are represented in terms of mean ± standard deviation (SD). By contrast, the categorical attributes are measured by a count. The characteristics of the features in the dataset are presented in Table 3.

2.2. Preprocessing. Preprocessing is one of the key steps in data analysis and prediction. Several preprocessing techniques were applied on the dataset. The dataset contains data of all the patients admitted in the hospital. Some symptoms or vital signs occurred with very low frequency and were therefore removed from the dataset. All symptoms with occurrences at 50% or above were selected to be added to the feature set, while the symptoms with occurrences in the range from 2% to 49% were cumulated as one feature the was assigned a unique code. The first three vital signs: fever, cough, and shortness of breath (SOB) were defined as symptom features, while the remaining features were incorporated as a new attribute "sym_others." 5% of the patients in the study were asymptomatic at the time of initial diagnosis and considered as a part of the sym_others attribute. Similarly, the chronic top three (3) diseases (i.e., diabetes, high blood pressure, and cardiac) with the highest frequency were included as features. However, all other chronic disease types with more than 1 occurrence were incorporated as one feature "chr_others." After the initial preprocessing data, an encoding scheme was applied on the categorical features. As the dataset contains a small number of missing values, imputation was performed using the *K*-means technique.

2.3. Prediction Model. In the study, three classification algorithms were used: logistic regression (LR), random forest, and extreme gradient boosting (XGB). A brief description of the classification algorithms is given below.

2.3.1. Logistic Regression. Logistic regression is one of the widely used statistical classification algorithms for binary and multiclass problems. For predicting the probability of the class label, logistic function is used [18]. The functional form of the hypothesis is

$$\mathbf{Y} = \mathbf{C}^{\mathbf{T}}(\mathbf{X}), \tag{1}$$

where C is the list of regression coefficients and X is the list of the features.

$$\mathbf{C} = \begin{bmatrix} \boldsymbol{\beta}_0 \\ \boldsymbol{\beta}_1 \\ \boldsymbol{\beta}_2 \\ \dots \\ \boldsymbol{\beta}_\mathbf{n} \end{bmatrix}, \quad \mathbf{X} = \begin{bmatrix} \mathbf{X}_1 \\ \mathbf{X}_2 \\ \mathbf{X}_3 \\ \dots \\ \mathbf{X}_\mathbf{n} \end{bmatrix}, \tag{2}$$

where $\beta_\mathbf{i}$ represents the regression estimators also known as predicted weights for the selected features in the data and β_0 represents the intercept of the equation.

Table 1: Related studies on mortality prediction for COVID-19 patients.

Reference	Technique	Dataset	Target class	Result
[10]	XGB	404 patients	Death, survived	0.95 precision 0.90 accuracy
[11]	XGB	1747 COVID-19 patients	Fatal, severe	Accuracy 0.668 (fatality) 0.712 (severe)
[12] [13]	SVM SVM	336 COVID-19 patients 137 COVID-19 patients	Severe, critical Severe, nonsevere	0.775 accuracy 0.815 accuracy
[14]	LR	115 COVID-19 patients	Severe, nonsevere	0.881 AUROC 0.839 sensitivity 0.794 specificity
[5]	SVM	303 patients	Negative, positive cases	0.967 accuracy
[15]	XGB	3927 COVID-19 patients	—	0.85 accuracy 0.90 AUC
[16]	LR	1696 COVID-19 patients	Home, deceased	0.89 AUC 0.82 sensitivity 0.81 specificity
[17]	SVM (linear)	8000 COVID-19 patients	Mortality, recovered	0.962 AUC 0.92 sensitivity 0.91 specificity

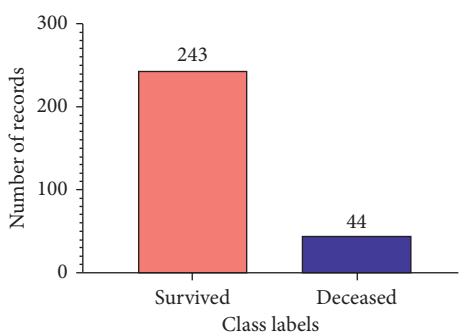


Figure 1: Number of records per class label.

$$H(x) = Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_n x_n. \quad (3)$$

Since the dataset used in the study consists of 25 features in total, the logistic regression algorithm for our study is

$$\mathbf{h}(\mathbf{x}) = (\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_n x_n). \quad (4)$$

The model will predict the record as survived or death if the value of

$$\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_n x_n \geq 0. \quad (5)$$

For optimal selection of regression estimator, maximum-likelihood ratio concept is used.

Sigmoid function (logistic function) is used to map the attributes with the class label. The functional form of the sigmoid equation is given in the following equations:

$$\mathbf{S}(\mathbf{g}) = \frac{1}{(1 + \mathbf{e}^{-y})}, \quad (6)$$

$$\mathbf{S}(\mathbf{g}) = \frac{1}{\left(1 + \mathbf{e}^{-\mathbf{C}^{\mathrm{T}}(\mathbf{X})}\right)}, \quad (7)$$

where e is a numeric constant Euler's number. In LR, a regularization parameter is used to reduce the chance of model overfitting. The logistic regression was optimized using grid search to get hyperoptimized parameters. The parameter set for logistic regression used in our study is shown in Table 4.

2.3.2. Random Forest. Random forest is an ensemble-based classification and regression model initially proposed by Zhang [19]. Random forest can be used for feature selection as well. It uses the bootstrapping data sampling method for partitioning of the data into training and testing sets. The model iteratively generates the trees for every bootstrap. The final prediction is made using the mean vote for each class. It is the combination of all generated decision trees. A decision tree is the hierarchical classification algorithm. The selection of the decision node is made using entropy, information gain, gain ratio, and Gini-index, respectively. In our study, we used information gain and entropy, as shown in the following equations:

$$\mathbf{E}(\mathbf{Y}) = \sum_{\mathbf{i}=1}^{\mathbf{n}} -\mathbf{p_i}\mathbf{log_2}\mathbf{p_i}, \quad (8)$$

$$\mathbf{E}(\mathbf{X}, \mathbf{Y}) = \sum_{\mathbf{n} \in \mathbf{X}} \mathbf{P}(\mathbf{n})\mathbf{E}(\mathbf{n}), \quad (9)$$

where $E(Y)$ represents the entropy of the target, while $\mathbf{Entropy}(\mathbf{X}, \mathbf{Y})$ is the entropy of the attributes with the target, in which $X = \{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n\}$ is the set of attributes in the dataset. The attribute with the highest information gain will be the root attribute, as follows:

$$\text{Information\_Gain} = E(Y) - E(X, Y). \quad (10)$$

It combines the predictions made by multiple trees using randomly selected vectors represented by $\theta_{\mathbf{T}}$. The selected

Table 2: Description of the dataset.

No.	Feature name	UOM	Data type	Missing values
1	Age	Years	Numeric	0
2	Gender	Male/female	Nominal	0
3	BodyTemp (1&2)	Celsius (°C)	Numeric	1%–11%
4	Pulse rate (1&2)	Beats per minute (BPM)	Numeric	7%–5%
5	Resp (1&2)	Breaths per minute (BPM)	Numeric	3%–3%
6	BP_Sys (1&2)	mm Hg	Numeric	10%–7%
7	BP_Dsys (1&2)	mm Hg	Numeric	5%–5%
8	OX (1&2)	mm Hg	Numeric	4%–5%
9	Fever	Yes/no	Nominal	0
10	SOB	Yes/no	Nominal	0
11	Cough	Yes/no	Nominal	0
12	Symptoms_Others	—	Nominal	0
13	chr_dm	Yes/no	Nominal	0
14	chr_htn	Yes/no	Nominal	0
15	chr_cardiac	Yes/no	Nominal	0
16	chr_dlp	Yes/no	Nominal	0
17	Chr_ckd	Yes/no	Nominal	0
18	Chr disease_others	—	Nominal	0

vectors are independent with the previously selected vectors. This results in the collection of trees represented by $h(x)$. The generalization error of decision tree is represented as follows:

$$\mathbf{GE} = \mathbf{P}_{\mathbf{X,Y}} \quad (\mathbf{margin\_fuc}(\mathbf{X}, \mathbf{Y}) < 0), \tag{11}$$

where $\mathbf{P}_{\mathbf{X,Y}}$ is the probability of set of the attributes to map to class label Y.

The parameters used in our study for random forest classifier are shown in Table 5.

2.3.3. Extreme Gradient Boosting. Extreme gradient boosting (XGB) algorithm is an ensemble-based classification and regression technique. It is the regularized form of the gradient boosting algorithm. Gradient boosting algorithm due to the data imbalance sometimes suffers from model overfitting. However, in the XGB algorithm, the regularization parameter reduces the risk the model overfitting. Like random forest, XGB is also a tree-based ensemble classifier. The boosting data resampling method attempts to enhance the model accuracy by minimizing the misclassification error [19]. It is an iterative approach. The records that were not successfully predicted in the previous iteration were used in the next iteration for training the model. The model will repeat the process until the model achieved an optimal result.

The regularization parameter reduces the variance in the model by increasing the weights of the misclassified instances. The increase in weight decreases the model underfitting. However, for reducing the bias of the model, penalty regularization was used to control the model overfitting without leading to a high misclassification rate. The XGB algorithm is the combination of several parameters. The optimal combination of parameters enhances the performance of the model. For parameter optimization, the gird search technique was used. The parameter used in the XGB algorithm is represented in Table 6.

2.4. Performance Evaluation. The performance of the model was evaluated using the standard evaluation measures such as accuracy, precision, sensitivity, specificity, and F-score, respectively. Area under curve and receiver operating characteristic (ROC) were also used for comparing the classifiers. It is one of the widely used tests for exploring the trade-off between true-positive (sensitivity) and false-positive rate (specificity) for the diagnostic test.

$$\mathbf{accuracy} = \frac{\mathbf{TP} + \mathbf{TN}}{\mathbf{TP} + \mathbf{TN} + \mathbf{FP} + \mathbf{FN}}, \tag{12}$$

where the accuracy of the model represents the proportion of the test records that is correctly classified.

$$\mathbf{Sensitivity} = \frac{\mathbf{TP}}{\mathbf{TP} + \mathbf{FN}}. \tag{13}$$

Sensitivity is the proportion of the positive class labels that is correctly predicted. It is also known as the true-positive rate (TPR) or positive-predicted value (PPV).

$$\mathbf{Specificity} = \frac{\mathbf{TN}}{\mathbf{TN} + \mathbf{FP}}. \tag{14}$$

Sensitivity also known as the true-negative rate (TNR) or negative-predicted value (NPV) is the proportion of the negative class labels that are correctly predicted as negative.

$$\mathbf{F-score} = \frac{2 \times \mathbf{precision} \times \mathbf{recall}}{\mathbf{precision} + \mathbf{recall}}, \tag{15}$$

where F-score is the harmonic mean of precision and recall.

3. Experimental Setup and Results

Data imbalance is one of the challenges in data analysis and usually leads to model overfitting. The dataset in this study also suffers from data imbalance as presented in Figure 1. The number of records for the survived category is 243 and for death category is 44. K-nearest neighbor-(KNN-) based synthetic minority oversampling

Table 3: Characteristics of the samples in the dataset.

Feature type	Feature name	Survived ($n = 243$)	Death ($n = 44$)
Demographic	Age (years)	47.28 ± 15.84	59.3 ± 14.4
		Gender	
	Male	159 (65)	38 (86)
	Female	86 (35)	8 (18)
Preliminary investigation	BodyTemp 1	37.1 ± 4.25	33.9 ± 10.7
	Pulse rate-1	94.6 ± 21.6	93.4 ± 31.1
	Resp-1	94.6 ± 21.6	93.4 ± 31.08
	BP_Sys-1	126.45 ± 21.52	128.7 ± 41.27
	BP_Dsys1	78.4 ± 14.9	75.5 ± 26.6
	OX1	93.2 ± 11.9	77.0 ± 24.8
	Temp2	14.8 ± 230.9	0 ± 17.6
	Pulse2	77.1 ± 28.9	68.5 ± 47.7
	Resp2	17.8 ± 6.6	17.8 ± 6.6
	BP_Sys2	109.7 ± 38.9	73.4 ± 55.8
	BP_Dsys2	68.0 ± 24.3	45.2 ± 35.1
	OX2	68.0 ± 24.4	45.2 ± 35.1
Symptoms		Fever	
	Yes	144 (59)	22 (50)
	No	99 (40)	22 (50)
		SOB	
	Yes	111 (45)	31 (70)
	No	134 (55)	13 (30)
		Cough	
	Yes	131 (53)	20 (45)
	No	114 (46)	24 (55)
		Other symptoms	
	Fatigue_weakness	16 (7)	4 (9)
	Sore_throat	10 (4)	3 (7)
	Pain	17 (7)	2 (5)
	Diar	12 (5)	1 (2)
	Anorexia	5 (2)	3 (7)
	Dizz	9 (4)	1 (2)
	Headache	11 (5)	3 (7)
	Nausea	13 (5)	2 (5)
	Vomit	8 (3)	1 (2)
	Dyspnea	13 (5)	2 (5)
	Runny_nose	6 (2)	1 (2)
	Chill	5 (2)	3 (7)
	No	118 (49)	32 (32)

Table 3: Continued.

Feature type	Feature name	Survived ($n = 243$)	Death ($n = 44$)
		chr_dm	
	Yes	73 (30)	23 (52)
	No	170 (69.9)	21 (48)
		chr_htn	
	Yes	67 (27)	18 (41)
	No	176 (72)	26 (59)
		chr_cardic	
	Yes	24 (9.8)	9 (20)
	No	219 (90)	35 (80)
		chr_dlp	
	Yes	25 (10.2)	3 (7)
	No	218 (89.7)	41 (93)
		Chr_CKD	
	Yes	16 (6.5)	5 (11)
	No	227 (93.4)	39 (39)
Chronic disease		Other_ChrDis	
	Epilepsy	4 (2)	1 (2)
	Stroke	4 (2)	3 (7)
	Respiratory	5 (2)	1 (2)
	Bph	3 (1)	2 (5)
	Sle	4 (2)	1 (2)
	Obesity	2 (1)	2 (5)
	Hypothyroidism	5 (2)	1 (2)
	Sickle	4 (2)	1 (2)
	Anemia	4 (2)	1 (2)
	Asthma	2 (1)	3 (7)
	Bone	4 (2)	2 (5)
	Ba	4 (2)	1 (2)
	Dyslipidemia	3 (1)	2 (5)
	Sinusitis	3 (1)	1 (2)
	Dpl	3 (1)	1 (2)
	No	189 (78)	25 (57)

Table 4: Logistic regression parameters using grid search optimization.

Parameter name	Value
Penalty	L2
Random_state	777
Max_iter	10000
Tol	10

Table 5: Random forest parameters using grid search optimization.

Parameter name	Value
Random_state	1
N_estimators	100
Max_depth	15
Min_samples_split	5
Min_samples_leaf	1

Table 6: XGB parameters using grid search optimization.

Parameter name	Value
learning_rate	0.05
max_depth	3
max_features	0.5
random_state	42

technique (SMOTE) was used to alleviate the data imbalance. SMOTE is an algorithm developed by Chawla et al. [20] to overcome the issue of imbalanced datasets in machine learning. In the SMOTE algorithm, the k-nearest neighbor (KNN) is used to calculate the Euclidean distance between the minority class instances to generate new minority class samples in the neighborhood. For A is the minority class with x instances, $A = \{x_1, x_2, \ldots x_n\}$ and k-nearest neighbors of $x_1 = \{x_6, x_7, \ldots x_k\}$ and then A_1 of $x_1 = \{x_7, x_4, \ldots x_n\}$, where $xk€A1$ ($k = 1, 2, 3, \ldots, N$). $x' = x + \text{rand}(0, 1) * |x - xk|$, where x' is the generated point and rand (0, 1) represents the random number between 0 and 1.

The models were implemented in Python language using Jupyter notebook (6.1.4) and sklearn library (0.23.2). For partitioning the data, 10-fold cross-validation technique was used. Experiments were performed on the original dataset and the SMOTE-transformed dataset. Several feature sets were produced using Extratree classifiers with feature importance technique. The set of features was used in the experiments such as all features (25), top 20 features, top 15 features, and top 10 features, respectively. Figure 2 represents the feature ranking, using feature importance, for 20 features.

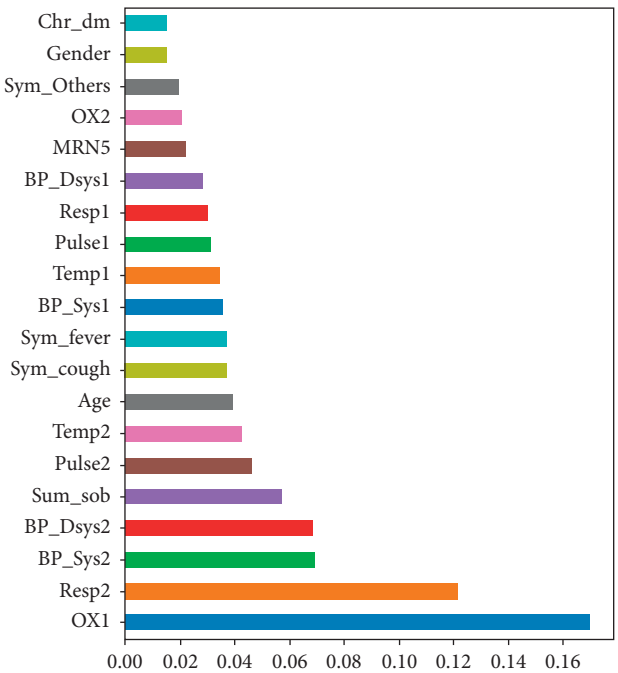


Figure 2: Correlation of top 20 features in the dataset.

The following tables present the performance of the classifiers in terms of accuracy, sensitivity, specificity, and *F*-score. The results showed that random forest outperformed the other models with SMOTE data. Table 7 presents the performance of the classifiers using all features. Table 8 presents the outcome using the top 20 features, Table 9 presents the results with the top 15 features, and Table 10 presents the comparison with the top 10 features, respectively.

Experimental results revealed that random forest outperformed the other classifiers using the top 20 features with SMOTE data with the accuracy of 0.952, sensitivity of 0.949, specificity of 0.956, and *F*-score of 0.955, respectively. Similarly, the AUC-ROC curves for logistic regression, random forest, and extreme gradient boosting are shown in Figures 3, 4, and 5, respectively, using the top 20 features. Random forest achieved the AUC of 0.99. However, the random forest achieved the highest specificity of 1 using the top 15 features.

Logistic regression, on the other hand, underperformed over other classifiers in the top 20, 15, and 10 features using SMOTE data with the accuracy of 0.86, 0.82, and 0.84, respectively. The AUC-ROC curve shows that LR achieved 0.91. However, LR in our study performed better than another study conducted by Yao et al. [13]. They used the LR model to identify the COVID-19 patients' severity and the results achieved an AUC-ROC of 0.881.

A number of studies focused on prediction of severity or mortality have noted that the age is one of the top features that helps to predict the severity of cases [10–13]. In our study, age was ranked among top 10 features across all 25 features used in our prediction model. In addition, our study outperformed other studies that are covered in the literature review with an accuracy of 0.952 and AUC-ROC curve of 0.99.

Table 7: Performance comparison of classifiers using all features (25) using original and SMOTE data.

Classifier	Sampling technique	Accuracy	Sensitivity	Specificity	*F*-score
LR	Without SMOTE	0.874	0.538	0.932	0.56
	With SMOTE	0.753	0.766	0.739	0.766
RF	Without SMOTE	0.908	0.75	0.924	0.6
	With SMOTE	0.938	0.947	0.929	0.941
XGB	Without SMOTE	0.885	0.6	0.922	0.545
	With SMOTE	0.925	0.923	0.926	0.929

Table 8: Performance comparison of classifiers using top 20 features using original and SMOTE data.

Classifier	Sampling technique	Accuracy	Sensitivity	Specificity	*F*-score
LR	Without SMOTE	0.874	0.538	0.932	0.56
	With SMOTE	0.863	0.82	0.93	0.88
RF	Without SMOTE	0.908	0.7	0.935	0.636
	With SMOTE	0.952	0.949	0.956	0.955
XGB	Without SMOTE	0.862	0.5	0.909	0.455
	With SMOTE	0.897	0.878	0.922	0.906

Table 9: Performance comparison of classifiers using top 15 features using original and SMOTE data.

Classifier	Sampling technique	Accuracy	Sensitivity	Specificity	*F*-score
LR	Without SMOTE	0.874	0.583	0.932	0.56
	With SMOTE	0.822	0.793	0.864	0.841
RF	Without SMOTE	0.908	0.7	0.935	0.636
	With SMOTE	0.911	0.856	1	0.922
XGB	Without SMOTE	0.851	0.455	0.908	0.435
	With SMOTE	0.932	0.894	0.984	0.938

This study covers the prediction of the survival and the death of COVID-19-positive patients using demographic, vital signs, and chronic diseases, respectively. The overall result demonstrates the significance of the proposed study with the accuracy of 0.95 and the AUC value of 0.99 using 20 features. The study was performed using a real dataset from the King Fahad University Hospital. Moreover, the dataset

Table 10: Performance comparison of classifiers using top 10 features using original and SMOTE data.

Classifier	Sampling technique	Accuracy	Sensitivity	Specificity	*F*-score
LR	Without SMOTE	0.862	0.5	0.909	0.45
	With SMOTE	0.849	0.867	0.831	0.855
RF	Without SMOTE	0.89	0.63	0.934	0.609
	With SMOTE	0.925	0.884	0.983	0.933
XGB	Without SMOTE	0.851	0.455	0.908	0.43
	With SMOTE	0.89	0.843	0.965	0.904

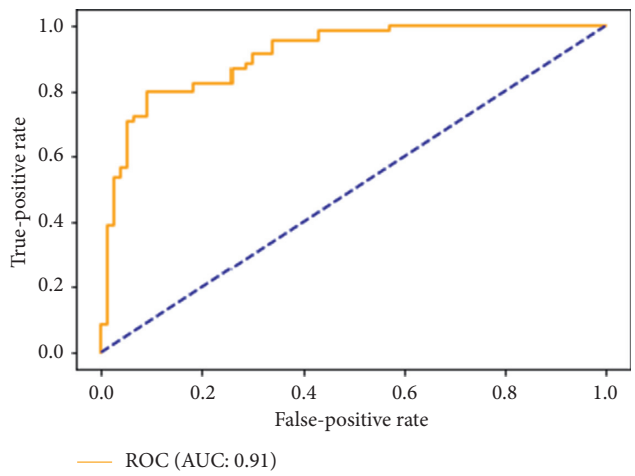


Figure 3: ROC curves of logistic regression using top 20 features.

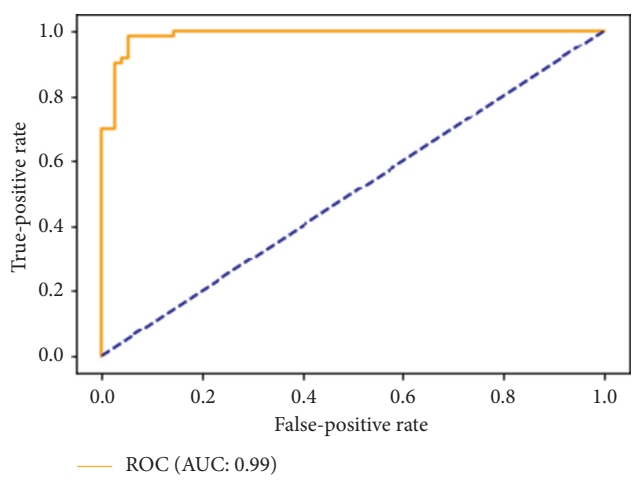


Figure 4: ROC curves of random forest using top 20 features.

contains a very small number of missing data. Despite the several advantages, the study can be further improved by increasing the number of patients. Furthermore, the study needs to incorporate other laboratory tests like lactate dehydrogenase (LDH), neutrophils, lymphocyte, and highly

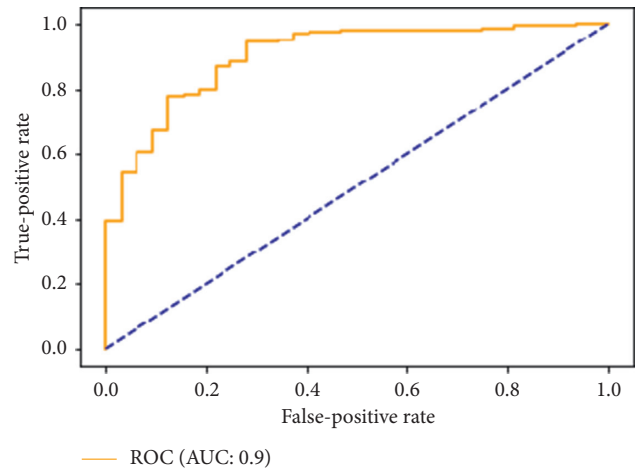


Figure 5: ROC curves of extreme gradient boosting using top 20 features.

sensitive C-reactive protein. Several identified significant features from the literatures need to be included for predicting the mortality risk in COVID-19 patients.

4. Conclusion

The COVID-19 pandemic outbreak has devastated the whole world and lead to a state of worldwide health emergency. Several efforts have been performed to combat this pandemic. In this study, we aimed to explore the impact of vital signs, chronic disease, preliminary clinical data, and demographic features to predict the mortality and survival of the COVID-19 patients using supervised machine learning algorithms. Due to the reduced mortality risk of the COVID-19 cases, the dataset suffers from data imbalance. SMOTE technique was used to alleviate the data imbalance. The results showed that random forest outperformed the other models using 10-fold cross-validation. Grid search technique was applied for parameter optimization. The study achieved the accuracy of 0.952 and AUC of 0.99. Despite the significant outcome achieved from this proposed model, there is still a need for improvement. The models need to be validated using multiple datasets. Furthermore, in the future, we will incorporate and explore the impact of other clinical features and laboratory results that were identified as significant in the previous studies.

Data Availability

The data used to support the findings of this study will be shared upon request to the corresponding author, and the IRB details of the data are available in the paper.

Conflicts of Interest

The authors declare that there are no conflicts of interest.

Acknowledgments

The authors extend their appreciation to the Deputyship for Research & Innovation, Ministry of Education in Saudi

Arabia for funding this research work through the project number Covid19-2020-059-CSIT at Imam Abdulrahman Bin Faisal University/College of Computer Science and Information Technology.

References

[1] "Worldometers-COVID-19 Coronavirus Pandemic." https://www.worldometers.info/coronavirus/?utm_campaign=homeAdvegas1? (accessed January 17, 2020).

[2] C. Gazzaruso, E. Paolozzi, C. Valenti et al., "Association between antithrombin and mortality in patients with COVID-19. A possible link with obesity," Nutrition, Metabolism and Cardiovascular Diseases, vol. 30, no. 11, pp. 1914–1919.

[3] Z. Malki, E.-S. Atlam, A. E. Hassanien, G. Dagnew, M. A. Elhosseini, and I. Gad, "Association between weather data and COVID-19 pandemic predicting mortality rate: machine learning approaches," *Chaos, Solitons & Fractals*, vol. 138, p. 110137, 2020.

[4] A. S. Albahri, R. A. Hamid, J. K. Alwan et al., "Role of biological data mining and machine learning techniques in detecting and diagnosing the novel Coronavirus (COVID-19): a systematic review," *Journal of Medical Systems*, vol. 44, no. 7, p. 122, 2020.

[5] R. Zagrouba, M. Adnan Khan, A. ur-Rahman et al., "Modelling and simulation of COVID-19 outbreak prediction using supervised machine learning," *Computers, Materials & Continua*, vol. 66, no. 3, pp. 2397–2407, 2021.

[6] A. Kumar, A. Sharma, and A. Arora, "Anxious depression prediction in real-time social data," in *Proceedings of the International Conference on Advances in Engineering Science Management & Technology (ICAESMT)-2019*, pp. 1–7, Dehradun, India, July 2019.

[7] S. Lalmuanawma, J. Hussain, and L. Chhakchhuak, "Applications of machine learning and artificial intelligence for Covid-19 (SARS-CoV-2) pandemic: a review," *Chaos, Solitons & Fractals*, vol. 139, p. 110059, 2020.

[8] M. E. H. Chowdhury, T. Rahman, A. Khandakar et al., "An early warning tool for predicting mortality risk of COVID-19 patients using machine learning," 2020, http://arxiv.org/abs/2007.15559.

[9] M. Nemati, J. Ansary, and N. Nemati, "Machine-learning approaches in COVID-19 survival analysis and discharge-time likelihood prediction using clinical data," *Patterns*, vol. 1, no. 5, p. 100074, 2020.

[10] L. Yan, H.-T. Zhang, Y. Xiao et al., "Prediction of criticality in patients with severe Covid-19 infection using three clinical features: a machine learning-based prognostic model with clinical data in Wuhan," *medRxiv*, 2020.

[11] K. C. Y. Wong and H.-C. So, "Uncovering clinical risk factors and prediction of severe COVID-19: a machine learning approach based on UK biobank data," *medRxiv*, 2020.

[12] L. Sun, F. Song, N. Shi et al., "Combination of four clinical indicators predicts the severe/critical symptom of patients infected COVID-19," *Journal of Clinical Virology*, vol. 128, p. 104431, 2020.

[13] H. Yao, N. Zhang, R. Zhang et al., "Severity detection for the coronavirus disease 2019 (COVID-19) patients using a machine learning model based on the blood and urine tests," *Frontiers in Cell and Developmental Biology*, vol. 8, pp. 1–10, 2020.

[14] C. Hu, Z. Liu, Y. Jiang et al., "Early prediction of mortality risk among patients with severe COVID-19, using machine learning," *International Journal of Epidemiology*, vol. 49, no. 6, pp. 1918–1929, 2020.

[15] D. Bertsimas, G. Lukin, L. Mingardi et al., "COVID-19 mortality risk assessment: an international multi-center study," *PLoS One*, vol. 15, no. 12, p. e0243262, 2020.

[16] M. Sánchez-Montañés, P. Rodríguez-Belenguer, A. J. Serrano-López, E. Soria-Olivas, and Y. Alakhdar-Mohmara, "Machine learning for mortality analysis in patients with COVID-19," *International Journal of Environmental Research and Public Health*, vol. 17, no. 22, pp. 8386–20, 2020.

[17] C. An, H. Lim, D.-W. Kim, J. H. Chang, Y. J. Choi, and S. W. Kim, "Machine learning prediction for mortality of patients diagnosed with COVID-19: a nationwide Korean cohort study," *Scientific Reports*, vol. 10, p. 18716, 2020.

[18] R. X. S. D. W. Hosmer and S. Lemeshow, *Applied Logistic Regression*, John Wiley & Sons, Toronto, Canada, 2013.

[19] Y. M. C. Zhang, *Ensemble Machine Learning*, Springer, New York, NY, USA, 2012.

[20] N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, "SMOTE: synthetic minority over-sampling technique," *Journal of Artificial Intelligence Research*, vol. 16, pp. 321–357, 2002.

Hindawi
Scientific Programming
Volume 2021, Article ID 5551255, 9 pages
https://doi.org/10.1155/2021/5551255

Research Article

Research on Grey Relational Clustering Model of Multiobjective Human Resources Based on Time Constraint

Xiaowen Wang and Yijun Mu

School of Economics, Lanzhou University, Lanzhou 730000, China

Correspondence should be addressed to Yijun Mu; muyj16@lzu.edu.cn

Received 2 March 2021; Revised 26 March 2021; Accepted 3 April 2021; Published 17 April 2021

Academic Editor: Shah Nazir

The existing grey relational clustering method has limitations in the application of multidimensional sequences and cannot directly calculate the grey correlation degree between unequal-length sequences. In this paper, by introducing the multidimensional dynamic DTW distance into the existing 3D grey relational model, a new grey relational analysis model that can be applied to multidimensional data is proposed, which is based on DTW distance. The model does not require one-to-one correspondence of data points but evaluates the similarity of its geometric curves by calculating the shortest distance between sequences. In addition, since the traditional grey correlation clustering method is implemented, the method first extracts the reference sequence from the observation sequence and then calculates the similarity between the observation objects by calculating the grey correlation degree between each sequence and the reference sequence, so each object only needs to be calculated once. The experimental results show that the multidimensional grey correlation degree based on DTW distance and the grey relational clustering model oriented to multidimensional data are more accurate than other existing methods. Finally, the grey relation clustering method of multidimensional data is used to analyze the multiobjective human resource grey relational clustering model under time constraints, and the validity of the model is verified.

1. Introduction

Clustering, as an important data analysis method, has attracted the extensive attention of many scholars and is a research hotspot in machine learning, statistics, computer science, and other fields. Most of the existing clustering methods are aimed at the data with massive information [1]. For those data with a small sample size, insufficient information, and unclear sample rules, these methods often cannot get accurate results. The main object of grey relational clustering is such data [2]. It measures the similarity of the observation systems by a grey relational analysis model and simplifies the complex system by dividing similar objects into the same class. In recent years, many scholars have improved and optimized it and applied it to different fields. Grey relational analysis and its model were proposed by Professor Deng Julong in 1985 [3]. As an important research field in grey system theory, it mainly judges the similarity between two sequences according to the geometric shape of the sequence curve [4]. Linear interpolation is usually used to transform the observed data of the discrete behavior of the observation system into piecewise continuous lines, and then a corresponding model is constructed according to the geometric characteristics of the lines to judge the similarity between the sequences, including the characteristics of distance, area, and slope [5]. The more similar objects are observed, the more similar their geometric characteristics are. Because the model is not affected by the sample size and the rule of the sample distribution. In addition, the model also has the characteristics of small computation and convenient application [6].

Most of the existing grey clustering methods can be used to deal with one-dimensional data, so the research objects are limited [7]. To broaden the application scope of grey clustering, some scholars tried to combine grey theory with panel data and put forward the grey clustering method under

panel data [8]. By combining the model with the traditional grey relational clustering method, good clustering results are achieved [9]. By using the principal component analysis method, the grey clustering coefficient matrix at different times is obtained by calculating the correlation degree between the score sequences of the principal components of two comprehensive factors [10]. However, the above method only considers the fluctuation of the indexes between adjacent objects, and if the panel data takes different time, redundant data often need to be supplemented or deleted, so the original data will be destroyed. In addition, some methods also have the problem of the small correlation between indexes in the class [10]. Aiming at these problems, some scholars also put forward a grey index correlation clustering model [11]. The model extracts the feature information from the original sequence by constructing the generating sequence and realizes the dimensionality reduction of the sequence [12]. The clustering rules of panel data can effectively avoid clustering sequences with small association degree into a group and can deal with unequal data. However, this method is easy to be affected by dimensionality reduction results, and the introduction of new uncertainties will affect the final clustering accuracy [13].

2. Basic Principles of Grey Relational Analysis and Cluster Analysis

2.1. Basic Principles of Grey Relational Analysis. We should analyze its components, distinguish the behavior factors and related factors, and make clear the characteristics of the system behavior; then, we should focus on the relationship between the system behavior and related factors; and finally, we should estimate and predict the system behavior [14]. When describing the system behavior and its related factors, that is, the feature mapping quantity of the system behavior and related factors, the feature mapping quantity is used to indirectly reflect the system behavior and related factors [15].

Definition 1. The sequence of behavioral characteristic data of behavioral factors $X_0 = (x_0(1), x_0(2), ..., x_0(n))$ is assumed. The sequence of behavioral characteristic data of related factors is as follows [16]:

$$\begin{aligned} X_1 &= (x_1(1), x_1(2), ..., x_1(n)), \\ X_i &= (x_i(1), x_i(2), ..., x_i(n)), \\ X_m &= (x_m(1), x_m(2), ..., x_m(n)). \end{aligned} \tag{1}$$

The main calculation steps of Deng's grey correlation are [17] as follows:

Step 1: order:

$$\begin{aligned} X_i' &= \frac{X_i}{x_i(l)} = (x_i'(1), x_i'(2), ..., x_i'(n)), \\ i &= 0, 1, 2, ..., m. \end{aligned} \tag{2}$$

Step 2: the difference sequence. Remember:

$$\begin{aligned} \Delta_i(k) &= \left|x_0'(k) - x_i'(k)\right|, \\ \Delta_i &= (\Delta_i(1), \Delta_i(2), ..., \Delta_i(n)), \quad i = 1, 2, ..., m. \end{aligned} \tag{3}$$

Step 3: remember:

$$\begin{aligned} M &= \max_i \max_k \Delta_i(k), \\ m &= \min_i \min_k \Delta_i(k). \end{aligned} \tag{4}$$

Step 4: correlation coefficient:

$$\begin{aligned} \gamma_{0i}(k) &= \frac{m + \xi M}{\Delta_i(k) + \xi M}, \quad \xi \in (0, 1), \\ k &= 1, 2, ..., n; i = 1, 2, ..., m. \end{aligned} \tag{5}$$

Step 5: calculate relevance:

$$\gamma_{0i} = \frac{1}{n} \sum_{k=1}^{n} \gamma_{0i}(k), \quad i = 1, 2, ..., m. \tag{6}$$

2.2. The Basic Principle of Grey Clustering Analysis Method

Definition 2. There are n clustering objects and m evaluation index values, and the j index is divided into s grey class, which is called the subclass of j index [18].

The s subclasses of each index have their own functions. Through these functions, different clustering objects are classified according to the values of the index. This function is called the whitening weight function [19]. A schematic diagram of a typical whitening weight function is shown in Figure 1.

The schematic diagram of whitening functions in four forms is shown in Figure 2.

(1) The typical whitening weight function can be described as a piecewise function:

$$f_j^k(x) = \begin{cases} 0, & x \notin \left[x_j^k(1), x_j^k(4)\right], \\ \dfrac{x - x_j^k(1)}{x_j^k(2) - x_j^k(1)}, & x \in \left[x_j^k(1), x_j^k(2)\right], \\ 1, & x \in \left[x_j^k(2), x_j^k(3)\right], \\ \dfrac{x_j^k(4) - x}{x_j^k(4) - x_j^k(3)}, & x \in \left[x_j^k(3), x_j^k(4)\right]. \end{cases} \tag{7}$$

(2) The lower bound measure whitening weight function can be described as a piecewise function:

$$f_j^k(x) = \begin{cases} 0, & x \notin \left[0, x_j^k(4)\right], \\ 1, & x \in \left[0, x_j^k(3)\right], \\ \dfrac{x_j^k(4) - x}{x_j^k(4) - x_j^k(3)}, & x \in \left[x_j^k(3), x_j^k(4)\right]. \end{cases} \tag{8}$$

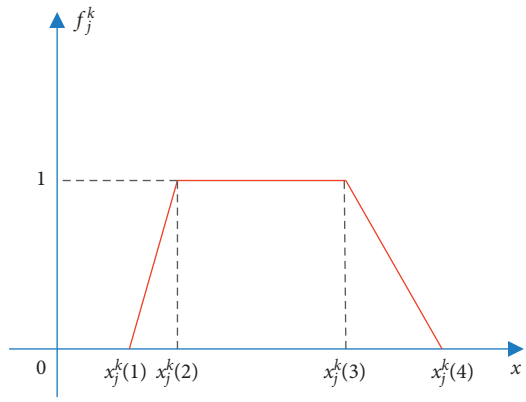


Figure 1: Schematic diagram of a typical whitening weight function.

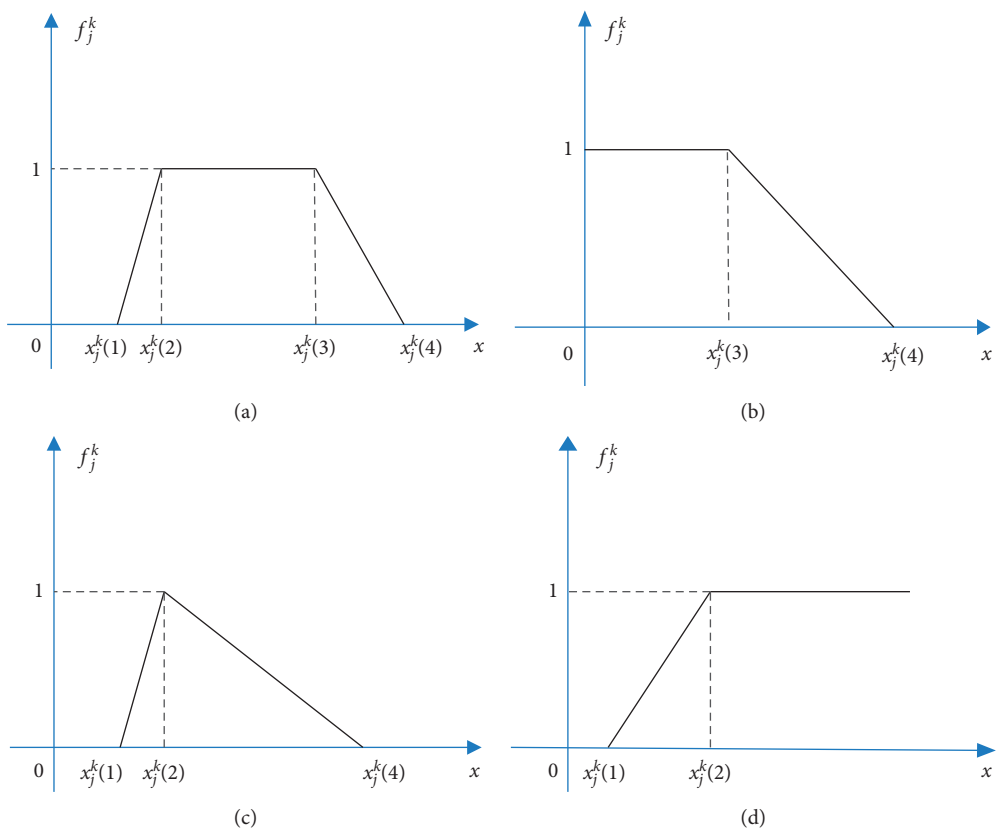


Figure 2: Sketch map of four whitening weight functions. (a) Typical whitening weight function. (b) Lower bound measure whitening weight function. (c) Moderate measure whitening weight function. (d) Upper bound whitening weight function.

(3) Moderate measure whitening weight function can be described as a piecewise function:

$$f_j^k(x) = \begin{cases} 0, & x \notin \left[x_j^k(1), x_j^k(4)\right], \\ \dfrac{x - x_j^k(1)}{x_j^k(2) - x_j^k(1)}, & x \in \left[x_j^k(1), x_j^k(2)\right], \\ \dfrac{x_j^k(4) - x}{x_j^k(4) - x_j^k(2)}, & x \in \left[x_j^k(2), x_j^k(4)\right]. \end{cases} \tag{9}$$

(4) The upper bound measure whitening weight functions can be described as piecewise function:

$$f_j^k(x) = \begin{cases} 0, & x < x_j^k(1), \\ \dfrac{x - x_j^k(1)}{x_j^k(2) - x_j^k(1)}, & x \in \left[x_j^k(1), x_j^k(2)\right], \\ 1, & x \geq x_j^k(2). \end{cases} \tag{10}$$

Panel data is an extended form of cross-sectional data in time dimension, which is composed of three dimensions: time, index, and object. The description of the panel data is shown in Table 1. Because panel data can better reflect the change of different attributes over time than cross-sectional data, panel data has more important research value [20].

3. Grey Relational Clustering Model for Multidimensional Data

3.1. Multidimensional DTW Distance and Its Multidimensional Grey Relational Analysis Model. Some scholars have extended the application of DTW distance to measure the similarity of multidimensional time series. There are two main research directions of multidimensional DTW. One is to convert multidimensional sequences into one-dimensional sequences and calculate the DTW distances between sequences. If the traditional dynamic time warping distance is used to process multidimensional sequences, the simplest way is to apply it to each dimension separately [19, 20]. The distance matrix is constructed by calculating the sum of the distances of each dimension of the sequence, and then the shortest curved path between the sequences is obtained from the distance matrix to match it. Before calculating the DTW distance between sequences, it is necessary to normalize the points of each dimension in the sequence, such as calculating its zero mean or variance [21].

Definition 3. Multidimensional dynamic time warping distance.

Given two sequences $t \in R^{K\times L_t}$ and $r \in R^{K\times L_r}$, assuming K is the number of attributes or dimensions of a data, tL and rL are the sequence lengths of the data sequences t and r, respectively. The first step is to preprocess the sequence, which is to calculate the zero mean or unit variance of each dimension of the sequence. The second step is to use the following formula to calculate the distance matrix:

$$d(i, j) = \sum_{k=1}^{k} |t(k, i) - r(k, j)|. \tag{11}$$

Using the above formula, the vector normal distance between the points of the sequence is calculated and the corresponding distance matrix is constructed. Then, the shortest path between the sequences is obtained from the matrix. The DTW distance between sequences t and r is

$$D(i, j) = d(i, j) + \min \begin{cases} D(i-1, j), \\ D(i, j-1), \\ D(i-1, j-1). \end{cases} \tag{12}$$

DTW distance considers a sequence as a point in m-dimensional space and then extracts a reference sequence from the original sequence so that the correlation between the two sequences depends not only on themselves but also on the sequence of other factors [21]. The three-dimensional grey relational degree model obtains the environmental parameters between two sequences by calculating the Min$h$ space distance of the corresponding time between two sequences. However, this leads to the value of environmental parameters sometimes greater than the distance between the two sequences, making the grey correlation degree greater than 1. At this point, the model cannot satisfy the four axioms of relevance. The reference sequences are selected as the maximum and minimum differences of the two poles of each sequence based on the multidimensional grey correlation degree of DTW distance. Among them, the minimum difference between the two poles is always less than the shortest distance between the two sequences, which guarantees the normality of the grey correlation degree. Finally, by introducing multidimensional DTW distance into Deng's grey relational degree calculation formula, a multidimensional grey relational analysis method based on DTW distance is proposed. The improvement idea of this method is shown in Figure 3. The figures contain three-dimensional grey relational grade on one side and multidimensional grey relational grades based on dynamic time warping distance on the other side [22].

The feasibility and validity of the multidimensional grey correlation degree based on DTW distance are verified by the corresponding grey correlation analysis of the data. The experimental procedures are as follows:

(1) Step 1: get the maximum value of each dimension of each time, divide the value of the corresponding time of each observation factor by this value, and then get the corresponding initial value image.

(2) Step 2: define parameter values.

(3) Step 3: calculate the extremum.

(4) Step 4: calculate the mingling distance and the multidimensional DTW distance between the observed objects and the reference sequences.

(5) Step 5: calculate the three-dimensional grey correlation degree of the observation objects and the multidimensional grey correlation degree based on DTW distance, respectively.

(6) Step 6: use the result of grey correlation degree to rank or cluster the observed objects.

The results of the multidimensional grey relational analysis model based on DTW distance are accurate and consistent when dealing with equal-length and unequal-length sequences. When the three-dimensional grey relational analysis model deals with unequal-length sequences, the results of correlation degree obtained by different data pretreatment methods are also different. When the missing data is filled with 0, the result of the correlation is relatively large. When the sequence is filled with the mean of adjacent data, because the new sequence obtained at this time is similar to the original sequence, the result is more accurate in economic benefit evaluation. However, when the model is clustered in five cities of Hunan Province, the result of the correlation is inaccurate. Three-dimensional grey relational grade and multidimensional grey relational grade calculation based on DTW distance are shown in Table 2.

Table 1: Panel data description form.

	t_1	...	t_k	...	t_T
	$u_1,\ldots,u_j,\ldots,u_m$	...	$u_1,\ldots,u_j,\ldots,u_m$	...	$u_1,\ldots,u_j,\ldots,u_m$
S_1	$x_{11}(t_1),\ldots,x_{1j}(t_1),\ldots,x_{1m}(t_1)$	...	$x_{11}(t_k),\ldots,x_{1j}(t_k),\ldots,x_{1m}(t_k)$	...	$x_{11}(t_T),\ldots,x_{1j}(t_T),\ldots,x_{1m}(t_T)$
	...	...	...	...	...
S_i	$x_{i1}(t_1),\ldots,x_{ij}(t_1),\ldots,x_{im}(t_1)$	...	$x_{i1}(t_k),\ldots,x_{ij}(t_k),\ldots,x_{im}(t_k)$	...	$x_{i1}(t_T),\ldots,x_{ij}(t_T),\ldots,x_{im}(t_T)$
	...	...	...	...	...
S_n	$x_{l1}(t_1),\ldots,x_{lj}(t_1),\ldots,x_{lm}(t_1)$	...	$x_{n1}(t_k),\ldots,x_{nj}(t_k),\ldots,x_{nm}(t_k)$	...	$x_{n1}(t_T),\ldots,x_{nj}(t_T),\ldots,x_{nm}(t_T)$

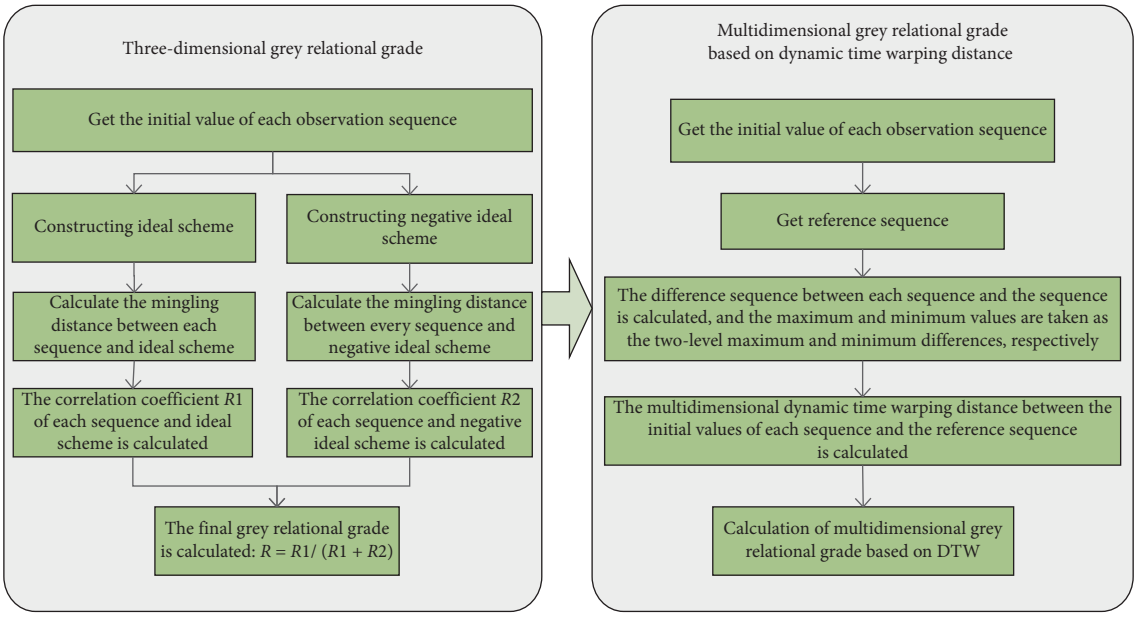


Figure 3: Flow chart of multidimensional grey relational analysis method based on DTW distance.

Table 2: Three-dimensional grey relational grade and multidimensional grey relational grade calculation based on DTW distance.

Grey relational model	Raw data set	Missing data set
Multidimensional grey relational grade based on DTW distance	$S2 > S3 >> S4$	$S2 > S3 >> S4$
Three-dimensional grey relational grade	$S2 > S3 > S1 > S5 > S4$	Missing data complement 0; $S2 > S1 > S5 > S4 > S3$ missing data complement mean
Accurate evaluation results		$S2 > S3 >> S4$

3.2. Grey Relational Clustering for Multidimensional Data. To realize the grey relational clustering of multidimensional data, a grey relational clustering method for multidimensional data is proposed based on the idea of extracting reference sequences from original sequences by referring to the three-dimensional grey relational analysis model. Using this method, each sequence only needs to calculate the grey correlation degree once, instead of comparing two or more sequences like the traditional grey correlation clustering method. Therefore, the calculation process of the method is more convenient and fast. In addition, this method does not require the same length of multidimensional sequence data, compared with the traditional grey relational clustering method, and its application scope is wider. The flow chart of the multidimensional grey relational clustering method is shown in Figure 4.

Definition 4. Grey relational clustering of multidimensional data.

Given the sequence group $X_1, X_2, \ldots, X_m$, suppose any sequence can be represented as $X^i = \{\{x_1^i(1), \ldots, x_1^i(n)\}, \{x_2^i(1), \ldots, x_2^i(n)\}, \ldots, \{x_m^i(1), \ldots, x_m^i(n)\}$, where $x_i(j)$ represents the j data of dimension i. The grey correlation degree between each sequence is calculated. According to these correlation values, multiple critical intervals are manually determined in interval $[0, 1]$. The flow chart of the multidimensional grey relational clustering method is shown in Figure 4.

Because the method does not need to preprocess the sequence to obtain the same length sequence as the traditional grey relational clustering method, it will not introduce new uncertainties into the data, so it can get better clustering results. In addition, because the traditional grey relational

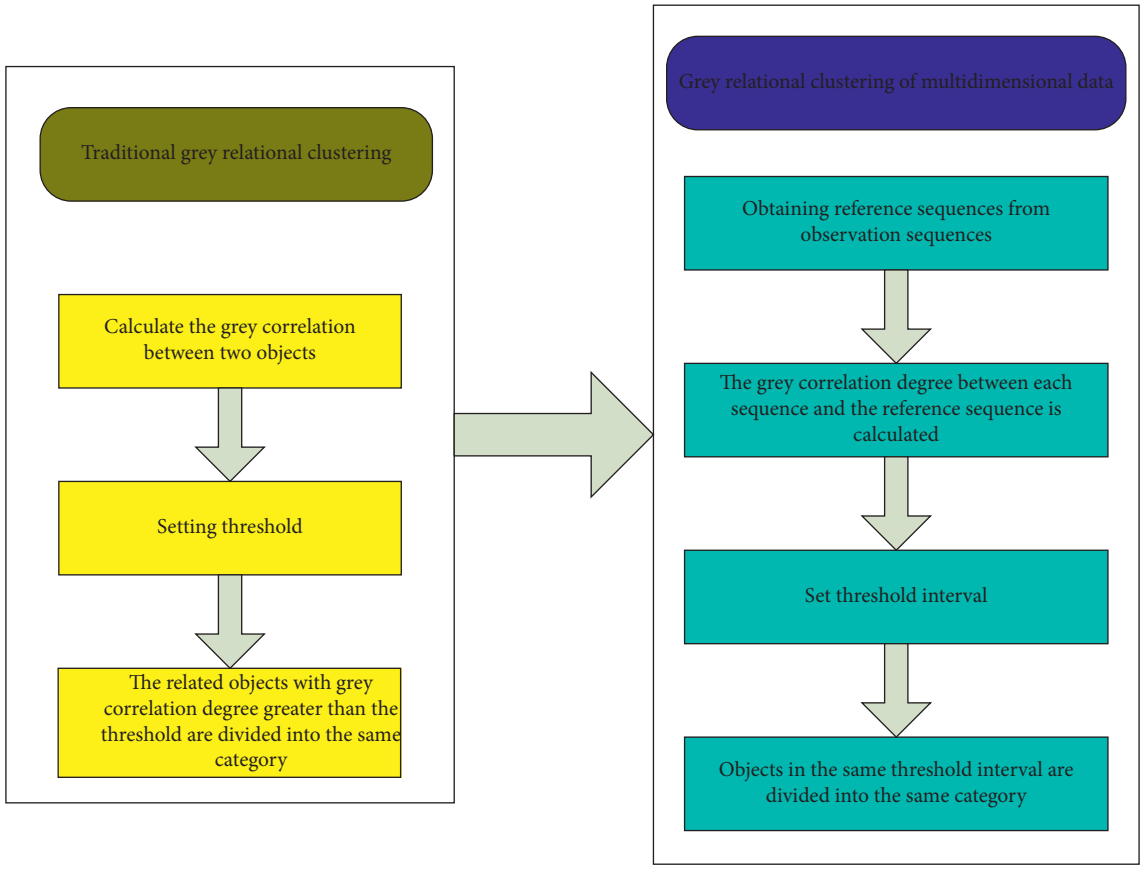


Figure 4: Flow chart of the multidimensional grey relational clustering method.

clustering method needs to compare all objects in pairs when clustering, when there are more data objects, the amount of calculation will be very large. However, this clustering method does not need two-to-two comparisons between sequences, each object only needs to be computed once, and the calculation process is more convenient.

4. Experimental Analysis and Application

4.1. Experimental Analysis of Grey Relational Clustering Model for Multidimensional Data. Because the traditional grey relational clustering method needs to compare two sequences, its time complexity is high. To evaluate the accuracy of the grey relational clustering method for multidimensional data, the clustering experiments of Iris and Wine data sets in UCI data set were carried out, and the clustering effects of the two methods were evaluated by Rand Index [22].

The corresponding experimental steps are as follows:

(1) Step 1: get the maximum value of each time dimension of the sequence, divide the value of the corresponding time of each sequence by this value, and then get the corresponding initial value image.

(2) Step 2: the grey relational clustering method of multidimensional data uses the initial value image and selects the minimum or maximum values of each dimension at the corresponding time to form a reference sequence according to the characteristics of the data set.

(3) Step 3: calculate the initial sequence values separately.

(4) Step 4: calculate the mingling distance and the multidimensional dynamic time warping distance between each sequence and the reference sequence.

(5) Step 5: according to the third and fourth steps, the multidimensional grey scale correlation of DTW is calculated.

(6) Step 6: set the threshold interval, and the sequences in the same interval are divided into the same category.

Grey relational clustering and comparison of traditional methods for multidimensional data are shown in Table 3. Grey relational clustering of multidimensional data and Rand Index value of traditional grey relational clustering are shown in Figure 5.

As can be seen from Table 3, the grey correlation clustering of multidimensional data is improved by nearly 30% compared with the traditional grey correlation clustering, and the clustering result is more accurate. In addition, the traditional grey relational clustering method needs two-to-two comparisons of sequences, and the grey

Table 3: Grey relational clustering and comparison of traditional methods for multidimensional data.

Data set name	Multidimensional grey relational clustering	Traditional grey relational clustering
Iris data set	0.9978	0.6983
Wine data set	0.9598	0.6434

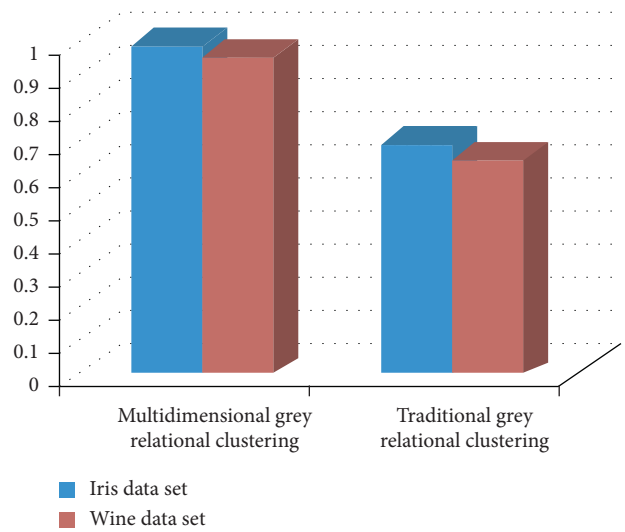


Figure 5: Grey relational clustering of multidimensional data and Rand Index value of traditional grey relational clustering.

relational clustering method of multidimensional data only needs to calculate the grey relational degree for each object in clustering, so the calculation process is more convenient.

4.2. Application of Grey Relational Clustering Method for Multidimensional Data. In this paper, the UCI data set "individual household electric power consumption" is used to design the experiment, and the grey relational clustering method for multidimensional data is used to realize the clustering of similar months of electricity consumption, and the household's electricity consumption law for four years is analyzed.

The experiment takes the monthly electricity consumption as the object of observation with 12 objects. However, because the original data has more than 2 million, and the grey system research object is "small sample, poor information" uncertain data. Therefore, we need to preprocess the data and extend the sampling interval from one minute to one day. Therefore, the sequence length of each object is the number of days observed by electricity. However, the number of days per month is different, resulting in different lengths of data for each object in the sequence. Because the three-dimensional grey relational model needs to calculate Minh's distance between sequences, it cannot deal with sequences of different lengths. Therefore, we need to delete long sequences before using this method. Multidimensional grey relational grade based on DTW distance is not needed.

Three-dimensional grey relational analysis model and its multidimensional grey relational analysis model clustering results based on DTW distance are shown in Table 4. The comprehensive score of fiscal input and output factors of science and technology is shown in Figure 6.

To more intuitively reflect the content of Table 4, the final clustering results are displayed in Figure 6. From the analysis of household electricity consumption based on the multidimensional grey relational analysis model of DTW distance, the clustering results in each month of the four years have little difference, but the clustering results obtained by the three-dimensional grey relational analysis model have great changes.

In view of the limitation of the existing grey relational clustering methods in the application of multidimensional sequences and not being able to directly calculate the grey relational degree between unequal-length sequences, in this paper, by introducing the multidimensional dynamic DTW distance into the existing three-dimensional grey relational model, a new grey relational analysis model, multidimensional grey relational degree based on DTW distance, is proposed, which can be applied to multidimensional data. The model does not need data point-to-point correspondence but evaluates the similarity of geometric curves by calculating the shortest distance between sequences. Finally, the grey relational clustering method of multidimensional data is used to analyze the multiobjective grey relational clustering model of human resources under time constraints, and the validity of the model is verified.

Table 4: Three-dimensional grey relational analysis model and its multidimensional grey relational analysis model clustering results based on DTW distance.

Year	Grey relational analysis values	Three-dimensional grey relational analysis model
2014	{13}; {2}; {7,5}; {3,12,4,11}; {6,9,10}; {8}	{2}; {13}; {3,4}; {7,10,5}; {12}; {6,11}; {8,9}
2015	{5}; {2}; {9}; {12,3,13,11,7}; {4,6,10}; {8}	{5}; {9}; {2}; {8,7}; {3,12,13}; {10,4}; {6,11}
2016	{12}; {13,10,11,4,2}; {9}; {5}; {6,7,3}; {8}	{13}; {10,11}; {2,4,12}; {3,6,5}; {7}; {9,8}
2017	{6,5}; {2}; {11,7,9,10,4}; {3,12}; {8}	{11}; {2}; {4,7,6}; {3,12}; {10,9,5}; {8}

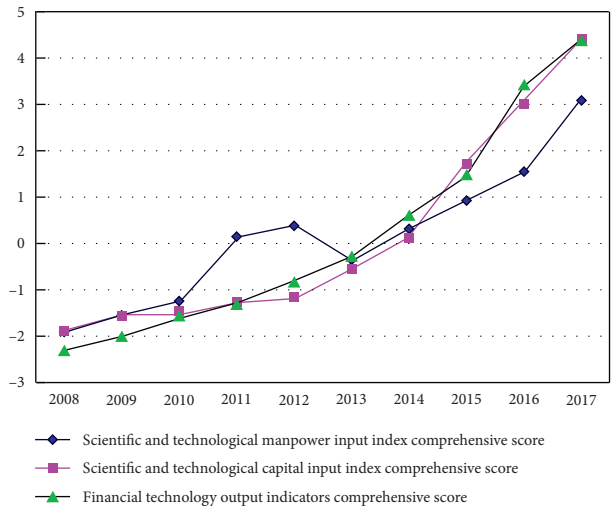


Figure 6: Comprehensive score of fiscal input and output factors of science and technology.

5. Conclusion

The existing models still have limitations in the application of uncertain data and multidimensional data. As the data is collected in real life, there will be a possibility of data missing, and the existing grey relational analysis model requires that the sequence length is equal in application. If the length of the sequence is not equal, the missing data need to be filled or long sequences deleted to preprocess data, which will increase the uncertainty of the original sequence, so the accuracy of the grey relational degree is not enough, thus affecting the accuracy of the final clustering results. In addition, the characteristics of many objects need to collect multiple dimensional data to reflect their characteristics. However, most of the existing grey relational analysis models are based on one-dimensional data. By introducing the multidimensional DTW distance into the three-dimensional grey relational analysis model, the problem that the three-dimensional grey relational analysis model cannot deal with the uncertain sequence directly is solved. Compared with the traditional three-dimensional grey relational analysis model, the experimental results show that the accuracy of this method is higher than that of the three-dimensional grey relational analysis model in dealing with unequal and equal-length sequences. On the basis of multidimensional grey relational degree based on DTW distance and traditional grey relational clustering method, a new grey relational clustering method is proposed. By improving the traditional grey relational clustering method, this method can be applied to multidimensional sequence clustering. The effectiveness of the proposed approach is demonstrated through experimental results.

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] Y. S. Yang, C. Y. Shih, and R. F. Fung, "Multi-objective optimization of the light guide rod by using the combined Taguchi method and Grey relational approach," *Journal of Intelligent Manufacturing*, vol. 25, no. 1, pp. 99–107, 2014.

[2] S. S. Anand and D. S. Vasudev, "Multi-objective optimization in WEDM of D3 tool steel using integrated approach of Taguchi method & Grey relational analysis," *Journal of Industrial Engineering International*, vol. 10, no. 4, pp. 149–162, 2014.

[3] X. Li, Z. Lv, Z. Zheng, C. Zhong, I. H. Hijazi, and S. Cheng, "Assessment of lively street network based on Geographic information system and Space Syntax," *Multimedia Tools and Applications*, vol. 76, no. 17, pp. 17801–17819, 2017.

[4] S.-x. Pan, W.-j. Sun, and Z. Zheng, "Video segmentation algorithm based on superpixel link weight model,"

Multimedia Tools and Applications, vol. 76, no. 19, pp. 19741–19760, 2017.

[5] M. Chahal, S. Harit, K. K. Mishra, A. K. Sangaiah, and Z. Zheng, "A survey on software-defined networking in vehicular ad hoc networks: challenges, applications and use cases," *Sustainable Cities and Society*, vol. 35, no. 11, pp. 830–840, 2017.

[6] S. Shivani, S. Tiwari, K. K. Mishra, Z. Zheng, and A. K. Sangaiah, "Providing security and privacy to huge and vulnerable songs repository using visual," *Multimedia Tools and Applications*, vol. 77, 2018.

[7] H. Malekpoor, K. Chalvatzis, N. Mishra et al., "Integrated grey relational analysis and multi objective grey linear programming for sustainable electricity generation planning," *Annals of Operations Research*, vol. 15, no. 8, pp. 1–29, 2017.

[8] N. D. Ghetiya, K. M. Patel, and A. J. Kavar, "Multi-objective optimization of FSW process parameters of aluminium alloy using taguchi-based grey relational analysis," *Transactions of the Indian Institute of Metals*, vol. 69, no. 4, pp. 917–923, 2016.

[9] S. H. Mahdaviani, M. Parvari, and D. Soudbar, "Simultaneous multi-objective optimization of a new promoted ethylene dimerization catalyst using grey relational analysis and entropy measurement," *Korean Journal Of Chemical Engineering*, vol. 33, no. 2, pp. 1–15, 2015.

[10] R. Porwal, V. Yadava, and J. Ramkumar, "Multi-Objective optimization of hole drilling electrical discharge micromachining process using grey relational analysis coupled with principal component analysis," *Journal Of the Institution Of Engineers*, vol. 94, no. 4, pp. 317–325, 2013.

[11] A. E. Hassanien, R. M. Rizk-Allah, and M. Elhoseny, "A hybrid crow search algorithm based on rough searching scheme for solving engineering optimization problems," *Journal of Ambient Intelligence And Humanized Computing*, vol. 12, 2018.

[12] S. Thakur, A. K. Singh, S. P. Ghrera, and M. Elhoseny, "Multilayer security of medical data through watermarking and chaotic encryption for tele-health applications," *Multimedia Tools And Applications*, vol. 78, 2019.

[13] H. Abbas, S. Shaheen, M. Elhoseny, A. K. Singh, and M. Alkhambashi, "Systems thinking for developing sustainable complex smart cities based on self-regulated agent systems and fog computing," *Sustainable Computing: Informatics and Systems*, vol. 19, 2018.

[14] M. Elhoseny, D. Oliva, V. Osuna-Enciso, and M. Aboul Ella Hassanien, "Parameter identification of two dimensional digital filters using electro-magnetism optimization," *Multimedia Tools and Applications*, vol. 79, 2020.

[15] P. Sethuramalingam and B. K. Vinayagam, "Multi objective optimization of multi wall carbon nanotube based nanogrinding wheel using grey relational and regression analysis," *Journal of the Institution of Engineers*, vol. 97, no. 3, pp. 407–416, 2016.

[16] E. Kuram and B. Ozcelik, "Multi-objective optimization using Taguchi based grey relational analysis for micro-milling of Al 7075 material with ball nose end mill," *Measurement*, vol. 46, no. 6, pp. 1849–1864, 2013.

[17] R. Vinayagamoorthy and M. A. Xavior, "Parametric optimization on multi-objective precision turning using grey relational analysis," *Procedia Engineering*, vol. 97, no. 7, pp. 299–307, 2014.

[18] Q. Tu, X. H. Li, and Y. C. LuoChen, "Multi-criteria evaluation of small-scale sprinkler irrigation systems using grey relational analysis," *Water Resources Management*, vol. 28, no. 13, pp. 4665–4684, 2014.

[19] N. S. Dwomoh, R. Sehgal, and V. S. Sharma, "Multi-Objective optimization using grey relational taguchi analysis in machining," *International Journal of Organizational and Collective Intelligence*, vol. 6, no. 4, pp. 45–64, 2016.

[20] S. Liu, "Multi-objective optimization design method for the machine tool's structural parts based on computer-aided engineering," *International Journal of Advanced Manufacturing Technology*, vol. 78, no. 5–8, pp. 1053–1065, 2015.

[21] S. S. Sonali, A. K. Singh, S. P. Ghrera, and M. Elhoseny, "An approach for de-noising and contrast enhancement of retinal fundus image using CLAHE," *Optics & Laser Technology*, vol. 110, 2019.

[22] A. Akbar, M. Ibrar, M. A. Jan, A. K. Bashir, and L. Wang, "SDN-enabled adaptive and reliable communication in IoT-fog environment using machine learning and multi-objective optimizatioSDN-enabled adaptive and reliable communication in IoT-fog environment using machine learning and multi-objective optimization," *IEEE Internet of Things Journal*, vol. 8, 2021.

Hindawi
Scientific Programming
Volume 2021, Article ID 5520662, 11 pages
https://doi.org/10.1155/2021/5520662

Research Article

Construction and Optimization of Higher Education Management System Based on Internet Video Online Technology

Jinhua Liu, Caiping Wang, and Yanhua Wu

Jiangsu Vocational College of Agriculture and Forestry, Jurong 212400, China

Correspondence should be addressed to Jinhua Liu; liujinhua@jsafc.edu.cn

Received 26 February 2021; Revised 25 March 2021; Accepted 1 April 2021; Published 16 April 2021

Academic Editor: Shah Nazir

With the rapid development of information technology, the integration of multimedia technology and traditional institutional systems is increasing. However, at present, the establishment and construction of the higher education system in China are not proportional, the function is vague, and the communication efficiency is not high. The reason for this phenomenon is that the traditional higher education management system has been unable to meet the needs of information society and information integration and it is unable to meet the increasing demand for higher education. Therefore, based on the above reasons, it is very important and has long-term strategic significance to establish an efficient and convenient higher education management system based on today's Internet multimedia technology in a timely and reasonable manner. This paper analyzes the current status of the higher education management system and the problems that are arisen and proposes a three-dimensional integrated management system based on the innovation of Internet multimedia technology. At the same time, aiming at the urgent needs of the current lifelong education of higher education, a reasonable system of higher education diversion is established to focus on solving the phenomenon of imbalanced proportions, ambiguous functions, and low communication efficiency. In the final practical part of this paper, the system model of the three-dimensional integrated education system and the corresponding practical solution to the education diversion are given. After discussion, it has practical application value. The results of the research have shown the effectiveness of the proposed study.

1. Introduction

With the improvement of China's economy and the implementation of the education first policy, the popularization of higher education has made great achievements in China. So far, the scale of China's higher education has reached the world's first level [1]. However, with the increasing popularity and development of higher education, the higher education system exposed in the process is too aging [2], and the quality, structure, and corresponding educational benefits of higher education development are increasingly prominent [3]. Furthermore, rationalization of communication efficiency has become an important topic in the construction of higher education systems today [4].

There are a lot of research studies and analyses on the system construction of higher education. The analysis mainly focuses on the design of higher education curriculum, the construction of higher education resources sharing, the research of higher education teaching mode, and the corresponding higher education management and evaluation mechanism [5]. The Higher Education Technical Cooperation Management Committee has systematically proposed the theory and practice of higher education, which systematically introduces the form and practice of the higher education system, but due to its age, its corresponding theory is no longer applicable in today's information society [6]. But its focus is on the teaching of online courses [4], which focuses on network distance teaching, so it has certain limitations, and at the same time, the implementation plan of the education diversion system is not given [7]. Foreign countries' construction of higher education systems is limited to their own educational traditions. Similar to Ralph's "Sponsored Mobility and Competitive Mobility and School Education System" proposal, this analysis believes that the school's teaching system has two modes, namely, sponsored mobility and competitive mobility [8], and it is believed that sponsorship mobility is dominant in the UK education system,

so its theory has no practical significance for the construction of China's education system [9]. It has also proposed the construction of an education system based on the current situation in the United States, which emphasizes the factors affecting higher education such as class mobility, but it has no practical significance in inspiring China's higher education or modern higher education [10].

Based on the current teaching situation, this paper proposes a three-dimensional integrated teaching system, proposes a three-in-one higher education system in combination with today's computer network multimedia technology, and introduces a quality management system in the higher education system [11] to realize the scientific management and standardized handling of the higher education system. Finally, the paper puts forward the actual model of the higher education management system in the practice part and conducts in-depth analysis and discussion [12].

This article is organized as follows. In Section 2, the key research works in the area are presented, which mainly analyzes the current problems in higher education and the relationship between computer technology and the current education system [13]. Section 3 mainly analyzes the principle of the three-dimensional integrated quality management system proposed in this paper and the corresponding application of computer network multimedia technology. Section 4 mainly constructs and analyzes the actual model of the higher education system based on multimedia technology proposed in this paper. Section 5 summarizes the content of this article [14].

2. Related Work

Research specifically analyzes the problems existing in the current higher education management system in China and analyzes them in detail [15]. At the same time, this section will specifically analyze the multimedia technology tools currently used in education to provide technical support and theoretical support for the follow-up higher education management mechanism. Figure 1 shows the framework of the traditional higher education management system. Corresponding to Figure 1, we can see that the traditional higher education management system is too mechanized, too complicated, and lengthy [16].

2.1. Analysis of Problems in Traditional Higher Education System. The comprehensive survey and analysis can summarize the problems existing in the current higher education management system as follows [17].

2.1.1. Higher Education Teaching Infrastructure Management. In this respect, there is a situation of uneven distribution of resources, which leads to the increasing distribution of educational resources in developed regions and the growing gap between infrastructure resources in underdeveloped regions and developed regions, especially in online higher education. The most obvious form is the low computer penetration rate in underdeveloped areas and at the same time limited by the speed of network transmission. The quality and stability of distance education teaching information such as broadband multimedia telecommunication networks are very poor, which leads to higher education [18]. The serious mismatch of teaching resources and the interaction and function of interactive technical media cannot play a role. At the same time, in terms of infrastructure construction management, better and better infrastructure funding has flowed to the place where the original teaching infrastructure is good, resulting in a large amount of waste of resources, and higher education institutions in underdeveloped regions cannot further improve their teaching condition.

2.1.2. Higher Education Teaching Resource Management. In the direction of teaching resource management that is too old, the relevant textbooks of higher education and the corresponding curriculum practice still stay in the existing teaching system, the teaching resources are equipped, and the use of reasonable information is not enough. This has caused a huge waste of human resources and time costs.

2.1.3. Higher Education Teacher Team Management System. There are serious problems in the distribution of teachers' resources. The teacher resources in different regions are seriously dysfunctional. The teaching ability of many teachers cannot match the teaching content and student needs. The fundamental reason is that there is a lack of reasonableness. And the teacher training management keeps pace with the times. It is impossible to be reasonable and efficient in the management of the title of teachers.

2.1.4. The Lack of Higher Education Diversion System. There are too many people in China who urgently need higher education, but because the level of people and needs are different, the traditional higher education management system is no longer applicable. This led to the diversion system in the higher education management system, but the domestic research is too little in theoretical construction. At the same time, the research content is single and not comprehensive. When the specific research is carried out, there are many descriptions, but the mechanism analysis is too small. The method is also too qualitative, and there are fewer persuasive examples. The specific corresponding framework for the problems in the higher education diversion system is shown in Figure 2.

Based on the abovementioned problems in the higher education management system, the performance of different categories and corresponding details can be summarized in the reason tree shown in Figure 3.

2.2. Analysis of the Relationship between Internet Multimedia Technology and Higher Education. Internet-based multimedia technology has gradually integrated into the higher education system. Before the relevant research, this paper made a questionnaire on the acceptance of relevant information about students and teachers in the integration of

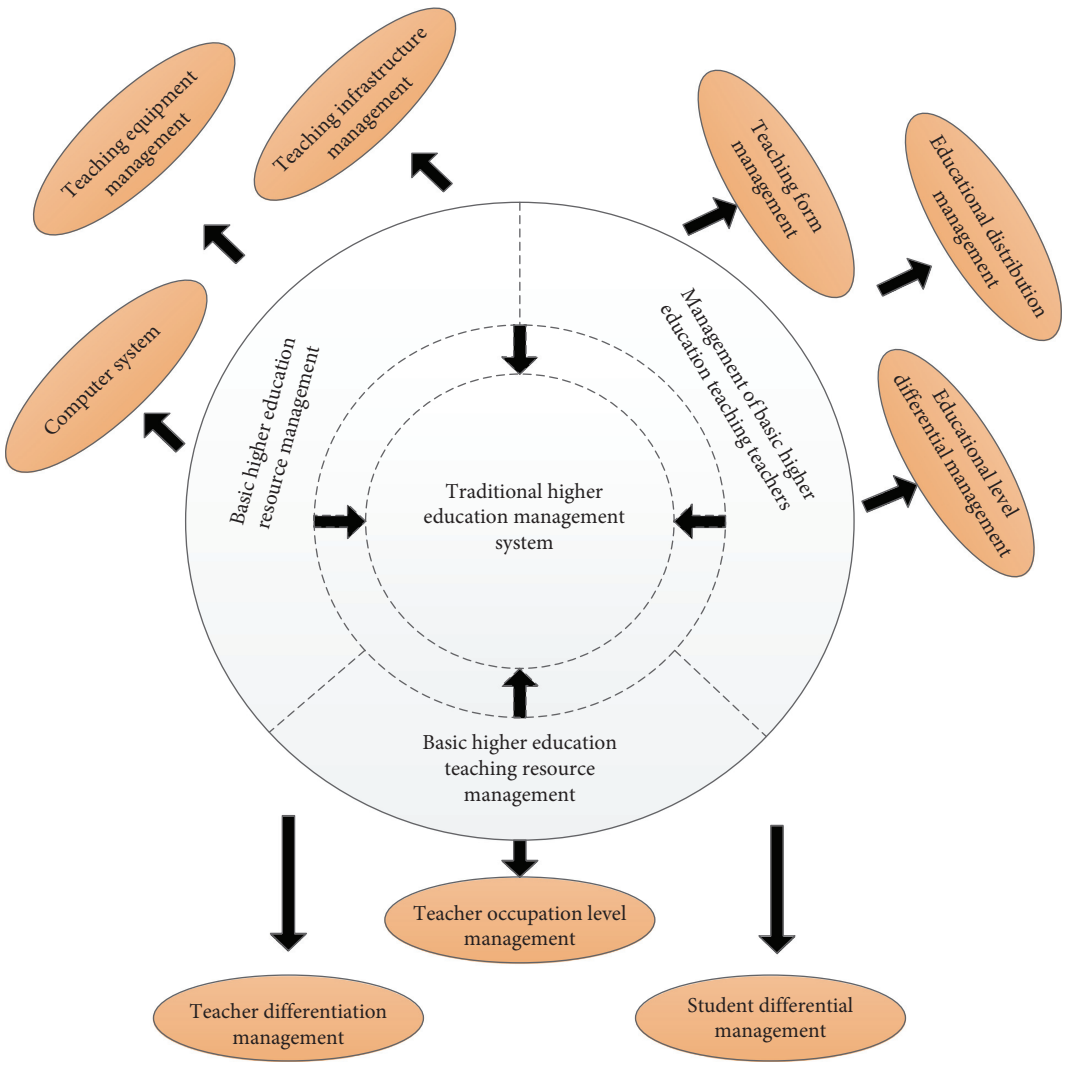


Figure 1: The framework of the traditional higher education management system.

information technology into the higher education system. The corresponding survey sample is a total of 2,000 colleges and universities in the east and the west, which fully guarantees the regionality of the sample and the sample size. The corresponding data table is shown in Table 1.

2.2.1. Internet Multimedia Education Technology. Internet multimedia teaching technology is limited to sound, text, and images, as well as the defects of recording through books. Multimedia education technology has greatly improved the informationization degree of education. Internet multimedia education technology is a convenient and rapid integration of a variety of media, and its corresponding educational and teaching content is presented in a more abundant form. Its main advantages are mainly as follows:

(1) Internet-based education and teaching technology can synthesize and stimulate a variety of media, which leads to management changes in textbooks in the higher education management system. For example, traditional textbook materials can become auxiliary textbooks under informatization, such as multimedia textbooks, teaching materials, and some hypertext technology of sound, image, and animation synthesis.

(2) Based on the convenience of Internet technology, the real-time opening of education and teaching can be realized. At the same time, the management system of higher education is also conducive to the release and timely delivery of information. In terms of the details of its implementation, it is dedicated to teaching. The information website can be updated and supplemented with teaching resources and new management trial methods at any time.

(3) The emergence of information technology teaching is conducive to the change of students' learning styles. Therefore, it also realizes the management of students' learning methods in the higher education management system to a certain extent. In this process, multimedia computer technology has become an auxiliary tool for students to conduct self-learning management, and it has become a

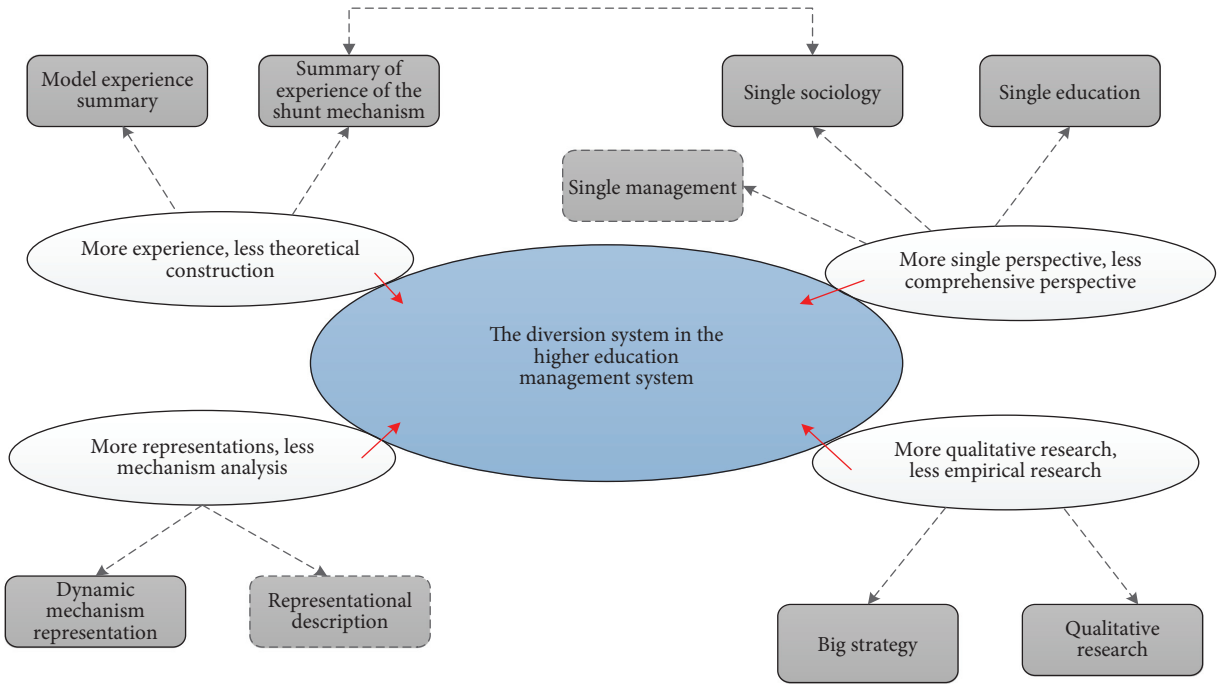


Figure 2: Specific corresponding framework for the problems in the higher education diversion system.

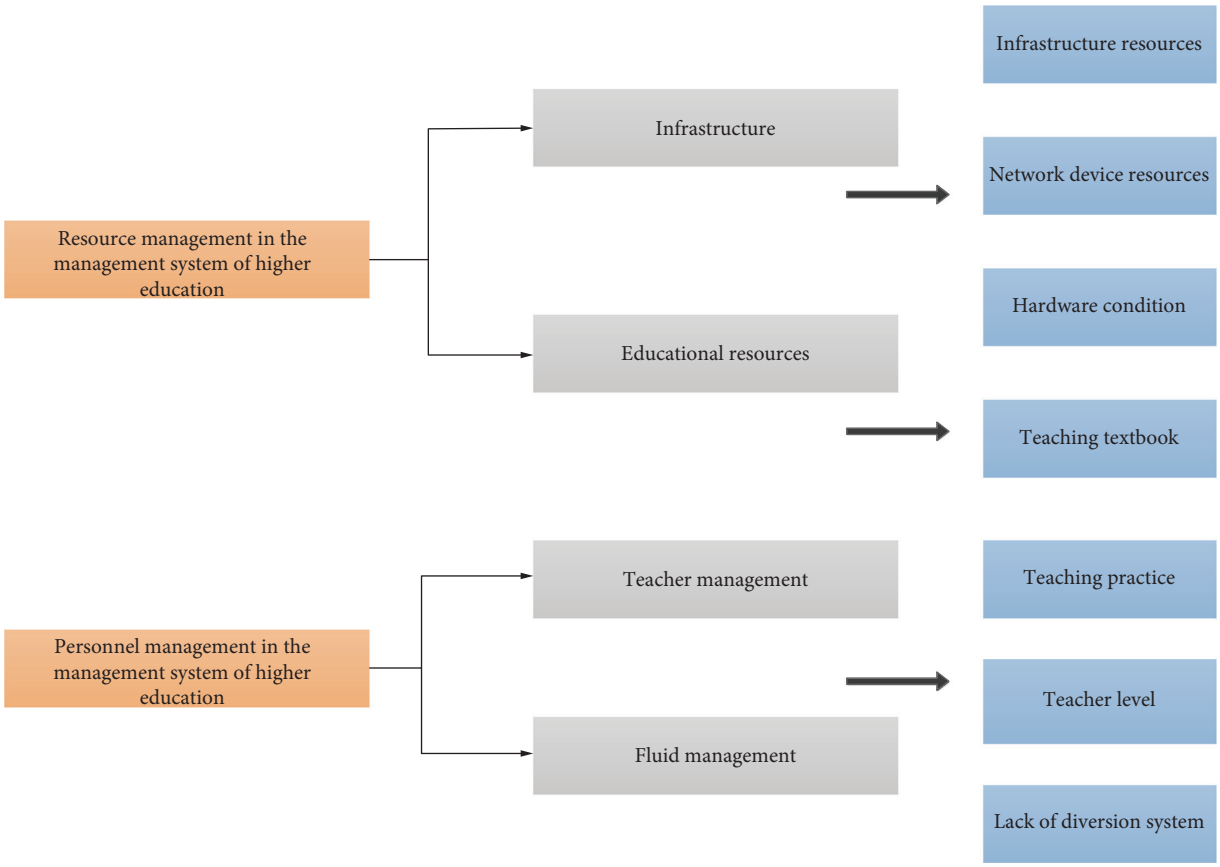


Figure 3: The performance of different categories and corresponding details.

Table 1: The statistics of the textbooks used in traditional English translation.

Total number of samples	Option	Proportion	Conclusion
2000	Fully accepted and expect a good form to appear	80	On the whole, I tend to look forward to the emergence of a new higher education management system
	Accept but stay optimistic	15	
	No feeling is not important	1	
	Do not expect to feel the current problem	4	

student's performance tool, communication tool, and research tool. At the same time, as a part of management in higher education, informatization makes the student management part more efficient.

(4) Informatization technology makes the management of teacher-student interaction module more vivid and timely. Education and teaching are essentially the interaction between teachers and students. The traditional mode of higher education is almost limited to verbal communication and is subject to the time and space of communication. The main applications include hierarchical teaching, cooperative teaching, and feedback teaching.

Based on the above analysis, the corresponding frame diagram is shown in Figure 4.

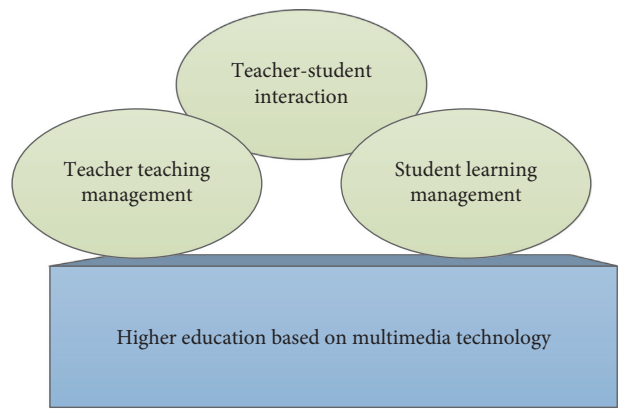


Figure 4: The corresponding frame diagram.

3. Construction of Higher Education Management System Based on Internet Multimedia Technology

3.1. Three-Dimensional Integrated Education Quality Management System Based on Multimedia Technology. This paper proposes a quality management system based on multimedia online technology, combined with the outcome management system evaluation system of teaching quality which effectively guarantees and continuously improves the management of education quality mode. The corresponding frame diagram is shown in Figure 5.

According to the optimal information gain attribute, the data classification and map feature reconstruction are carried out, and the univariate time series of big data time series is constructed as $\{x_n\}$. The length of the data sample is N, and the time series is reconstructed by phase space. It is obtained that the high-dimensional space trajectory vector of big data's characteristic distribution is described as follows:

$$\mathbf{L} = \begin{bmatrix} x_1^T, \\ x_2^T, \\ \vdots \\ x_N^T, \end{bmatrix} = \begin{bmatrix} x_1, & x_{1+\tau}, & \ldots, & x_{1+(m-1)\tau}, \\ x_2, & x_{2+\tau}, & \ldots, & x_{2+(m-1)\tau}, \\ \vdots & \vdots & \ddots, & \vdots \\ x_{N-1}, & x_{N-1+\tau}, & \ldots, & x_{N-1+(m-1)\tau}, \end{bmatrix}, \quad (1)$$

where in the embedding dimension is m and the time delay is τ, and in the large data distribution space, the trajectory matrix $\mathbf{L}$ of high-dimensional space is decomposed with the unitary matrix decomposition method. The $\mathbf{L} = \mathbf{U} * \mathbf{S} * \mathbf{C}$ is obtained by using the unitary matrix decomposition method, where $\mathbf{U}$ and $\mathbf{C}$ are orthogonal matrices, and

$$\mathbf{C} = (c_1, c_2, \ldots, c_n), \quad (2)$$

wherein $\mathbf{S}$ is an average measure matrix of $\mathbf{L}$, and the elements and eigenvalues of matrix $\mathbf{S}$ satisfy

$$\begin{aligned} \mathbf{S} &= \operatorname{diag}(\sigma_1, \sigma_2, \ldots, \sigma_n), \\ &\sigma_1 \geq \sigma_2 \geq, \cdots, \geq \sigma_n \geq 0. \end{aligned} \quad (3)$$

The phase space reconstruction locus matrix $\mathbf{L}$ is obtained by calculating the correlation degree, and the scale matrix $\mathbf{X}$ is given by

$$\mathbf{X} = \begin{bmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \\ \vdots \\ \mathbf{x}_3 \end{bmatrix} = \begin{bmatrix} \mathbf{a}_1^T\mathbf{c}_1 & \mathbf{a}_1^T\mathbf{c}_2 & \ldots & \mathbf{a}_1^T\mathbf{c}_m \\ \mathbf{a}_2^T\mathbf{c}_1 & \mathbf{a}_2^T\mathbf{c}_2 & \ldots & \mathbf{a}_2^T\mathbf{c}_m \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{a}_N^T\mathbf{c}_1 & \mathbf{a}_N^T\mathbf{c}_2 & \ldots & \mathbf{a}_N^T\mathbf{c}_m \end{bmatrix}. \quad (4)$$

The autocorrelation coefficient of big data information flow scalar time series is obtained as follows:

$$\rho_{XY} = \frac{\operatorname{Cov}(X, Y)}{\sqrt{D(X)}\sqrt{D(y)}}. \quad (5)$$

Based on the sparse iterative covariance estimation of big data information flow, the autocorrelation covariance is obtained as follows:

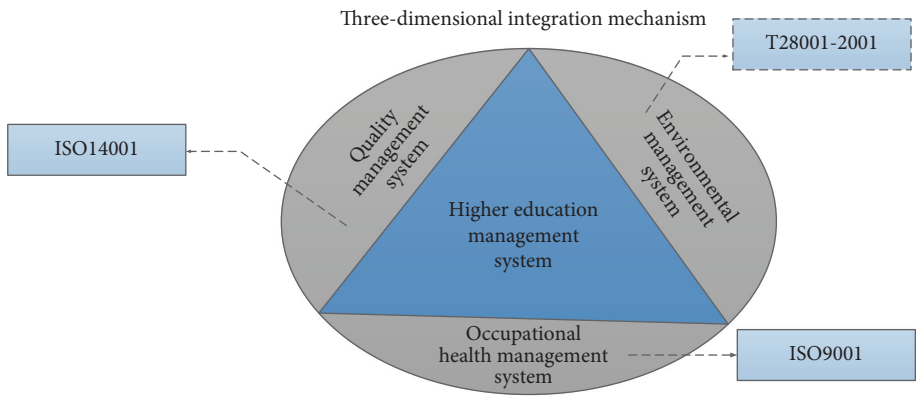


Figure 5: The corresponding frame diagram.

$$\mathrm{Cov}(X, Y) = E\{[X - E(X)][Y - E(Y)]\}. \quad (6)$$

With the above analysis, the nonlinear time series analysis and phase space reconstruction of big data information flow can be realized.

3.1.1. Analysis of the Quality Management System of Three-Dimensional Integrated Education. The three-dimensional integrated higher education management system proposed in this paper mainly contains five parts of understanding. This paper summarizes it as a series of numbers as "12345," in which different numbers correspond to "54321," respectively, and The corresponding explanation is based on the five concepts contained in the three-dimensional integrated concept system; the second is that the corresponding institutional system contains four levels of documents; the third monitoring system forms a three-way monitoring; the corresponding fourth is the technical level, which includes two platforms; the fifth corresponds to a continuous improvement system, which ultimately forms a mechanism. The corresponding concepts are shown in Table 2.

In the three-dimensional integrated education management system, the system can be compared to the automobile, in which the concept of the system is equal to the steering wheel, which plays a guiding role and has a very important position; the file system at the institutional level is similar to the traffic map and its planning. The route of the car also reflects the connection and limitation of the three-dimensional concept in various rules; the monitoring system is like the instrument panel, which reflects the running state of the entire education system in real time; the improved system is similar to the maintenance station and is adjusted in time. Technical support level is similar to the highway, which can ensure the timely, accurate, and efficient information processing of the entire higher education three-dimensional integrated management system. The corresponding logical framework diagram is shown in Figure 6.

The core concept of the three-dimensional integration proposed in this paper is to focus on the students, highlighting the importance of people, taking into account the prevention-oriented, pursuing a scientific management system, taking the facts as the basis for decision-making, and attaching importance to the research role. For the management foundation, it emphasizes the management of higher education system according to law; among them, the management of higher education management system should focus on teaching management, and all mechanisms should serve teaching.

In the three-dimensional integrated management system, three points are observed. One is to cherish and protect the health and safety of teachers and students, and the other is to respect the harmony between nature and society. The third is to provide quality educational products. In the higher education management system based on the three-dimensional integration, we must keep the school in harmony with nature, maintain the harmony between teachers and students, maintain the harmony of teaching, research, and management, and achieve a balanced and harmonious relationship between higher campus management and teaching management. In the quality management system of higher education, we must always be aware of the risk factors existing in the management system and conduct preventive assessments to prevent them from happening.

3.1.2. Integration of Multimedia Technologies Based on Higher Education. The three-dimensional integrated education management system proposed in this paper is based on Internet multimedia technology. In addition, the three-dimensional integrated higher education network management mechanism relies heavily on multimedia Internet technology. Based on the theory of this paper, the ecology of the Internet higher education management system can be constructed. The ecology is mainly based on the teaching content of the online course, the quality of education, the education group, and the form of education. In the course of network higher education teaching and the corresponding management system, such as 3D virtual environment technology and multimedia imaging technology, the corresponding application architecture technology is shown in Table 3.

3.2. Diversion Strategy in Higher Education System. The abovementioned problem of the increasing popularity of higher education is that people's demand for higher education is increasingly diversified. view of the dilemmas of my

Table 2: The corresponding concepts.

Three-dimensional integrated system	Main content
Concept level	5: 5 concepts, corresponding to "education as a service," "big teaching quality concept," teaching quality "three-dimensional view," "education quality risk concept," "full quality control concept"
Institutional level	4: management manual, document program, job guidance, record
Monitoring level	3: school, student, society
Technical level	2: management information system and educational administration network system
Improvement level	1: continuous improvement mechanism

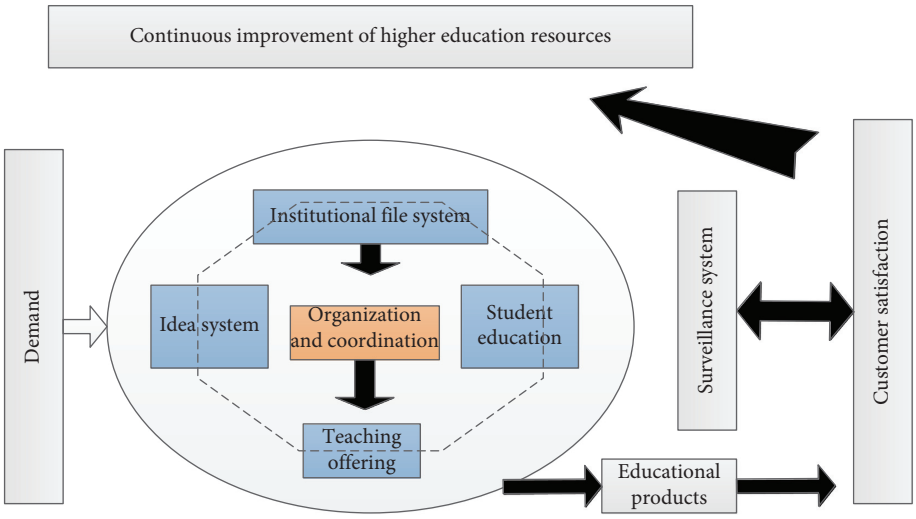


Figure 6: The corresponding logical framework diagram.

Table 3: Student and translation environment interaction questionnaire.

	Main medium	Multimedia technology application
Network higher education system	Before class	Smart device
	In the course	3D virtual technology, image multimedia technology
	After class	Intelligent device, artificial intelligence technology

country's higher education system, such as low system authority, poor system coordination, and imperfect system operation, this paper proposes a three-level education management mechanism, which classifies and refines the different needs of higher education. The corresponding framework is shown in Figure 7.

In the specific implementation strategy, it is necessary to promote the combination with market induction, promote the combination of top-level design and partial pilot, and combine the moderate breakthrough with steady advancement, and also based on the combination of national conditions and reasonable reference. In the specific implementation details, special laws and regulations should be formulated to promote the authority of the multidirectional diversion system of higher education. It is necessary to improve the horizontal and vertical coordination mechanism, improve the coordination of the multidirectional diversion system of higher education, and actively promote the transformation work and promote higher education. In the implementation of channels, we must actively build multiple communication channels, establish a two-way communication mechanism, and actively improve existing communication mechanisms. In terms of connection, we must actively establish higher vocational education undergraduate level education, develop higher vocational education graduate level, pay attention to top-level design, improve the systemic system, optimize the formulation process, ensure the scientific nature of the system, and enhance the operational efficiency of the system. In terms of legal guarantees, we must actively improve the promotion of legal and legislative work, establish a standard framework covering the whole country, and establish a supporting system for mutual assistance and harmony.

4. Practice and Analysis

Based on the abovementioned principle description of the higher education management system based on computer Internet technology and the corresponding analysis and research on its core ideas, this section will carry out the construction and verification analysis of the actual model for the above theory.

Figure 8 shows the corresponding platform block diagram of the higher education management system proposed in this paper.

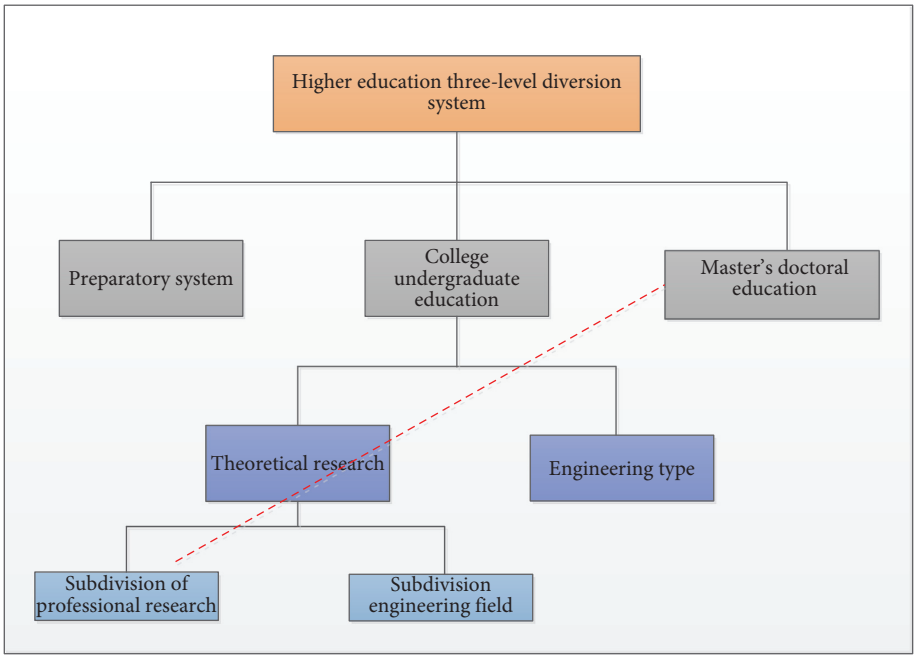


Figure 7: The corresponding framework.

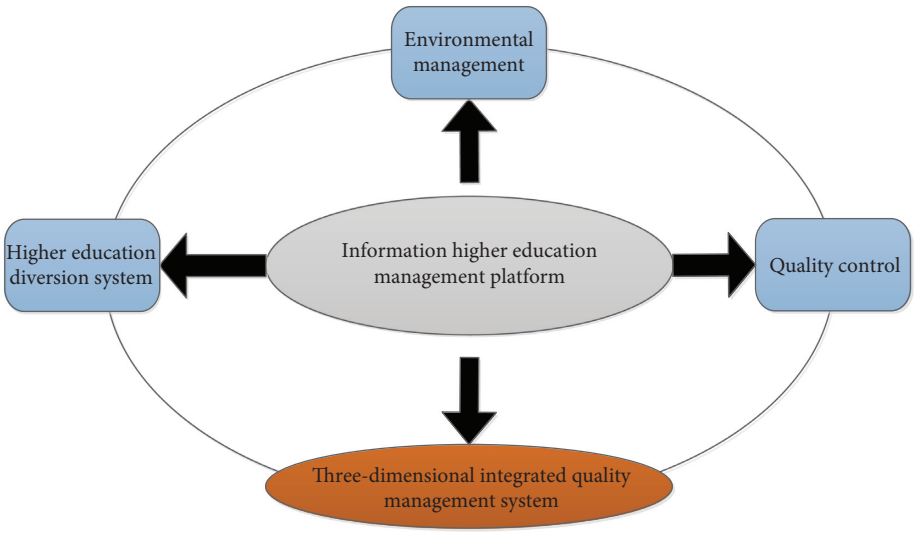


Figure 8: The corresponding platform block diagram of the higher education management system.

The corresponding operation mechanism is shown in Figure 9. In the actual operation process, the core idea of three-dimensional integration is always adhered to, and the operation system is monitored throughout the whole process. In practical application, the detailed process of higher education operation in the management system should be analyzed. Through process control to ensure that the quality is guaranteed, while correspondingly affecting the quality of education and the smooth operation of the system, in the aspect of the education subfluid system, it is necessary to pay attention to the policy of top-level design in a timely manner. The corresponding five management mechanisms are the teaching work management mechanism, the student work management mechanism, the scientific research work management mechanism, the school work management mechanism, and the logistics support management. The mechanism will be coordinated and fully utilized to give full play to the advantages and potential of the three-level diversion system.

Based on the analysis and establishment of the above theoretical model, this paper conducts a practical questionnaire analysis on the higher education management mechanism based on multimedia technology proposed in this paper. The corresponding data table is shown in Table 4. It can be seen from the table that it is certain. To a certain extent, it reflects the acceptance of the management mechanism proposed in this paper. The corresponding sample of the questionnaire is 4000. The corresponding

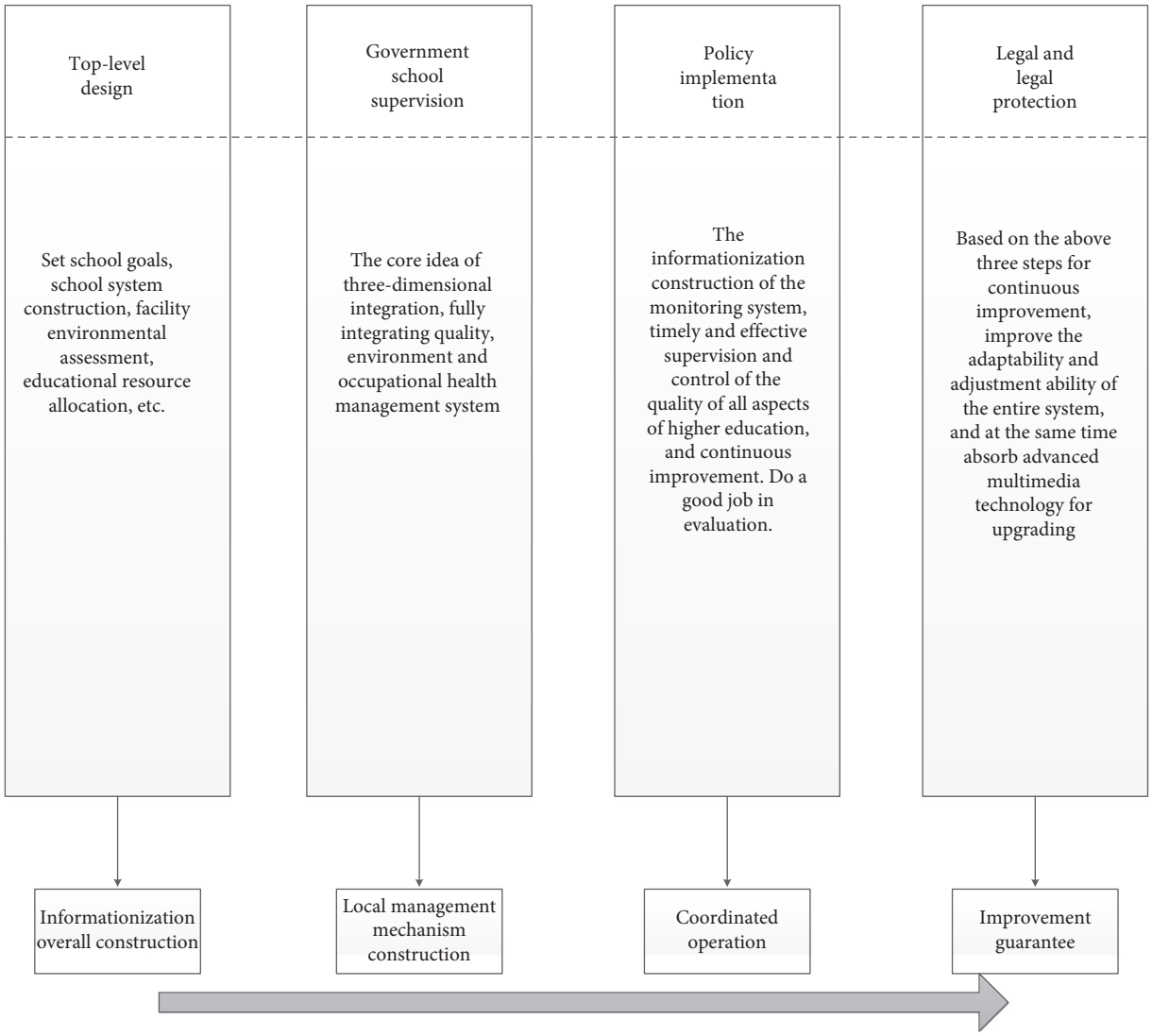


Figure 9: The corresponding operation mechanism.

Table 4: The corresponding data table.

Total number of samples	Option	Proportion (%)	Conclusion
4000	Fully accepted and expect a good form to appear	78	On the whole, I tend to look forward to the emergence of a new higher education management system
	Accept but stay optimistic	15	
	No feeling is not important	3	
	Do not expect to feel the current problem	4	

population includes teachers, students, and related educators, covering the eastern and western regions in the coverage area.

The corresponding actual number of people is shown in Figure 10. From the bar chart, the higher education management system based on Internet multimedia

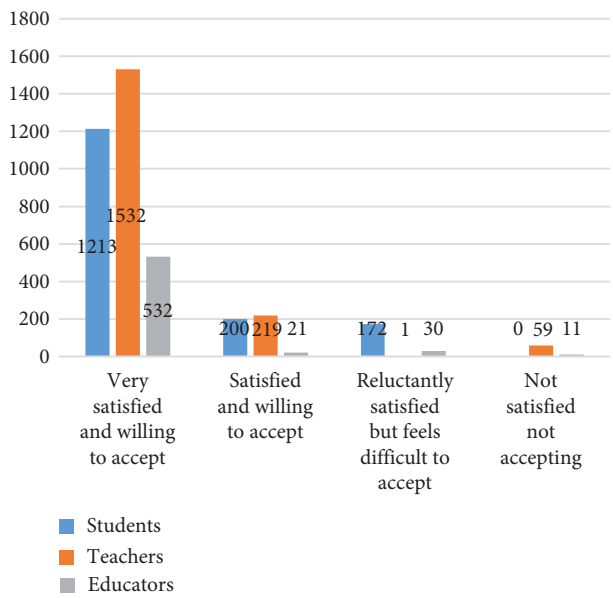


Figure 10: The corresponding histogram.

technology proposed in this paper has been recognized and paid attention.

5. Conclusion

With the improvement of China's economy and the implementation of the education first policy, the spread of higher education has made excessive successes in China. This paper proposes an online education management system based on Internet multimedia technology, which realizes the network management of education through computer technology. At the same time, aiming at the urgent needs of the current lifelong education of higher education, a reasonable system of higher education diversion is established to focus on solving the phenomenon of imbalanced proportions, ambiguous functions, and low communication efficiency. In the final practical part of this paper, the system model of the three-dimensional integrated education system and the corresponding practical solution to the education diversion are given. After discussion, it has practical application value. The results of the research have shown the effectiveness of the proposed study.

Data Availability

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

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This study was supported by the general project of philosophy and social science research of Jiangsu University in 2019 "on the strategy of Zhenjiang higher vocational education industry education integration and school-enterprise cooperation innovation platform under the background of vocational education reform (no. 2019sja1951)" and Science and Technology Project of Jiangsu Vocational College of Agriculture and Forestry in 2019 "Research on the Development Countermeasures of Zhenjiang Higher Vocational Education Cooperative Innovation and School Enterprise Cooperative Education (no. 2019kj023)".

References

[1] S. B. Dias and J. A. Diniz, "Towards an enhanced learning management system for blended learning in higher education incorporating distinct learners' profiles," *Journal of Educational Technology & Society*, vol. 17, no. 1, pp. 307–319, 2014.

[2] A. Decramer, C. Smolders, and A. Vanderstraeten, "Employee performance management culture and system features in higher education: relationship with employee performance management satisfaction," *The International Journal of Human Resource Management*, vol. 24, no. 2, pp. 352–371, 2013.

[3] J. Dalling and P. Rafferty, "Open source, open minds?: An investigation into attitudes towards open source library management systems in UK higher education libraries," *Program-electronic Library and Information Systems*, vol. 47, no. 4, pp. 399–423, 2013.

[4] J. B. Abugre, "Institutional governance and management systems in sub-Saharan Africa higher education: developments and challenges in a Ghanaian research university," *Higher Education*, vol. 75, no. 2, pp. 323–339, 2017.

[5] K. A. Al-Busaidi and H. Al-Shihi, "Key factors to instructors' satisfaction of learning management systems in blended learning," *Journal of Computing in Higher Education*, vol. 24, no. 1, pp. 18–39, 2012.

[6] A. Clarke, "The campus environmental management system cycle in practice: 15 years of environmental management, education and research at Dalhousie University," *International Journal of Sustainability in Higher Education*, vol. 7, no. 4, pp. 374–389, 2006.

[7] W. Y. W. Lo, "The recalibration of neoliberalisation: repoliticising higher education policy in Hong Kong," *Higher Education*, vol. 73, no. 5, pp. 759–773, 2017.

[8] A. Smolentseva, "Universal higher education and positional advantage: soviet legacies and neoliberal transformations in Russia," *Higher Education*, vol. 73, no. 2, pp. 209–226, 2017.

[9] A. Bagerelsborg and L. Greve, "Establishing a method for analysing metaphors in higher education teaching: a case from business management teaching," *Higher Education Research & Development*, vol. 38, no. 7, pp. 1329–1342, 2017.

[10] P. A. M. Maassen and H. P. Potman, "Strategic decision making in higher education: an analysis of the new planning system in Dutch higher education," *Higher Education*, vol. 20, no. 4, pp. 393–410, 2017.

[11] J. Mampaey, J. Brankovic, and J. Huisman, "Inter-institutional differences in defensive stakeholder management in higher education: the case of Serbia," *Studies in Higher Education*, vol. 44, no. 6, pp. 978–989, 2017.

[12] M. Seyfried and M. Ansmann, "Unfreezing higher education institutions? Understanding the introduction of quality management in teaching and learning in Germany," *Higher Education*, vol. 75, no. 6, pp. 1061–1076, 2018.

[13] C. C. Hsieh and J. Huisman, "Higher education policy change in the European higher education area: divergence of quality

assurance systems in England and the Netherlands," *Research Papers in Education*, vol. 32, no. 1, pp. 71–83, 2017.

[14] M. Ekman, M. Lindgren, and J. Packendorff, "Universities need leadership, academics need management: discursive tensions and voids in the deregulation of Swedish higher education legislation," *Higher Education*, vol. 75, no. 2, pp. 299–321, 2017.

[15] B. Broucker, K. D. Wit, and J. C. Verhoeven, "Higher education for public value: taking the debate beyond new public management," *Higher Education Research & Development*, vol. 37, no. 2, pp. 227–240, 2017.

[16] A. Parvin, "Leadership and management in quality assurance: insights from the context of Khulna university, Bangladesh," *Higher Education*, vol. 77, no. 2, pp. 739–756, 2018.

[17] N. Dreamson, G. Thomas, A. L. Hong, and S. Kim, "Policies on and practices of cultural inclusivity in learning management systems: perspectives of Indigenous holistic pedagogies," *Higher Education Research & Development*, vol. 36, no. 5, pp. 947–961, 2016.

[18] F. Marimon, M. Masmachuca, and J. Berbegalmirabent, "Fulfilment of expectations on students' perceived quality in the Catalan higher education system," *Total Quality Management & Business Excellence*, vol. 31, no. 5, pp. 483–507, 2018.