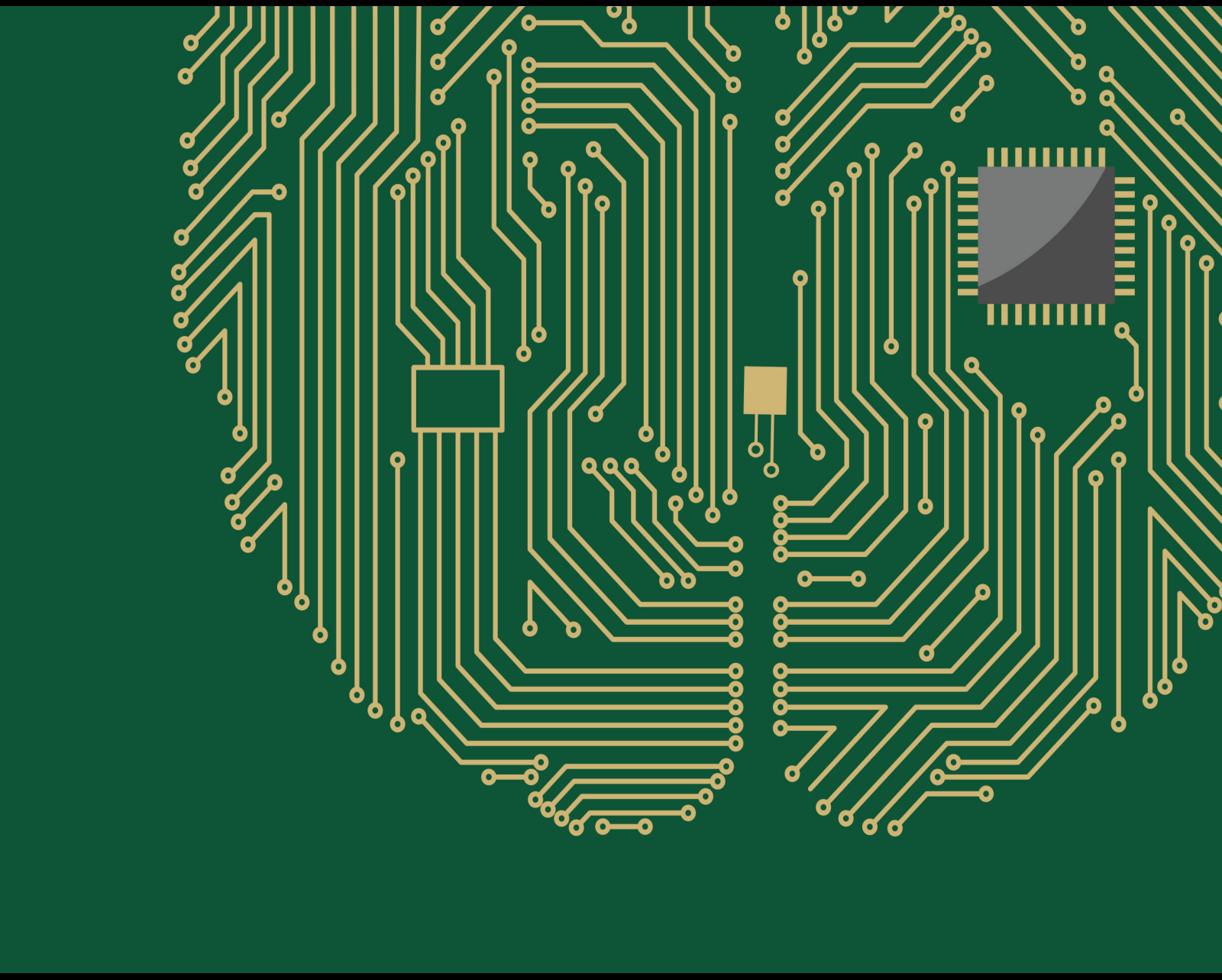
Computational Intelligence and Neuroscience

Advances in Machine Learning for Computational Neural Modeling for Visual Recognition

Lead Guest Editor: Ning Cao
Guest Editors: Luobing Dong and Russell Higgs



Advances in Machine Learning for Computational Neural Modeling for Visual Recognition

Computational Intelligence and Neuroscience

Advances in Machine Learning for Computational Neural Modeling for Visual Recognition

Lead Guest Editor: Ning Cao
Guest Editors: Luobing Dong and Russell Higgs

Contents

Contents

Contents

Hindawi
Computational Intelligence and Neuroscience
Volume 2023, Article ID 9843832, 1 page
https://doi.org/10.1155/2023/9843832

Retraction

Retracted: Application of CNN-Based Machine Learning in the Study of Motor Fault Diagnosis

Computational Intelligence and Neuroscience

Received 17 October 2023; Accepted 17 October 2023; Published 18 October 2023

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1].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

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

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

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

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

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

References

[1] X. Peng, L. Wei, and W. Gao, "Application of CNN-Based Machine Learning in the Study of Motor Fault Diagnosis," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 9635251, 9 pages, 2022.

Hindawi
Computational Intelligence and Neuroscience
Volume 2022, Article ID 2047576, 10 pages
https://doi.org/10.1155/2022/2047576

Research Article

Classification System of National Music Rhythm Spectrogram Based on Biological Neural Network

Dan Mi [1] **and Lu Qin** [2]

[1] *Department of Music, Xinxiang University, Xinxiang, Henan 453003, China*
[2] *Department of Sports, Xinxiang University, Xinxiang, Henan 453003, China*

Correspondence should be addressed to Lu Qin; qinlu0727@xxu.edu.cn

Received 1 August 2022; Revised 15 September 2022; Accepted 28 September 2022; Published 12 October 2022

Academic Editor: Ning Cao

National music is a treasure of Chinese traditional culture. It contains the cultural characteristics of various regions and reflects the core value of Chinese traditional culture. Classification technology classifies a large number of unorganized drama documents, which are not labeled, and to some extent, it helps folk music better enter the lives of ordinary people. Simulate folk music of different spectrum and record corresponding music audio under laboratory conditions Through Fourier transform and other methods, music audio is converted into spectrogram, and a total of 2608 two-dimensional spectrogram images are obtained as datasets. The sonogram dataset is imported into the deep convolution neural network GoogLeNet for music type recognition, and the test accuracy is 99.6%. In addition, the parallel GoogLeNet technology based on inverse autoregressive flow is used. The unique improvement is that acoustic features can be quickly converted into corresponding speech time-domain waveforms, reaching the real-time level, improving the efficiency of model training and loading, and outputting speech with higher naturalness. In order to further prove the reliability of the experimental results, the spectrogram datasets are imported into Resnet18 and Shufflenet for training, and the test accuracy of 99.2% is obtained. The results show that this method can effectively classify and recognize music. The experimental results show that this scheme can achieve more accurate classification. The research realizes the recognition of national music through deep learning spectrogram classification for the first time, which is an intelligent and fast new method of classification and recognition.

1. Introduction

Music classification is an important part of multimedia applications. With the rapid development of data storage, compression technology, and Internet technology, music type data have increased dramatically. Traditional manual retrieval methods can no longer meet the needs of massive music information retrieval and classification. The computer automatic classification of music is to solve the above problems. The computer automatic recognition of music is a new interdisciplinary subject. Its research involves physics, signal processing, human-computer interaction, music theory, music psychology, and many other disciplines. Its main task is to obtain the relevant information of music content through the processing and feature extraction of audio signals, which can be used for comparison, classification, and even automatic recording [1]. Therefore, it is urgent to adopt new technologies to effectively manage massive music resources. However, this method requires manual annotation of massive databases, and the workload is very huge. Another method is to manage according to the content of music, that is, index and retrieve from the music database according to the rhythm, sound quality, and melody characteristics of music [2].

Music classification is essentially a problem of pattern recognition, which mainly includes two aspects: feature extraction and classification. Many researchers have done a lot of work in this field. Using univariate autoregressive integrated moving average (ARIMA) model and nonlinear autoregressive exogenous (NARX) model to classify music, the average absolute error and mean square error of the database are 5.5% and 10.6%, respectively. The average absolute error and mean square error of 10 min database are 2.3% and 12.8%, respectively. The support vector machine

(SVM) is adopted to music classification and has achieved high prediction accuracy and convergence speed. Based on the modeling technology of chaotic time series and the wind music classification strategy of Apriori algorithm, the chaotic characteristics of music rhythm time series are analyzed, and its significance to music classification is discussed [3]. Literature [4] classifies music using a double-layer emotional classification model based on AdaBoost. The paper points out that the first layer of the system classifies according to the intensity and time-domain characteristics of music signals, and the second layer classifies according to the rhythmic characteristics of music signals. Finally, according to the classification results of the two-level classifiers and the weights of each level of classifiers, the final classification results are obtained by matching and accumulating. Literature [5] attributes the spectrum classification of music to the problem of regression. According to quality, melody, and rhythm characteristics of music, SVR regression algorithm is used to calculate the arousal and value values of each song and locate them in the emotion plane proposed by Thayer. Search the music songs you need according to the explicit values of arousal and valance. Literature [6] proposed using BP neural network algorithm to classify MIDI music files. Literature [7] uses GMM machine learning algorithm to detect and track music emotion. Siphocly and others [8] proposed a double-layer SVM model for music emotion classification. It can be seen that the classification of music emotion has been fully concerned by the academic community [9].

Since its birth, deep learning has been favored by many researchers. It has been widely used in many fields, such as food detection [10], medical detection [11], face recognition [12], emotion recognition [13], crop recognition [14], image enhancement [15], and machine control [16], and researchers have also begun to apply it to the field of music classification and recognition. Literature [16] uses neural network to predict the uncertain input data expressed in the form of interval. The purpose is to quantify the uncertainty generated by the input data and prediction model in the prediction and finally train a high-performance multilayer perceptron neural network, so as to realize the classification of music. Literature [17] uses deep learning self-encoder for short-term music prediction. The proposed CNN model is superior to the classical DNN model and the DNN model with shallow structure. Literature [18] predicts music based on time series prediction and extreme value optimization. Based on empirical wavelet transform, long-term and short-term memory neural network, Elman neural network and other deep learning methods for music type recognition, the results show that the model has good performance in high-precision music recognition [20, 21]. Although more and more researchers are aware of the importance of deep learning technology in the field of music type recognition, they have not found any research that combines deep learning technology with spectrogram.

Spectrogram is a two-dimensional image that can reflect the essence of sound and can fully reflect the time-domain and frequency-domain information of sound [23]. The characteristics of spectrogram of different sounds are different, which is the premise of sound recognition using deep learning image processing method. There are differences in time-domain and frequency-domain information of different music types, which will be reflected in their spectrograms to a certain extent. This research combines deep learning with spectrogram and proposes a new music type recognition method. The feasibility of the method is verified by constructing the music spectrogram dataset and importing it into the deep learning model for training, so as to find a new and intelligent method for the research of national music recognition.

In this paper, from the perspective of multimodal data, music is decomposed into multidimensional data, and corresponding feature vectors are extracted for data fusion. At the same time, based on the depth confidence network, the classification of music is improved, that is, the fine-tuning node is added. The training set obtained from the fusion is trained in the improved deep confidence network, which can further improve the accuracy of music score classification. The main contributions are summarized as follows. (a) Simulate folk music of different spectrum, record corresponding music audio by Fourier transform and other methods under laboratory conditions, convert music audio into spectrum diagram, and obtain 2608 two-dimensional spectrum diagram images as datasets. Import the spectrum diagram dataset into the deep convolutional neural network GoogLeNet for music type recognition, with a test accuracy of 99.6%. (b) Parallel GoogLeNet technology based on inverse autoregressive flow is used. The unique improvement is that acoustic features can be quickly converted into corresponding speech time-domain waveforms, reaching the real-time level, improving the efficiency of model training and loading, and outputting more natural speech.

2. Design of the Classification System of the Spectrogram of the Folk Music

2.1. The Overall Framework Design of National Music Classification System. This system adopts MVC mode, which is popular in the industry, and develops the system based on B/S structure. It realizes the development goal of "high cohesion and low coupling," separates page display, business logic, and data access, and is conducive to the hierarchical development of the system. The overall frame structure design of the system is shown in Figure 1. The user operates the system through the browser, and the operation request is transmitted from the view layer to the control layer through HTTP protocol. The control layer forwards these received requests to the business layer, which adopts different business logic to respond according to the different contents of user requests. After that, the response results are transferred to the view layer through the control layer and presented to the user in the form of HTTP pages.

Taking the user's classification request for opera as an example, the user submits the folk music information to be classified to the control layer through the view layer, and then the control layer transfers it to the business layer. After determining that the request is a classification request, the business layer calls the corresponding opera classification

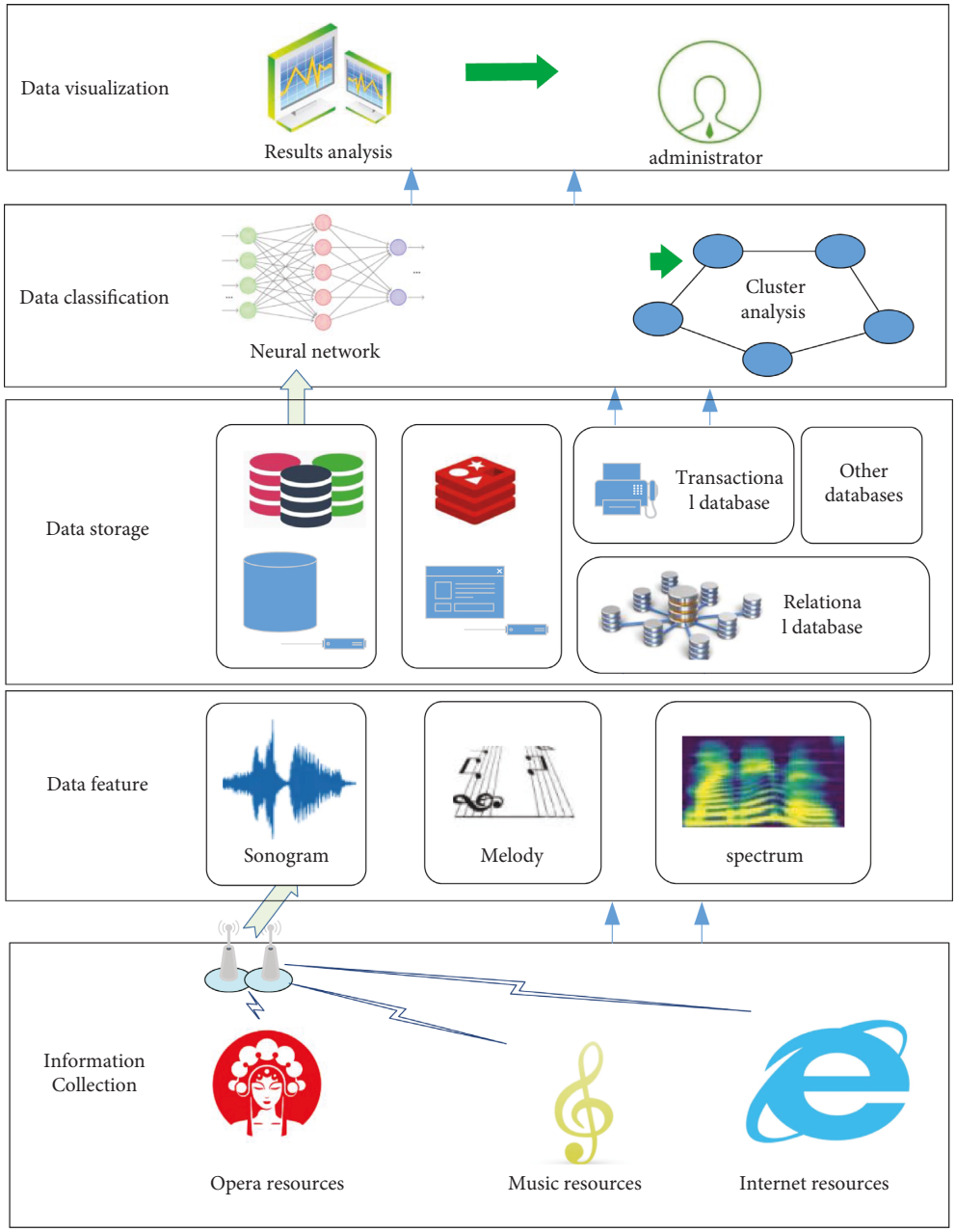


Figure 1: System architecture design figure.

module to process it. After that, the classification results are transferred to the view layer through the control layer and displayed to the user at the specified position in the browser. Because the average power spectrum of the voice part of the music signal is affected by glottic excitation and oronasal radiation, it falls by 6 dB/octave in the high-frequency band above 800 Hz. In order to make the music signal obtain the spectrum with the same signal-to-noise ratio in the whole frequency band, it is necessary to strengthen the high-frequency part of the signal, which is the purpose of pre-emphasis processing on the signal. The spectrum can be made relatively flat by preemphasis processing, which is conducive to spectrum analysis and channel parameter analysis of the signal. The process of preemphasis is as follows: first, the music signal is digitized and then filtered by a preemphasis digital filter with 6 dB/octave high-frequency characteristics. In addition, the main goal of the development of this system is to realize a national music classification system, which can provide batch opera classification and annotation services for any user who has the needs of national music sorting and classification. This system selects the mel cepstrum coefficient that reflects the timbre and the pitch frequency, formant, and band energy distribution that reflect the melody. Combined with the deep confidence network model, it trains a large number of labeled samples to obtain the ability to classify national music and saves the classification model that meets the classification requirements through the test set to the system, so as to achieve the

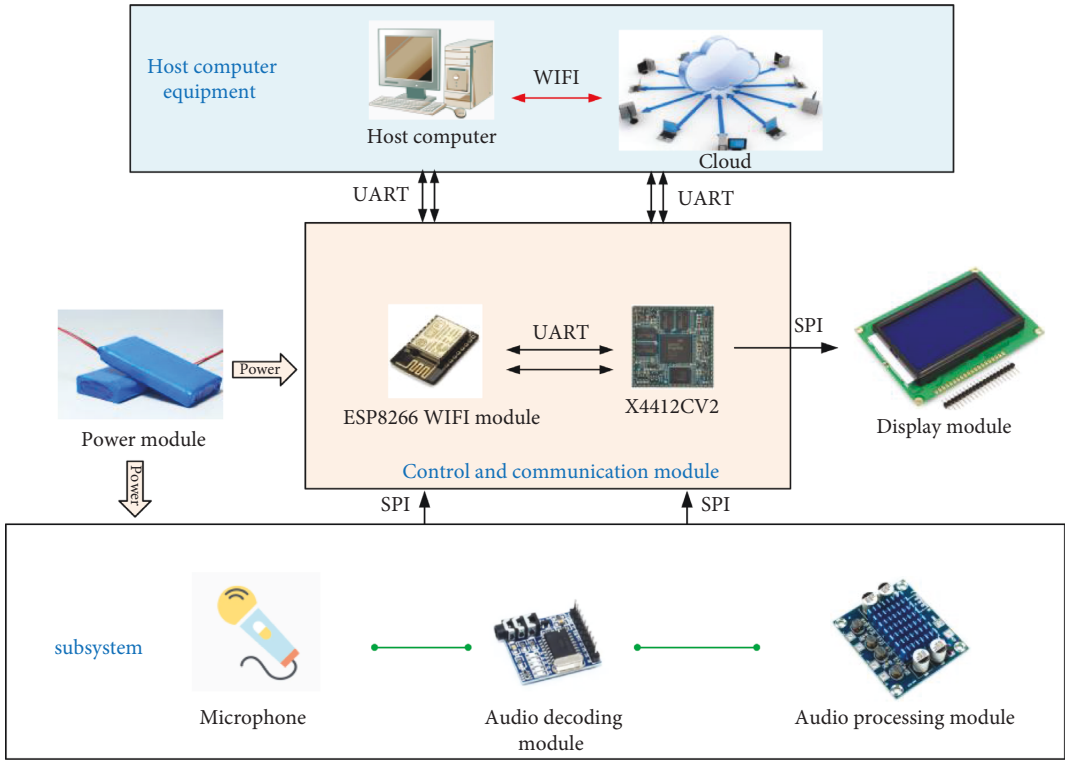


Figure 2: Schematic diagram of system hardware system.

classification goal of unknown style national music. In addition, Chinese pronunciation has very complex prosodic rules, including whether the generated speech fully combines semantics with sound correspondence to express the real semantic relationship. The main function of the prosodic processing module is to label the text prosodically and process the prosodic features that can express the semantics such as the correct tone, stress, and tone of the language in the sentence.

2.2. System Hardware Design for Reorganization System. The main controller of the wireless distributed background music system is composed of core modules and functional modules. The hardware structure is shown in Figure 2. The core module adopts X4412CV2 as the core board, and the functional modules mainly include power circuit module, SD card and USB module, audio module, WiFi module, human-computer interaction module, and so on. The LCD screen in the human-computer interaction module adopts 7-inch TFT LCD screen, which supports backlight brightness adjustment.

The core module of this design uses Chuangzhan x4412cv2 core board as the development platform, which is a core board with low power consumption, high performance, and strong scalability. Exynos 4412 with Samsung cortex-a9 architecture is used as the main processor, and the running speed is as high as 1.5 GHz. It supports EMMC of various brands and capacities, uses dual-channel DDR3 design, has power sleep wake-up function, supports Android, Linux, and Ubuntu operating systems, and supports wired Ethernet expansion. The size of the core board is only 55 mm ∗ 55 mm. Furthermore, in this system, the reception and transmission of audio data are mainly completed through the WiFi module. In embedded systems, there are many kinds of WiFi interfaces, including SPI, SDIO, and USB. Based on the analysis of comprehensive performance, this design adopts the mt6620 chip of WiFi module MTK and uses the SDIO interface type as the connection type between WiFi module and x4412 board.

3. Research on the Classification Algorithm of National Music Rhythm Spectrogram Based on Biological Neural Network

3.1. Overall Structure Design of Classification Algorithm. Using vector sequence to represent the characteristic information of some aspects of music signal is called feature extraction. The performance of music classification system can be improved only by reasonably selecting the characteristic parameters of music signals. The overall framework diagram design of the algorithm is shown in Figure 3.

As shown in Figure 3, the classification of music rhythm mainly includes the following parts: (1) the determination of music rhythm category; (2) feature extraction of music clips; (3) training of music spectrum classification model; and (4) classification of the test set using the generated model. The main research content of this paper will also be structured according to these modules. In most content-based music classification systems, music content refers to the characteristic parameters of music itself. Therefore, for a

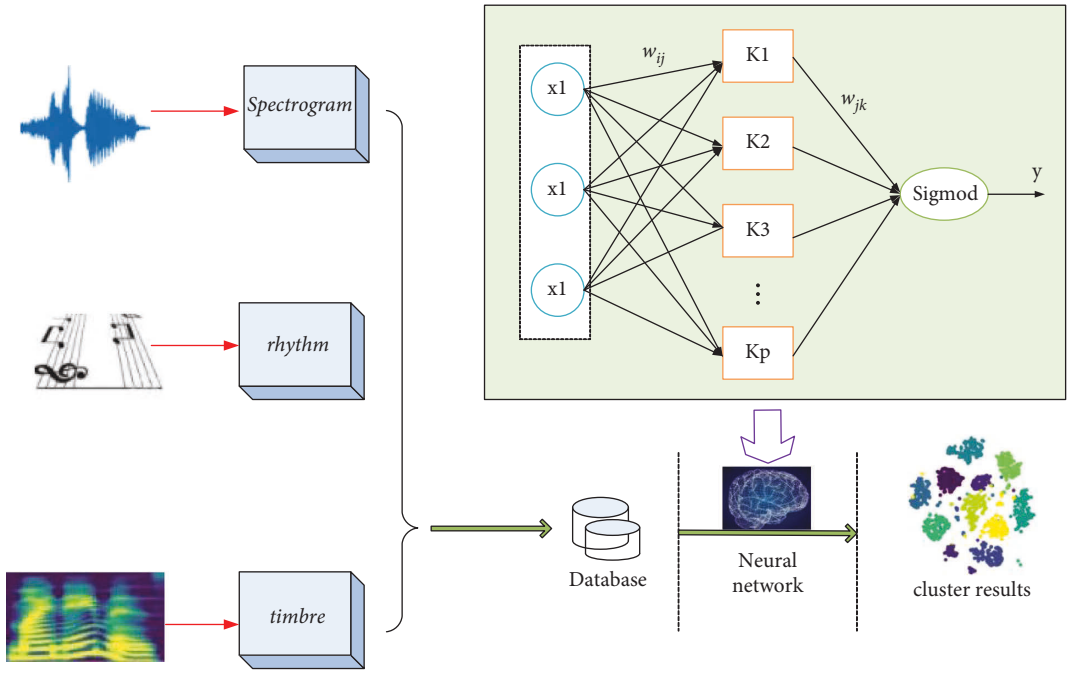


Figure 3: Overall framework of the algorithm.

classification system, the feature extraction of music signal is particularly important. This paper divides music features into three parts: bottom features, middle features, and advanced tags. The underlying music features can be sub-divided into timbre features and timing features. Timbre features are the most basic sound elements. Different sound sources produce different sound harmonics, and their timbre features are also very different. Experiments show that using timbre features can better reflect the music characteristics of the sound source, and it is easier to extract timbre features. Therefore, timbre features are usually selected as a parameter to reflect the music characteristics. Commonly used timbre feature parameters include SR, SC, and MFCC. The middle-level music features are based on the bottom-level music features, which reflect the melody, pitch, chord, and other characteristics of music. Prosody feature is another important parameter of music feature extraction, in which a large part of important information in music signal is included. Melodic features can be sub-divided into pitch, frequency, amplitude, pronunciation duration, rhythm, and so on. High-level labels can be divided into music emotion, music genre, etc. It is different from the basic and intermediate characteristics of music. The content of this part is usually not directly available, and the problem that it cannot directly correspond to the music signal is called the semantic gap problem. The purpose of extracting music signal features is to abstract the high-level label information of people on this music signal through these signal features.

In order to ensure sufficient model representation space and improve the recognition effect of the model, the basic prediction unit in this paper is audio segment, that is, it does not use full length music samples for prediction but uses short audio segments with a certain length for prediction. The model predicts each audio segment during training and prediction and takes the average value of the prediction results as the prediction result of the whole training sample. This method not only greatly reduces the requirement for input size but also enables the model to focus more on the capture of local features.

The general music classification system only selects one of these features as the music feature parameter. Although it also contains some music feature information, it will inevitably have some shortcomings, such as monotonous music information form, incomplete information content, and so on, which will affect the subsequent classification accuracy. This paper adopts the method of combining multiple features to extract the timbre features of the underlying music and the melody features (pitch frequency, formant, and band energy) of the middle-level music features from the original songs and then introduces the training set composed of these features to improve the accuracy of the classification system. The extraction of these two types of features is the integration of feature parameters extracted from music signals in time domain and frequency domain.

CNN deep neural network mainly combines generative pretraining and discriminant methods to process DBN, which is used to classify tasks and change all weights, which can improve the performance of the music classifications. The principle of discriminant scheme is as follows. Based on the existing network, attach a layer of nodes at the last layer, so as to achieve the purpose of fine-tuning.

3.2. Feature Extraction of Music Spectrogram. Audio signals convey linguistic and nonlinguistic information. The former refers to the semantic expression content delivered by the speaker, while the latter refers to the acoustic changes

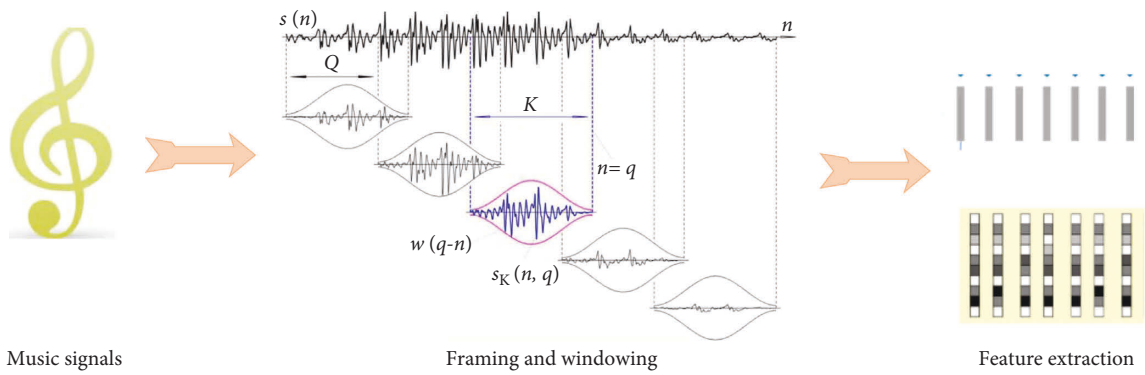


Figure 4: Flowchart of music feature extraction.

produced by the vocal system while the speaker completes the semantic expression. The two levels of information conveyed by voice and audio signals play a certain role in human emotional expression. In this paper, the audio of folk music is converted into corresponding spectrograms to construct datasets. Spectrogram is a two-dimensional image with color or gray changes that can fully reflect the essence of sound. The flowchart of music feature extraction is shown in Figure 4.

Spectrogram can reflect the time-domain information and frequency-domain information of sound at the same time, and it is highly intuitive. There are certain differences in the characteristics of spectrogram of different sounds. Convert audio into spectrogram, and the parameters used in the function are as follows: sampling frequency f_s, overlap length (L_0), Hamming window of Fourier transform (w), and number of points of Fourier transform n. Establish the short-time Fourier transform function of input signal x:

$$\text{specgram}\left(x\left(:,1\right), N, f_s, w, L_0\right). \tag{1}$$

Taking time as the horizontal axis, the size k of the horizontal axis data is calculated by using the dimension N_x of the input data, the Hamming window length $l(w)$ during Fourier transform, and the window overlap length (L_0):

$$k = \frac{\text{fix}\left(N_x - L_0\right)}{l(w) - L_0}. \tag{2}$$

Taking frequency as the vertical axis, the magnitude of frequency t is determined by the number of points N of Fourier transform:

$$t = \begin{cases} \dfrac{N}{2+1}, & N = \text{odd number}, \\ \dfrac{(N+1)}{2}, & N = \text{even number}. \end{cases} \tag{3}$$

In speech analysis, pitch generally refers to the sound produced by the vocal cord vibration when people pronounce voiced sound. When people pronounce unvoiced sound, because the vocal cord does not vibrate, it can be considered that this audio does not contain pitch. The pitch frequency is the reciprocal of the vocal cord vibration frequency when a person utters voiced sound. Fundamental frequency is widely used in speech analysis and synthesis and speech recognition. Especially, some researchers are also analyzing the relationship between pitch frequency and spectral prosody. Since intonation is a time function of the fundamental frequency of voice and intonation can reflect the emotional state of the speaker, the change of fundamental frequency is also closely related to the spectrum. Next, the effect of the fundamental frequency of the audio signal on the extraction of spectral features is analyzed from the mean, range, variance, and other aspects of the fundamental frequency trajectory. This paper selects the method of autocorrelation function detection to extract pitch frequency. The function $R_n(k)$ of signal S_n is defined as

$$R_n(k) = \sum_{m=0}^{N-k-1} S_n(m) S_n(m+k). \tag{4}$$

That is,

$$S_n(m) = s(m) w(n-m), \tag{5}$$

where $w(n-m)$ is a window function.

3.3. Algorithm Design of Deep Neural Network. This paper designs the CNN structure figure as shown in Figure 5. Its core architecture includes nine inception modules. The classification algorithm reduces the number of channels through convolution, and computes after aggregating information. It effectively uses the computing power and obtains good classification performance while controlling the amount of computation and parameters. CNN can save parameters, speed up operations, and reduce overfitting The convolution and pooling operations of different scales are fused to fuse multidimensional features to further improve the recognition and classification effect. The network model of classification algorithm can fuse multiscale features and obtain good classification results. Therefore, this experiment applies it to this music classification task. Deep confidence network, abbreviated as DBN [14], is usually composed of multilayer RBW models, and the number of layers determines the accuracy of training, but it is also prone to overfitting. After a RBM is trained, the activation probability

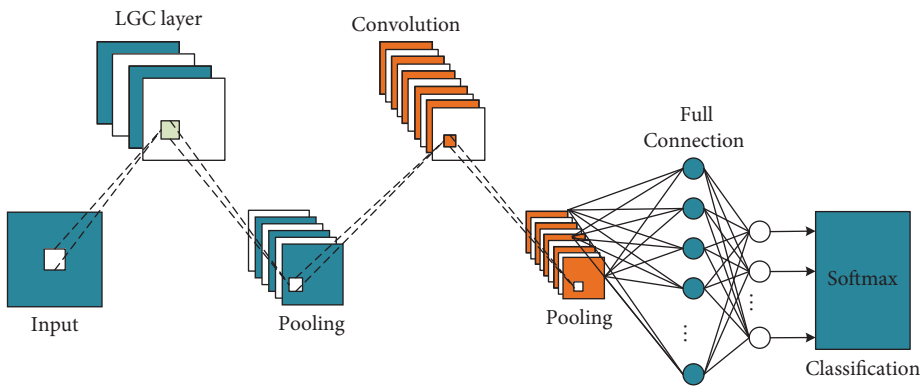


Figure 5: Structural design figure based on convolutional neural network.

of its hidden unit will be used as the input data of the next RBM, in this order until all RBM models are trained. This training process is called generative pretraining. According to the data characteristics required for music emotion analysis, this paper improves the traditional DBN training method and adds the process of discrimination and fine-tuning on the basis of generative pretraining. The specific method is to add a node in each RBM hidden layer, improve the training accuracy of the whole model by fine-tuning the value of the ownership value, and provide labels of training data. The DBN model is composed of N layers of improved RBM, one layer of traditional RBM, and one layer of softmax. The first layer of RBM is the input layer of the DBN model and has n input vectors. Softmax layer is the output layer of DBN model, with m nodes representing m categories of music emotion.

LGC layer includes a set of predefined sparse Gabor filters, a nonlinear activation function, and a set of learnable linear weights. Due to the directional selectivity of Gabor, LGC layer can extract spatial texture features in all directions. LGC layer is actually composed of two convolution layers. The first layer performs convolution on the input image. Its convolution kernel is a set of fixed sparse Gabor kernels defined in advance, and the weight of this layer is fixed, which is not learnable in the process of training. The convoluted feature map is mapped by the nonlinear activation function, and the activation function used here is relu. In the second layer, the mapped features are linearly weighted to form the final feature map, which is used as the input of the next layer. This step consists of 1×1, and the weight of this layer can be learned in the process of training. Compared with the standard convolution layer, LGC layer has fewer learnable parameters with the same convolution kernel size and the same number of input and output channels. Suppose the number of channels of input data is p, the number of output channels is Q, and the size of convolution kernel is $n \times m$. According to the convolution mechanism, the standard convolution layer requires $P \times n \times m \times Q$ learnable parameters, while the LGC layer requires $p \times n \times m \times W$ fixed parameters and $W \times Q$ learnable parameters, where W represents the channels of the middle layer and $w \times Q$ corresponds to a convolution kernel size of $1 \times$ convolution operation of 1.

4. Experimental Results and Analysis

4.1. Data Source and Simulation Environment Settings. At present, there are several mature music datasets in the market, but most of these datasets only contain English tracks. Datasets specially built for Chinese traditional folk music are rare in the market. Therefore, to experiment with the classification model, we first need to build a folk music dataset. This paper uses web crawler tools to crawl music information and files from Chinese opera website, drama website, drama house, music, and other websites to the local way to build the database. Due to the size of the dataset, the MSD dataset itself does not provide original audio samples but provides audio feature data such as mel frequency cepstrum coefficient that has been calculated according to preset parameters, which does not conform to the experiment in this paper. Therefore, this paper uses the metadata provided by the MSD dataset to build a sub-dataset containing original audio samples.

The music clips used in this article are unified as 30 seconds, wav format, 16000 Hz sampling rate, mono. Because the first paragraph of the song has too much accompaniment, the last paragraph is almost less than 30 seconds. In this way, there are 6200 30 s music clips in total. At this time, the 3200 music clips generated are considered our preliminary music database because there may be duplicate clips for each music. By further deleting the duplicate clips of 6200 music clips, 4800 music clips are finally retained. The folk music files crawled from the network in this paper contain 4 types and 4800 tracks. Among them, there are 2600 songs and dance music, 800 kinds of rap music, 400 kinds of opera music, and 1000 kinds of instrumental music. After the selection of the music library is completed, each music segment will be manually calibrated next. Due to the subjectivity of emotional calibration of music, in order to enhance the credibility of the whole music library, the authors of this paper adopted the method of multiple people calibrating the same music segment.

In order to further verify the reliability of the experiment, the sonogram datasets were imported into the pretrained Resnet18 and Shufflenet for training. Resnet18 introduces the depth residual module to solve the gradient disappearance problem with the deepening of the network layers. After the

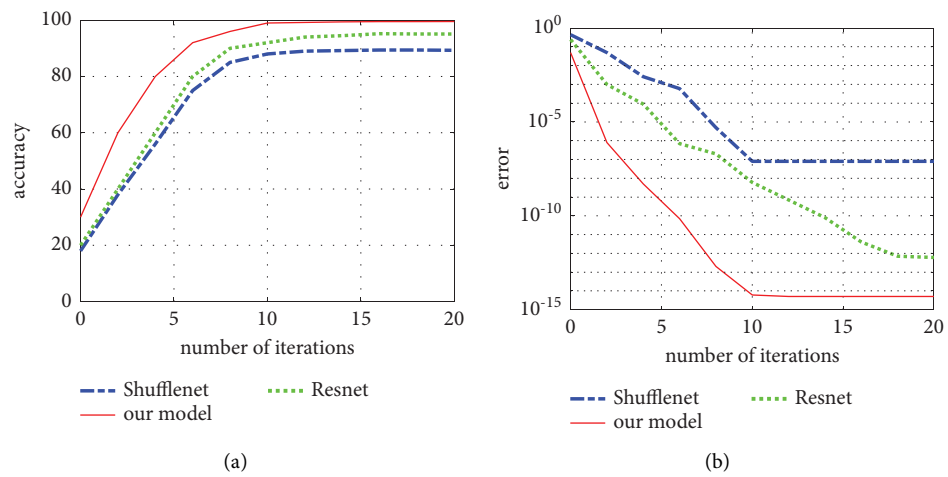


Figure 6: Comparison of training results of the three models. (a) Accuracy. (b) Convergence error.

convolution operation, a jump connection is added. When the accuracy decreases due to the deepening of the network layers, it can return to the shallow network to solve the gradient disappearance problem. Although Resnet18 has advantages in classification tasks by virtue of residual module, its network is relatively deep and its parameters are large. Shufflenet is one of the mainstream lightweight network models. The core operation of Shufflenet is to shuffle different channels, so as to give better play to the advantages of group convolution and further realize the balance between model lightweight and model performance improvement.

Software environment: using Python language, with the help of pycharm development environment and librosa speech extraction toolkit. The DBN model consists of N layers of RBM with fine-tuning nodes and one layer of unmodified RBM. Among them, the bottom input has 10 nodes, corresponding to 10 dimensional original data feature vectors. The number of cycles of pretraining is set to 100 times, and the number of fine-tuning cycles is set to 150 times. The top output layer has 5 nodes, corresponding to five types of folk music. Hardware configuration includes 16g memory, Intel Core i7-7700 processor, and NVIDIA 3050ti graphics card.

4.2. Convergence Verification of the Model. The system can test songs in two formats. Firstly, all songs under the whole testing folder are adopted, and finally count the number of classification errors of different categories of songs and the average accuracy of the classification of the last four categories of songs. The second test format can classify a music clip with unknown emotion category and get the category attribute of this music clip.

In this experiment, Resnet18 and Shufflenet are introduced into the music type recognition task to verify the reliability of the experiment, and the recognition effect of the model in this paper is compared with the large-scale model Resnet18 and the lightweight model Shufflenet. The data are imported into Resnet18 and Shufflenet, respectively, and the corresponding validation accuracy training chart, loss error

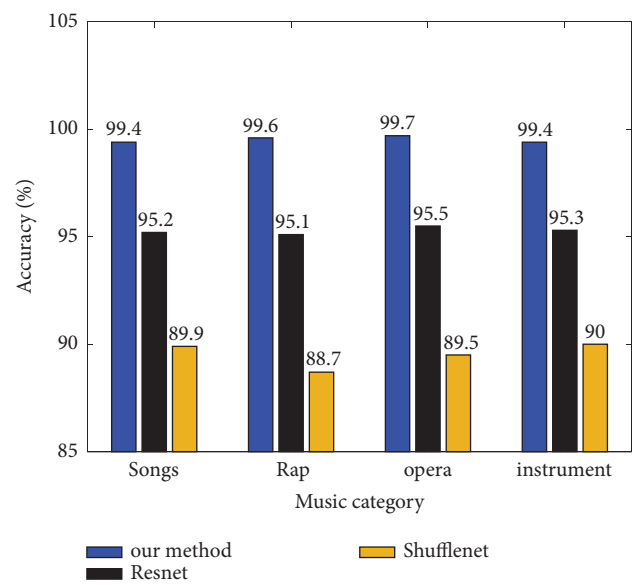


Figure 7: Accuracy verification of three network models.

training chart, and confusion matrix are obtained after the training of 20 epochs. Three network specific training parameters: the minimum batch is 64 times, the maximum iteration is 20 times, and the initial learning rate is 1×10^{-4}, the verification frequency is 36 Hz, and the image input size is 224px × 224px.

The accuracy curve of validation accuracy training and validation is shown in Figures 6(a) and 6(b). The validation accuracy of thenetwork models in this paper, Resnet18 and Shufflenet, is 99.60% and 98.85%, both of which have high classification accuracy. The training process of the network model in this paper tends to be stable after the 20th epoch, and the accuracy of 99.60% is finally obtained, which shows that the method can effectively recognize music categories. Only two epochs are needed to improve the accuracy from 20% to 90%, which shows that the model has a high convergence rate. The model training accuracy curve tends to be consistent with the verification accuracy curve, and the loss rate loss tends to be

Table 1: Model performance evaluation index.

Index	Our model			Resnet			Shufflenet		
	Sensitivity (%)	Specificity (%)	Accuracy (%)	Sensitivity (%)	Specificity (%)	Accuracy (%)	Sensitivity (%)	Specificity (%)	Accuracy (%)
Group 1	100	100	99.4	100	97.0	95.2	96.1	93.2	89.9
Group 2	100	100	99.6	99	96.0	95.1	94.3	92.4	88.7
Group 3	100	99.4	99.7	98	95.4	95.5	91.6	90.8	89.5
Group 4	99.3	99.5	99.4	96	95.0	95.3	90.4	89.7	90.0

close to 0 after the second epoch, which indicates that there is no fitting or underfitting phenomenon in the model training, which verifies the reliability of the experimental data. The simulation results in the above two cases (that is, the case of distinguishing the singer's gender and the case of not distinguishing the singer's gender) show that the performance of the GoogLeNet classification algorithm is better than that of the Shufflenet and Resnet classification algorithms. This is because the SVM method directly looks for an optimal hyperplane to distinguish the two categories of music and uses it as the final classifier. However, emotion classification of music itself has great fuzziness, and artificial emotion calibration of music has great subjectivity. The same music segment may be classified into different emotions. It is difficult to find a strong classifier with high classification accuracy directly for music emotion classification with large fuzziness, and the advantage of Resnet algorithm is to combine multiple weak classifiers to generate a strong classifier, which greatly improves the final classification performance. In order to verify the effectiveness of the two-tier classification system in this paper, the experimental results using the system structure in literature [18] and the system structure in this paper are compared, and it is found that the classification effect of the two-tier classification system structure in this paper is better.

In order to further verify the reliability and superiority, this paper classifies and verifies four types of folk music. Accuracy verification results of three network models are shown in Figure 7. The accuracy of GoogLeNet classification algorithm is significantly higher than that of Shufflenet, which also proves that the classification accuracy of finding a combination of multiple weak classifiers to form a strong classifier is significantly improved than that of finding a strong classifier alone. Of course, it is also a difficult task to distinguish them. The accuracy of this model is the highest whether using GoogLeNet algorithm or Shufflenet algorithm. For example, the first column of the confusion matrix indicates that 2480 samples are predicted as song types. From the data of confusion matrix, it can be calculated that the model test in this paper has achieved a high accuracy of 99.5%. Only 120 song type samples were mispredicted. The results show that this model has high accuracy in four categories: song, rap music, opera music, and instrumental music. Furthermore, in the process of training and verification, in order to ensure the consistency with the training parameters of GTZAN dataset, all samples with a time length of less than 30 seconds at the sampling frequency of 22050 Hz are eliminated in this paper. According to statistics, after the elimination according to the above rules, the original training set, development set, and evaluation set are eliminated by 5, 5, and 3 samples, respectively. In addition, other training details and training parameters are the same as those used in the GTZAN dataset. In the 5×10 fold cross-validation, the classification accuracy was the highest, reaching 87.68%, $5\times$ the average accuracy rate of 10-fold cross-validation is 87.11%. The number of national music classifications in the ISMIR2004 dataset is 5. The classification effect of Google network is better than that of Shufflenet and Resnet. This observation result is similar to that in the GTZAN dataset.

The sensitivity, specificity, accuracy, and other indicators of the test results of the three network models are shown in Table 1. By analyzing the model performance evaluation indicators such as sensitivity, specificity, and accuracy, it is found that the performance indicators of all models are above 97%, which indicates that the network model used in this experiment has good performance, which further verifies the accuracy of the network model in this paper.

5. Conclusion

With the development of various emerging art forms, the situation of traditional folk music is becoming increasingly severe. How to continue our unique traditional folk music is an urgent problem. This paper constructs a classification system of national music rhythm spectrogram based on biological neural network and uses classification technology to classify a large number of unorganized national music documents without labels, which will help national music better enter the vision of ordinary people to a certain extent. Firstly, the music information collected in the laboratory is converted into sound spectrum images, and the dataset containing 4800 sound spectrum images is obtained. Then, import the dataset into the deep convolution neural networks Resnet18 and Shufflenet proposed in this paper, and the test accuracy of 99.6%, 99.2%, and 99.2% is obtained. The experimental results show that the proposed network model has high recognition accuracy and can complete the classification of folk music. Music classification has important application value in multimedia applications. It shows that this method is reasonable and has better classification performance. The main reason is that there are fewer types of audio signal features, which limits the type resolution of music emotion. In the future, we will add other audio signal features and combination methods to the research of music emotion classification.

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 Department of Music, Xinxiang University.

References

[1] E. B. Panganiban, A. C. Paglinawan, W. Y. Chung, and G. L. Paa, "SECG diagnostic support system (EDSS): a deep learning neural network based classification system for detecting ECG abnormal rhythms from a low-powered wearable biosensors," *Sensing and Bio-Sensing Research*, vol. 31, Article ID 100398, 2021.

[2] J. R. Castillo and M. J. Flores, "Web-based music genre classification for timeline song visualization and analysis," *IEEE Access*, vol. 9, pp. 18801–18816, 2021.

[3] Y. J. Liao, W. C. Wang, S. J. Ruan, Y. H. Lee, and S. C. Chen, "A music playback algorithm based on residual-inception blocks for music emotion classification and physiological information," *Sensors*, vol. 22, no. 3, p. 777, 2022.

[4] B. Tiple and M. Patwardhan, "Multi-label emotion recognition from Indian classical music using gradient descent SNN model," *Multimedia Tools and Applications*, vol. 81, no. 6, pp. 8853–8870, 2022.

[5] Y. Khalifa, D. Mandic, and E. Sejdić, "A review of hidden markov models and recurrent neural networks for event detection and localization in biomedical signals," *Information Fusion*, vol. 69, pp. 52–72, 2021.

[6] K. A. Pati, S. Gururani, and A. Lerch, "Assessment of student music performances using deep neural networks," *Applied Sciences*, vol. 8, no. 4, p. 507, 2018.

[7] K. Zhang, G. Xu, Z. Han et al., "Data augmentation for motor imagery signal classification based on a hybrid neural network," *Sensors*, vol. 20, no. 16, p. 4485, 2020.

[8] N. N. J. Siphocly, E. S. M. El-Horbaty, and A. B. M. Salem, "Top 10 artificial intelligence algorithms in computer music composition," *International Journal of Computing and Digital Systems*, vol. 10, no. 1, pp. 373–394, 2021.

[9] M. Radha, P. Fonseca, A. Moreau et al., "Sleep stage classification from heart-rate variability using long short-term memory neural networks," *Scientific Reports*, vol. 9, no. 1, pp. 14149–14211, 2019.

[10] F. Merchan, A. Guerra, H. Poveda, H. M. Guzman, and J. E. Sanchez-Galan, "Bioacoustic classification of antillean manatee vocalization spectrograms using deep convolutional neural networks," *Applied Sciences*, vol. 10, no. 9, p. 3286, 2020.

[11] S. Nag, M. Basu, S. Sanyal, A. Banerjee, and D. Ghosh, "On the application of deep learning and multifractal techniques to classify emotions and instruments using Indian classical music," *Physica A: Statistical Mechanics and Its Applications*, vol. 597, Article ID 127261, 2022.

[12] K. Zhao, H. Jiang, Z. Wang, P. Chen, B. Zhu, and X. Duan, "Long-term bowel sound monitoring and segmentation by wearable devices and convolutional neural networks," *IEEE Transactions on Biomedical Circuits and Systems*, vol. 14, no. 5, pp. 985–996, 2020.

[13] S. Hong, Y. Zhou, J. Shang, C. Xiao, and J. Sun, "Opportunities and challenges of deep learning methods for electrocardiogram data: a systematic review," *Computers in Biology and Medicine*, vol. 122, Article ID 103801, 2020.

[14] D. Shah, A. Narayanan, and J. I. Espinosa-Ramos, "Utilizing the neuronal behavior of spiking neurons to recognize music signals based on time coding features," *IEEE Access*, vol. 10, pp. 37317–37329, 2022.

[15] M. S. Fernandes, W. Cordeiro, and M. Recamonde-Mendoza, "Detecting *Aedes aegypti* mosquitoes through audio classification with convolutional neural networks," *Computers in Biology and Medicine*, vol. 129, Article ID 104152, 2021.

[16] C. Ieracitano, N. Mammone, A. Bramanti, A. Hussain, and F. C. Morabito, "A convolutional neural network approach for classification of dementia stages based on 2D-spectral representation of EEG recordings," *Neurocomputing*, vol. 323, pp. 96–107, 2019.

[17] J. Gauer, A. Nagathil, K. Eckel, D. Belomestny, and R. Martin, "A versatile deep-neural-network-based music preprocessing and remixing scheme for cochlear implant listeners," *Journal of the Acoustical Society of America*, vol. 151, no. 5, pp. 2975–2986, 2022.

[18] G. Korvel, P. Treigys, and B. Kostek, "Highlighting interlanguage phoneme differences based on similarity matrices and convolutional neural network," *Journal of the Acoustical Society of America*, vol. 149, no. 1, pp. 508–523, 2021.

[19] J. Ramírez and M. J. Flores, "Machine learning for music genre: multifaceted review and experimentation with audioset," *Journal of Intelligent Information Systems*, vol. 55, no. 3, pp. 469–499, 2020.

[20] J. Li, L. Han, and Y. Wang, "Combined angular margin and cosine margin softmax loss for music classification based on spectrograms," *Neural Computing & Applications*, pp. 1–17, 2022.

[21] H. Tang, Y. Zhang, and Q. Zhang, "The use of deep learning-based intelligent music signal identification and generation technology in national music teaching," *Frontiers in Psychology*, vol. 13, Article ID 762402, 2022.

[22] C. Liu, L. Feng, G. Liu, H. Wang, and S. Liu, "Bottom-up broadcast neural network for music genre classification," *Multimedia Tools and Applications*, vol. 80, no. 5, pp. 7313–7331, 2021.

[23] W. W. Y. Ng, W. Zeng, and T. Wang, "Multi-level local feature coding fusion for music genre recognition," *IEEE Access*, vol. 8, pp. 152713–152727, 2020.

Hindawi
Computational Intelligence and Neuroscience
Volume 2022, Article ID 3929110, 10 pages
https://doi.org/10.1155/2022/3929110

Research Article

Application of Neural Network Based on Visual Recognition in Color Perception Analysis of Intelligent Vehicle HMI Interactive Interface under User Experience

Dongxin Zhao

Huita Information Technology Consulting (Shanghai) Co.,Ltd, Shanghai 201210, China

Correspondence should be addressed to Dongxin Zhao; anni.zhao@facecar.org

Received 27 July 2022; Revised 30 August 2022; Accepted 19 September 2022; Published 12 October 2022

Academic Editor: Ning Cao

As a bridge of human-computer communication, the color design of intelligent vehicle HMI interactive interface is particularly important. It is also the first guide to the driver during the driving process. The quality of its design will also directly affect the driver's senses and the driving safety of the vehicle. Therefore, this paper introduces the current situation, design principle, and future development of the vehicle interaction interface from multiple perspectives. Through the neural network system (condition generation countermeasure network model) of visual recognition, the color of the intelligent vehicle HMI interactive interface under the user experience is analyzed. According to the analysis of the psychological cognition and behavior operation of the automobile user, the correlation analysis of the human, vehicle, environment, and various elements of the interface is carried out, and how the vehicle interactive interface can meet the expected physiological and psychological needs of the user more and improve the operability is discussed in order to design an on-board HMI interactive interface that can be intelligently perceived according to weather, driver's interests, and other factors and then improve the current backward operation mode of the on-board interactive interface, so that the interaction between people and vehicles is more smooth and pleasant.

1. Introduction

Color perception is an important function of the human visual system, and it is also one of the most important ways for the human nervous system to receive external information. The research entity formed by separating the part responsible for color perception in the human visual system is called the color perception system.

The early research and attention on the human color vision system can be traced back to the relevant literature on the discovery of color blindness by Dalton, a famous British scientist, at the end of the 18th century [1–3], which first revealed and described the specific subtle differences between the color vision system of a small number of human individuals and normal people, thus initially forming the concept of normal color vision and abnormal color vision, attracting the attention of academia to the differences between color vision systems [4, 5].

We live in a world full of color. Anything needs color decoration. Color can directly or indirectly affect people's emotions [6–10]. In the long-term life of human beings, the application of color makes it a visual symbol system, and it has certain symbolism and stability. Color is a physical phenomenon, which should have no emotional expression, but in people's long-term application, many experiences and memories are formed in people's minds. When this intuitive perception and memory collide with external color stimulation in the heart, it gives color the corresponding emotion [2, 11, 12]. Color itself is expressionless, but in the long-term life practice and experience, human beings give color unique expression characteristics. As the color changes, the color expression will change accordingly. People use some typical color expressions to create things that meet people's needs [13–16].

With the rapid development of the computer system, the automobile has become more and more intelligent. From the original simple gasoline engine, it has developed into a new

product combining science and technology with wisdom. With the diversification of functions, complicated operation modes have been added. More and more car factories and emerging companies have paid attention to the informatization, interconnection, and intelligence of cars. It is no longer a concept but reflected in products and services. As an information-based product, on-board equipment has been widely used by users in daily life [17–19]. Interaction design in the car environment is very different from mobile phones and computers. The time that eyes can stay on the screen is short, and the complex environment during driving requires designers to constantly study and deepen in the process of interface design [20–22]. However, at present, most automotive systems and on-board products in the domestic market cannot meet the needs of today's users, making many seemingly powerful on-board devices difficult to operate and understand in the actual use of users. The unreasonable software interface will cause cognitive burden and operation difficulties to the driving driver when used, thus posing a certain threat to the driving safety of the driver [23, 24]. Through the research on the information visualization design of the on-board interactive interface, the usability level of the on-board interactive interface can be improved, and the interface operation of the on-board product is simpler and more effective. At the same time, it also has a certain guiding significance for the design and research of the software interface [25]. The design of automobile on-board interactive interface is relatively novel in the field of interaction. It is of certain research value to locate the interactive scenario in the automobile and deeply explore the interaction process between the driver and the on-board interface [26].

Nowadays, cars have become indispensable in people's daily life. With the rapid development of science, cars are also gradually moving towards information and intelligence, and the human-computer interaction mode has also been fully used in the automotive industry. In order to improve the safety and comfort of drivers when driving, automobile manufacturers have improved and optimized the design of on-board systems, making them integrate navigation, entertainment, and other functions. However, at present, most on-board systems only meet the functional requirements and lack attention to availability and ease of use.

Therefore, based on the elaboration of interaction design and user experience, this paper analyzes the design principles of vehicle interaction interface and uses the neural network of visual recognition to explore the design method of vehicle interaction interface that can combine user characteristics and meet user needs, so as to improve user experience.

2. Research on the Theory of Conditional Generation Countermeasure Network and Vehicle Interaction

2.1. Conditional Generation Countermeasure Network. Generating confrontation network is an unsupervised learning method of deep learning, which was first proposed by Goodfellow et al. According to its design idea based on game theory, it is composed of two parts, one is called a generator and the other is called a discriminator. The generator is used to generate false data, and the discriminator determines whether the source of the data is true or false. In the original generation countermeasure network proposed by Goodfellow et al., the input of the generator is a one-dimensional random noise vector, which is finally output as an image through a multilayer fully connected neural network.

The generation countermeasure network is widely used in image generation tasks because it can synthesize images with higher accuracy and quality and has made breakthroughs in image generation, image style conversion, and image segmentation [27, 28]. The generation countermeasure network consists of two parts, one is the generator (g), the other is the discriminator (d). The generation network G and the discrimination network D iterate with each other until the generation network G and the discrimination network D are in balance with each other.

Generate a confrontation network and define the following formula to update the weight as a learning rule:

$$L_{GAN}(G,D) = \min_G \max_D \Big[E_{x\sim P_{data}(x)}\log(D(x)) + E_{z\sim P_{data}(z)}\log(1-D(G(z)))\Big], \tag{1}$$

where $\min_G$ is the minimum value of the generated network error; $\max_D$ is the maximum value of network error; $z \sim P_{data}(z)$is the real sample distribution; $x \sim P_{data}(x)$ is the random noise sample distribution; and E is the calculated probability mean.

In the existing generation countermeasure network, the generation network G collects noise Z as input in the hidden space. When it passes through the full connection layer, it will generate images. The discrimination network D is used to judge the real images and pseudo images, but there is no strict correspondence between the input and output [28, 29]. Therefore, the conditional generation countermeasure network updates the weight of the existing generation countermeasure network through the following learning rules:

$$L_{cGAN}(G,D) = \min_G \max_D \Big[E_{x\sim P_{data}(x)}\log(D(y,x)) + E_{z\sim P_{data}(z)}\log(1-D(G(z,y)))\Big], \tag{2}$$

where y is the label information corresponding to the real sample data. Adding y label to Gaussian noise z can make the result of generating countermeasure network controllable.

Pix2pix replaces noise and condition information with the mask image on the basis of *cGAN* to realize image to image conversion. In order to ensure that the similarity between input and output images is large, after generating the countermeasure loss, the countermeasure network error is generated through the regular term error control [28].

$$G' = \arg\min_G \max_D L_{cGAN}(G,D) + \lambda L_{L1}, \tag{3}$$

where G' is the total loss function of Pix2Pix.

The conditional generation countermeasure network is a variant of the original generation countermeasure network architecture. Its main feature is that the condition information is added to the input of the generator so that the generator generates data according to the conditions. In the discriminator, it can also be divided into conditional discrimination and nonconditional discrimination. The conditional discrimination is to fuse the generated data of the conditional information horse and then input it to the discriminator for correlation discrimination. The nonconditional discrimination is to input the generated data to the discriminator separately, and the discriminator makes a separate judgment on it. In the process of model construction and training, the network can generate the generated data consistent with the condition information by designing a reasonable loss function for learning.

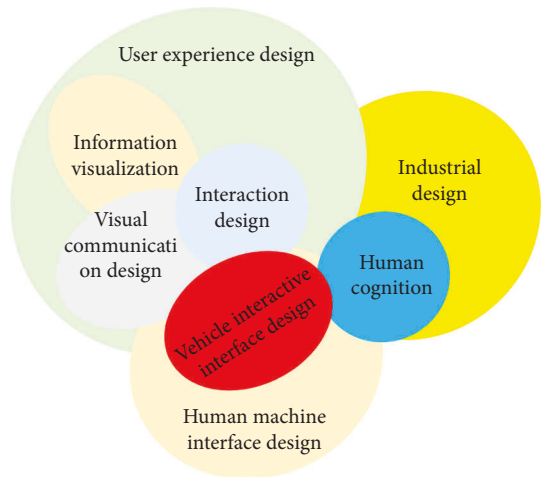


Figure 1: The relationship between vehicle interactive interface design and various disciplines.

2.2. Vehicle Interaction Design Theory

2.2.1. Characteristics of Information Visualization in Vehicle Interactive Interface. The basic way for people to communicate in life is words and graphics. Since birth, people have used senses, imagination, and emotions and just have some form of dialogue with the surrounding products and environment. The design of vehicle interaction interface is a part of human-computer interaction design, and its relationship with various design disciplines is shown in Figure 1, forming an overlapping relationship with most mathematics disciplines. Information visualization design and interaction design are closely linked. Information visualization design solves the problem of information expression. On this basis, vehicle interaction design provides users with a better way to interact with information.

James Jerome Gibson, an American perceptual psychologist, believes that all the characteristics of objects can be intuitively perceived, and their information is directly expressed in vision. For example, people judge how to open a door according to the shape of the door handle. An object has more than one purpose. People will arrange their behavior according to the characteristics of the object itself. In the vehicle interactive interface, information visualization not only needs to follow the principle of information visualization design visually but also conforms to the object entity, environment, and information in the vehicle in terms of information architecture. The design of vehicle interaction interface is based on functional interface, with environmental interface as the premise and emotional interface as the center. It is clear, concise, familiar, easy to respond, easy to operate, efficient, and fault-tolerant. In the process of information transmission in the design of vehicle interactive interface, people form behavior through brain judgment and transmit it to the vehicle machine. The vehicle machine processes the input information through CPU and outputs information through color, sound, animation, shape, text, image, and so on in the interactive interface. The general information transmission process of on-board interactive interface is shown in Figure 2.

2.2.2. Definition of On-Board Interactive Interface. The original car was controlled by the tiller, but in 1894, Alfred Vacheron introduced a gear system between the driver and the wheels and installed a steering wheel, which is considered to be one of the earliest human-computer interaction control devices. With the progress of time, speed indicators and fuel indicators have also been used in automobiles, becoming the earliest human-computer interactive display devices. Nowadays, all kinds of information and equipment are pouring into the vehicle, and the human-computer interaction in the vehicle becomes more and more troublesome. However, because the iteration speed of the software system is far faster than that of *R&D* of the automobile enterprises, the human-computer interaction development in the automobile is not as fast and perfect as the mobile phone interaction. Since the automobile has been well known by the public, the automobile, as a means of transportation for people, has not changed much in essence, but users' expectations for the automobile will be higher and higher, and car companies tend to be more functional and intelligent when designing cars. Panoramic sunroof, fingerprint identification, head up display, heated seats, and other new technologies are used in cars. All kinds of operations in the whole process of using the car from the user belong to the category of car interaction, and the design of on-board interactive interface pays more attention to the interactive interface. The interface is the contact level in the process of interaction between the human and machine. The interactive interface is expressed through sensory experience, emotional experience, and cultural experience. The sense of security and control obtained by the user through visual hearing reflects the user's concept, consciousness, lifestyle, and other cultural connotations. The user carries out an operation, and the operation result or state is feedbacked by the vehicle machine. After the user perceives the change of information through vision, hearing, touch, etc., he analyzes and explains the meaning of information through thinking and finally makes a decision. The on-board

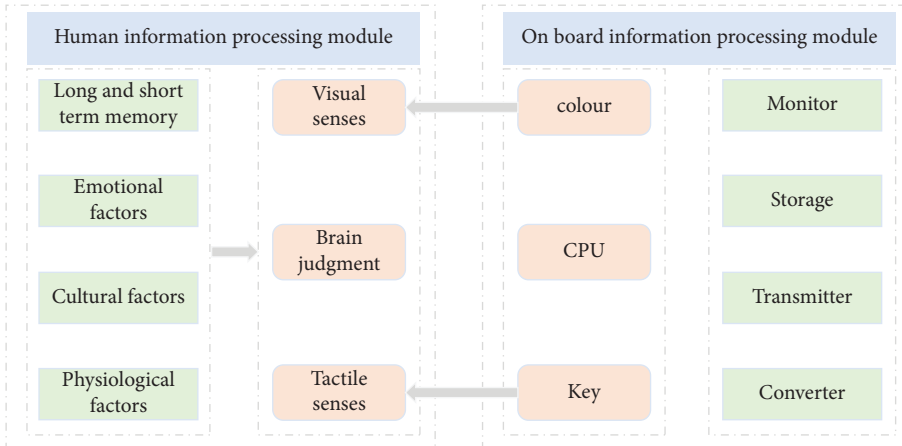


Figure 2: Information transmission process in vehicle interactive interface.

interactive interface is the link between people and vehicles. Good interface design can make communication more accurate and pleasant.

Vehicle interface can be roughly divided into three types: hard interface, soft interface, and multichannel interface:

(1) Hardware interaction refers to the user's active operation behavior through the physical controller, such as using the knob to adjust the volume, using the door handle to open the door, and so on. It includes traditional controllers and parts of product entities that can be operated by users. This kind of interactive interface is more accurate in operation, and the button layout position is relatively fixed, which makes it easier to form habitual operation actions. However, the feedback is relatively low, and the feedback after the operation is not clear enough.

(2) Software interaction interface refers to the behavior of presenting information through the electronic screen, and users interpret information or operate, such as setting the navigation destination through the on-board central control screen, and users observe the sudden speed information in the instrument panel, etc. There are touch operation interfaces and information only interfaces in the software interaction interface. Generally, high-resolution touch display devices are used to carry information and realize operation. This kind of interactive interface is more diverse in form, and it is not restricted by the shape and size compared with the hardware interactive interface. However, the accuracy of operation is relatively low, and the operation feedback is good. It is more in line with the user habits under the rapid development of mobile interaction.

(3) Multichannel interaction refers to the process of interacting with the car through multiple sensory organs in the car environment, such as voice control, mobile phone interconnection, somatosensory control, line of sight control, and so on. It is not only through the button generator screen to interact with the car. In the car environment, when the user's hands and eyes are in use, they can use multichannel senses to interact with the car, make full use of human multisensory channels, and enrich the possibility of human car interaction at a deeper level.

2.3. Classification of Color like Perception. Human perception of color vision has individual differences. Generally speaking, human color vision can be divided into normal color vision and abnormal color vision. The color perception system is mainly responsible for the perception of color by the special photoreceptor on the retina, namely the cone cells. Different types of cone cells have different absorption of visible light. The superposition value of the response results of cone cells stimulated by various bands of light in the visible spectrum determines the color form of the spectrum in human eyes. In the human color vision system, there are mainly three kinds of cone cells, namely L-cone cells, M-cone cells, and S-cone cells [9]. For the normal color vision system, L-cone cells are most sensitive to the long wavelength of 535 nm–575 nm in the visible spectrum, M-cone cells are most sensitive to the medium wavelength of 500–550 nm in the visible spectrum, and S-cone cells are most sensitive to the short wavelength of 400–450 nm in the visible spectrum. When these three cones are absent or their sensitivity to their respective wavelength bands is shifted, the perception of color by the color vision system will change, resulting in abnormal color vision. According to the deviation degree of L, M, and S-cone cells in the abnormal color vision system after stimulation of different wavelength components in the spectrum or due to their own lacking, abnormal color vision can be divided into the following categories:

(1) Monochromatic color vision: It is also called total color blindness. This color vision system lacks all visual cones or only one type of visual cones. According to the three primary color theory, its perception of different colors cannot be superimposed like normal color vision. It can only perceive the changes in the brightness of natural light and can only distinguish the brightness changes from

black to gray and then to white but cannot or can hardly perceive other forms of color.

(2) Dichromatic color vision: It is also a typr of color blindness. This color vision system lacks a type of cone cells, so it cannot effectively perceive the spectral information of the corresponding band, and there is a perceptual deviation. According to the type of visual cone cells missing, those missing L-cone cells are generally called red blindness, which is mainly manifested in the inability to normally perceive the red band spectrum. The absence of *M*-cone cells is called green blindness, which is mainly manifested in the inability to normally perceive the green band spectrum. Those who lack *S*-cone cells are called to suffer from blue blindness, and their main performance is that they cannot normally perceive the blue purple spectrum. In abnormal color vision groups, the first two kinds of deficiencies are in the majority, which are often collectively referred to as red and green color blindness; The number of blue blindness caused by the third condition is very small.

(3) Trichromatic color vision: It is color weakness. In this type, cone cells is almost the same as the normal color vision system, but one type of cone cells has some defects or abnormalities, and the perception of the spectrum of the corresponding band has changed, resulting in confusion and deviation in color resolution, and dichromatic color vision types can be subdivided into red weak, green weak, and blue weak according to their defective cone cell types, and they can be classified according to their severity. From the perspective of color perception, when the color weakness is slight to the extreme, its condition is close to normal color perception. When it reaches the most serious level, its condition is close to color blindness.

3. Structure Design of Generative Antagonism Neural Network Based on Visual Recognition

In order to solve the problem of color perception in the vehicle HMI interactive interface, this paper proposes an image enhancement model based on conditional generation countermeasure network, as shown in Figure 3. It is mainly composed of two parts, generation network and discrimination network. The former is used to learn the data distribution of clear images and obtain the nonlinear transformation relationship between the vehicle color image and this distribution. Then generate a pseudo clear image that removes the influence of other factors.

The generated network structure is shown in Figure 4, where k_i is $i \times i$ core size, s_j is the moving stride with length j, D is the expansion rate in cavity convolution, Conv is convolution operation, and Conc is a splicing operation. The latter discriminates the authenticity of the input image, the clear reference image label is true, and the pseudo clear image label obtained from the generated network is false. In addition, the discriminant network adopts the double discriminant form, which can supervise the global style and detail edges, respectively. In the loss function, this paper designs an unsupervised form that only considers the quality of the image itself, which together with confrontation and content loss constrains the training direction of the model and helps to alleviate the dependence on data. The network design will be described in detail in Figure 4.

3.1. Generate Network Structure Design. The network input is $256 \times 256 \times 3$ pixels of image data through 32 channels 1×1 convolution operation, and the ReLU activation function gets the characteristic distribution of $256 \times 256 \times 32$. First, in order to avoid the gradient sparsity problem, this model uses basic convolution operation to replace the pooling layer in the original se res module, so as to prevent information loss and help improve the stability of the network. At the same time, in order to realize the fusion of multiscale feature information, this paper, based on the original se res module and different sizes of convolution kernels are selected according to the convolution position. In addition, the kernel size of convolution operation will affect the parameters and calculation of the network. Therefore, in this paper, 3×3 hole convolution is used to replace 7×7 traditional convolution in the dark green module, which not only ensures the receptive field but also reduces the complexity of the network [30]. D is set to 2. After four se res module operations, convolution processing with the same size is used in the symmetrical position. At the same time, in order to improve the utilization of information and prevent feature loss, this paper adds multiple direct channels and uses the splicing operation to ensure the complete transmission of low-level features. After splicing, the feature distribution is 64 channels, which are used as the input of the last three se res modules. In addition, before the splicing operation, add three-layer networks, including yellow, gray, and blue ones which are convolution operation, nonlinear activation, and normalization, respectively. The convolution selection of the output layer is 4×4. The activation function is set as tanh, and the tanh activation output can play the role of both activation and normalization. It normalizes the calculation results between [−1, 1], which can avoid too large or too small values. During training, all image data are normalized to [−1, 1] and then inversely adjusted to [0, 255] when the generated results are displayed, so the value of Tanh activation function is adjusted as the pixel value of the three channels of the output image, Thus, the final enhancement result can be obtained.

3.2. Discriminant Network Structure Design. In the process of network training, the ability of generating network and distinguishing network cannot always maintain a balance, resulting in the inability of generating network to distinguish network to compete, and limiting the network depth of distinguishing network limits the ability of distinguishing network to a certain extent. The discriminant network structure is shown in Figure 5, which is composed of five

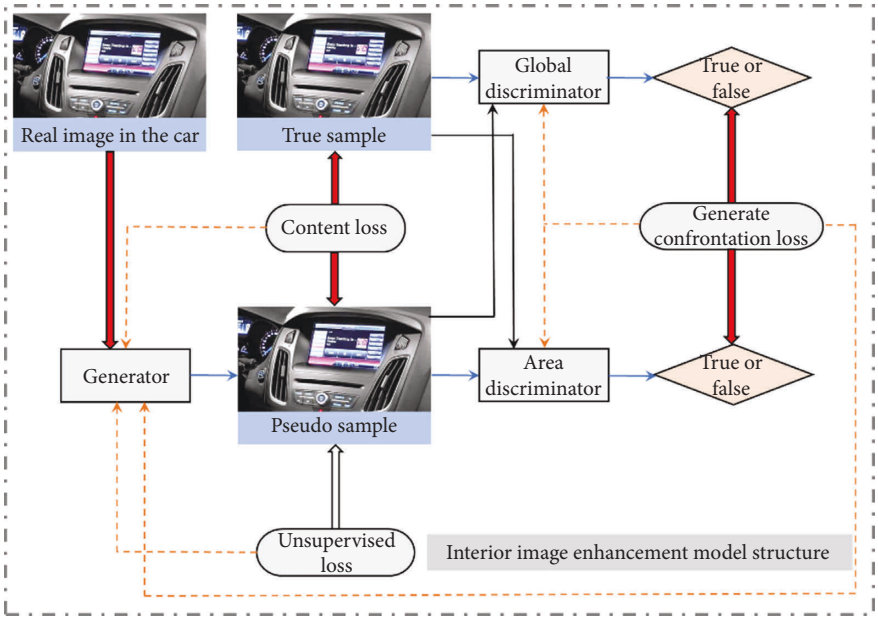


Figure 3: Interior image enhancement model structure.

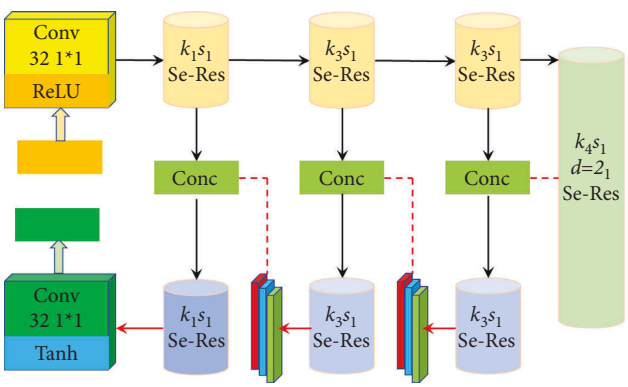


Figure 4: Design of generation countermeasure network structure.

convolution layers without using full connection and pooling layers [31, 32].

According to Figure 5, since the batch normal layer may bring artifacts to the generated pictures during network training, and the network training performance cannot be stabilized, it is judged that the batch normal layer is not used in the network, and it is replaced by the instance normal layer. Using the instance normal layer helps to enhance the network performance and reduce the computational complexity. In addition, the instance normal layer can subtract the depth information in the network input direction from the mean divided by the standard deviation, accelerate the network training speed, and increase the nonlinear fitting ability. Therefore, it is replaced by the instance normal layer.

Optimize the activation function of the last layer. The activation function in CNN network generally adopts the Sigmoid function or ReLU function. When the network training ability is strong, it can 100% distinguish the false image of the generated network. RelU, which is often used as the activation function in deep learning, can speed up the calculation of convolution. However, when a very large gradient flows through ReLU neurons, it may disturb the distribution of data, resulting in some neurons not activating any data and turning off. For this situation, use the leaky ReLU function, which can avoid this situation. The definition of the leaky ReLU function is as follows:

$$y = \max(0, x) + leaky * \min(0, x), \tag{4}$$

where leaky is the coefficient of (0, 1).

To sum up, the overall framework of realizing the color perception of the on-board HMI interactive interface through the condition generation countermeasure network is shown in Figure 6.

4. Color Perception Experiment and Result Analysis of Vehicle HMI Interface Based on Generative Antagonism Neural Network

4.1. Experimental Preparation. In order to get the user's favorite car interface color design scheme in a more objective way, the color perception experiment of car instrument and central control interface based on eye movement physiological data is designed. The experiment needs to be carried out under quiet conditions, and the electronic equipment irrelevant to the experiment in the same space should be closed in the laboratory. The experiment should be carried out in an environment with good light, smooth ventilation, and comfortable room temperature, and the subjects should be in a happy state to avoid the impact of measurement on the experimental results due to tension, anxiety, or depression.

This experiment was improved and arranged on the basis of existing research. There were 20 subjects in the test, of which the male to female ratio was 1 : 1. The age range of the subjects is 20–27 years old. Among them, the ratio of

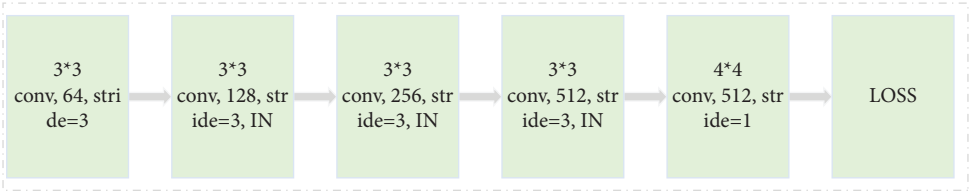


Figure 5: Discrimination network structure design diagram.

Figure 6: Overall frame diagram of color perception.

researchers engaged in automobile styling design to non-professionals is 1 : 1. This test requires that the test personnel participating in the test shall not engage in any violent activities, drink alcohol, keep regular work and rest, and shall not smoke, drink strong tea, drink coffee, or take drugs related to the nervous system before the test.

A total of 120 images of instruments and central control of 60 different models are classified, and these images are optimized for image processing such as de reflection and size correction. By consulting official pictures and video materials, we can grasp the color design style of each picture and carry out appropriate color correction, so as to ensure the true color of the picture as much as possible and reduce the color difference. Each sample picture includes six models, so 20 samples of instruments and central control are produced, including 10 samples of instruments and 10 samples of central control. The background of the experimental sample is pure black, and the size of each sample is 973 mm × 598 mm. Figure 7 is an example of an experimental sample picture.

4.2. Experimental Scheme. In the preparation stage of the experiment, it is necessary to check and debug the experimental equipment, confirm that the storage space of the power and data of the equipment is enough to meet the needs of this experiment, and prepare the corresponding emergency plan. The detailed experimental steps are as follows:

Step 1: Adjust the posture of the subjects to ensure that the sitting posture of all subjects is correct

Step 2: The subjects are required to sit at the designated position and adopt a normal and comfortable sitting posture, with their backs on the backrest and their eyes parallel to the center of the screen

Step 3: Present the pictures prepared in advance to the subjects, and let the subjects record the pictures they have selected and describe the colors that have been washed and matched;

Step 4: Use all the pictures selected in the test to generate the countermeasure network for learning analysis and extract the features of the analyzed pictures

Step 5: Count the most popular group matching colors

Step 6: Summarize the experimental records.

4.3. Evaluation Index Design. The experimental evaluation of the two proposed methods of color vision test map synthesis is mainly carried out from two perspectives. First, evaluate the quality of a single synthetic image. Using the two color vision test chart synthesis methods proposed in this paper, to generate a group of images, randomly select one from the images for subjective and objective evaluation and illustrate the effectiveness of the method proposed in this paper and its advantages compared with the traditional test chart picture book through numerical comparison with the classic images in the color vision test chart picture book in terms of compliance and specificity.

4.4. Analysis of Experimental Results. Figures 8 and 9, respectively, show the compliance results of the on-board HMI interactive interface color perception model based on the antineural network model generated in this paper and the traditional color perception model for the above color perception test chart. Figure 8 shows the compliance results of the model constructed in this paper, and Figure 9 shows the compliance results of the traditional color perception model. From the analysis results of the compliance of the

Figure 7: Legend of the experimental sample.

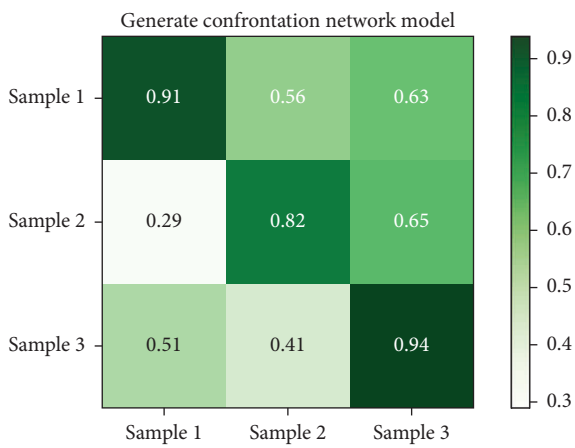


Figure 8: Generate compliance results for the confrontation model.

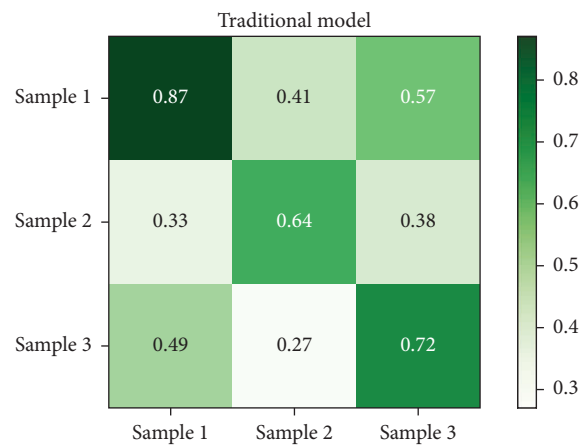


Figure 9: Legend of the experimental sample.

models in Figures 8 and 9 for color perception, it can be seen that the generation antagonism neural network model constructed in this paper has better compliance for color, and its maximum compliance is close to 0.94, while the traditional model is only 0.87.

Figures 10 and 11, respectively, show the specificity results of the vehicle HMI interactive interface color perception model based on the antineural network model generated in this paper and the traditional color perception model for the above color perception test chart, of which Figure 10 is the specificity results of the model constructed in this paper and Figure 11 is the specificity results of the traditional color perception model. It can be seen from the specificity analysis results of the models in Figures 10 and 11 that the specificity of the generation antagonism neural network model constructed in this paper for color is basically equivalent to that of the traditional model, and the maximum specificity of both models is above 0.9.

Based on the above compliance and specificity results, it is found that, first of all, from the compliance of Figures 8 and 9, the compliance of the two different models for color matching has a large deviation, in which the opportunity

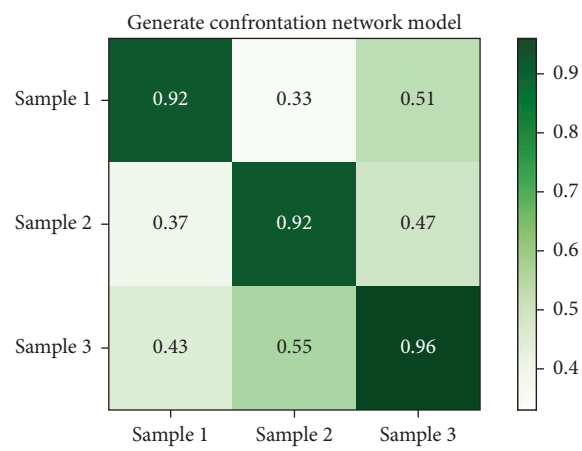


Figure 10: Generate specific results of the confrontation model.

generation confrontation neural network model constructed in this paper is significantly better than the traditional model. In terms of specificity, it can be seen from Figures 10 and 11 that the specificity of the two models is similar. It can be seen that the generative antagonism neural network

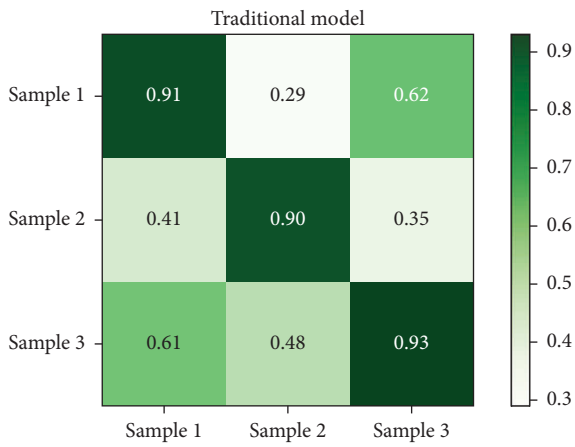


Figure 11: Specific results of traditional models.

model based on this paper not only retains the color perception ability of the original model but also can match the colors well, which verifies the applicability of the model constructed in this paper.

5. Conclusion

There are many colors on the vehicle mounted HMI interactive interface. However, among the many colors on the vehicle mounted interactive interface, many colors do not conform to the human body design, or the user experience is insufficient, which may easily lead to the reduction of driving pleasure and even increase the probability of causing traffic accidents. In view of the above problems, this paper analyzes the color of the intelligent vehicle HMI interactive interface under the user experience through the neural network system of visual recognition (condition generation confrontation network model), analyzes the correlation of various elements of human, vehicle, environment, and interface according to the analysis of the psychological cognition and behavioral operation of vehicle users, and discusses how the vehicle interactive interface can better meet the expected physiological and psychological needs of users and improve the operability, so as to design an on-board HMI interactive interface that can be intelligently perceived according to factors such as weather and driver's interest so as to improve the current backward status of the operation mode of the on-board interactive interface. On the basis of elaborating the interactive design and user experience, by analyzing the design principles of the on-board interactive interface, the neural network of visual recognition is used to explore on how to combine the characteristics of users. The design method of vehicle interaction interface meets the needs of users so as to improve the user's sense of experience.

Data Availability

The dataset used in this paper are available from the corresponding author upon request.

Conflicts of Interest

The author declares that there are no conflicts of interest regarding this work.

References

[1] G. Marin, A. Haimovici, F. l Serban, M. Pavel, N. Gheorghita, and R Vataman, "Biochemical changes in periodontal structures in juvenile diabetes," *Revista de Chirurgie Oncologie Radiologie O R L Oftalmologie Stomatologie Seria Stomatologie*, vol. 25, no. 3, pp. 181–184, 1978.

[2] G. Buchsbaum, "A spatial processor model for object color perception," *Journal of the Franklin Institute*, vol. 310, no. 1, pp. 337–350, 1980.

[3] E Ozgen and R L Ian, "Acquisition of categorical color perception: a perceptual learning approach to the linguistic relativity hypothesis," *Journal of Experimental Psychology General*, vol. 131, 2002.

[4] F. A. Stuart and C. A. Marisa, "Exogenous attention and color perception: performance," *Vision Research*, vol. 46, 2006.

[5] L. Arend, *Surface Colors, Illumination, and Surface Geometry: Intrinsic-Image Models of Human Color Perception*, 1994.

[6] N. Schwabe, A. Robertson, and A. Godwin, "P220Evaluating the influence of method, seat positioning, and anthropometric scaling on mobile machinery visibility profiles obtained using computer simulation," *Occupational and Environmental Medicine*, vol. 73, no. Suppl 1, pp. A194.3–A195, 2016.

[7] N. Kwallek, C. M. Lewis, and A. S. Robbins, "Effects of office interior color on workers' mood and productivity," *Perceptual & Motor Skills*, vol. 66, no. 1, pp. 123–128, 1988.

[8] C. K. Yang and L. K. Peng, "Automatic mood-transferring between color images," *IEEE Computer Graphics and Applications*, vol. 28, no. 2, pp. 52–61, 2008.

[9] K. W. Schaie and K. A. Warner, "A Q-sort study of color-mood association," *Journal of Projective Techniques*, vol. 25, no. 3, pp. 341–346, 1961.

[10] Y. SAITO and H Tada, "Effects of color images on psychosomatic state —effectiveness of color images as mood stimulants as assessed by using an image selection system," *Japanese Journal of Complementary and Alternative Medicine*, vol. 5, no. 3, pp. 225–232, 2008.

[11] N. S. Burner, "El Greco's application of color and mood," *Dissertations & Theses - Gradworks*, 2015.

[12] M. L. Huang, Y. C. Zhou, and C. M. Wang, "An improved color mood blending between images via fuzzy relationship," in *Proceedings of the International Conference on Computer Vision/computer Graphics Collaboration Techniques*, Springer-Verlag, France, March 2007.

[13] S. Kekkonen–Moneta and G. B. Moneta, N. Lewis and A. S Robbins, Effects of office interior color on workers' mood and productivity: Kwallek," *British Journal of Educational Technology*, vol. 33, no. 4, pp. 423–433, 2002.

[14] J. Shen, H. Sun, X. Mao, Y Guo, and X Jin, "Color-mood-aware clothing re-texturing," in *Proceedings of the International Conference on Computer-Aided Design & Computer Graphics*, IEEE Computer Society, Jinan, China, September 2011.

[15] N. Ikoma, "Color mood grasping in video by state estimation over color space with particle filter," in *Proceedings of the International Conference on Control*, IEEE, Chiang Mai, Thailand, November 2017.

[16] N. Kwallek, H. Woodson, C. M. Lewis, and C. Sales, "Impact of three interior color schemes on worker mood and performance relative to individual environmental sensitivity," *Color Research & Application*, vol. 22, no. 2, pp. 121–132, 1997.

[17] G. Solas, L. J. Valdivia, J. Anorga et al., "Virtual laboratory for on-board etcs equipment," in *Proceedings of the IEEE International Conference on Intelligent Transportation Systems*, IEEE, Gran Canaria, Spain, September 2015.

[18] Y. Wang, Y. Zhao, and Y. Yue, "Communication protocol conversion on electrical equipment intelligent monitoring system based on iec61850 standard," in *Proceedings of the International conference on future communication and computer technology*.

[19] S. Jin, J. Yin, M. Tian, S. Feng, S. G. Thompson, and Z. Li, "Practical speed measurement for an intelligent vehicle based on double radon transform in urban traffic scenarios," *Measurement Science and Technology*, vol. 32, no. 2, Article ID 025114, 2021.

[20] Z. Li, "Research on HMI user interface design of networked cars," *China Computer & Communication*, 2018.

[21] Y. Forster, F. Naujoks, and A. Neukum, "Your turn or my turn?: design of a human-machine interface for conditional automation," in *Proceedings of the International Conference on Automotive User Interfaces & Interactive Vehicular Applications*, ACM, 2016.

[22] G. Deep, M. C. Mahatme, and V. Megha, *HMI for Interactive 3D Images with Integration of Industrial Process Control*.

[23] F. Glinka, A. Raed, S. Gorlatch, and A. Ploss, "A service-oriented interface for highly interactive distributed applications," in *Proceedings of the Euro-Par 2009 - Parallel Processing Workshops, HPPC, HeteroPar, PROPER, ROIA, UNICORE, VHPC, Delft*, Springer-Verlag, The Netherlands, Auguest 2009.

[24] G. Wei, H. Tan, and J. H. Zhao, "User Participatory Design of Automobile Navigation Human-Machine Interface Design Based on Reference Object," *Packaging Engineering*, 2014.

[25] E. D. Cutright, C. Mokkapati, and R. D. Pascoe, *Method for Adjusting Braking Parameters of a Train to Account for Train Characteristic Parameter Variations*, 2015.

[26] C. Zhang and A. Zhu, "Research on application of interaction design in head-up Display on Automobile," in *Proceedings of the ICITEE-2019: 2nd International Conference on Information Technologies and Electrical Engineering*, Zhuzhou Hunan China, December 2019.

[27] Z. Wang, L. Chen, L. Wang, and G. Diao, "Recognition of audio depression based on convolutional neural network and generative antagonism network model," *IEEE Access*, vol. 8, no. 99, p. 1, 2020.

[28] L. Deutsch, "Generating Neural Networks with Neural Networks," 2018, https://arxiv.org/abs/1801.01952.

[29] B. Fan, L. Yuan, and X. U. Youxiang, "Image restoration method based on residual mechanism generative adversarial networks," *Video Engineering*, 2019.

[30] J. Liu, "Modeling of confrontation network image registration based on fuzzy mathematics," *Science Technology and Engineering*, 2019.

[31] W. Gao, Y. F. Guo, and X. U. De-Heng, "Application research of generation confrontation network in the field of computer vision," *Value Engineering*.

[32] A. Menon, F. Herwig, P. A. Denissenkov et al., "Reproducing the observed abundances in RCB and HdC stars with post-double degenerate merger models - constraints on merger and post-merger simulations and physics processes," *Physics*, vol. 772, no. 1, pp. 57–61, 2013.

Hindawi
Computational Intelligence and Neuroscience
Volume 2022, Article ID 3688797, 13 pages
https://doi.org/10.1155/2022/3688797

Research Article

Interactive Display of Images in Digital Exhibition Halls under Artificial Intelligence and Mixed Reality Technology

Xu Liu[1] **and Nan Zhang**[2]

[1]*Animation Art College, Jilin Animation Institute, Changchun, Jilin 130012, China*
[2]*Department of Movie, Jilin Animation Institute, Changchun, Jilin 130012, China*

Correspondence should be addressed to Xu Liu; 1503038@sust.edu.cn

Received 7 August 2022; Revised 11 September 2022; Accepted 15 September 2022; Published 12 October 2022

Academic Editor: Ning Cao

The attractiveness of traditional exhibition halls to young people is gradually decreasing. Combining modern digital technology to improve the display effect of the exhibition hall can effectively enhance the effect of cultural publicity. This article introduces the technology of image interaction and mixed reality (MR) to improve the historical and cultural propaganda level of the Shaanxi exhibition hall. The advantages of MR technology in applying digital exhibition halls are theoretically expounded. A theoretical plan for Shaanxi history and culture-related display areas is designed using artificial intelligence combined with MR technology. In addition, the survey respondent's evaluation of the effect of the new exhibition hall is obtained using a questionnaire survey. The survey results show that 97% of people like the history and culture of Shaanxi but only 13% of the people say they know or know very well about the history and culture of Shaanxi. In addition, 60% of the tourists say they are satisfied with the cultural experience of Shaanxi, and only 27% of the tourists are very satisfied. Also, 96% of tourists are willing to experience Shaanxi's history and culture through digital exhibition halls, and 93% are willing to participate in cultural experience activities based on MR technology. The survey results prove that tourists are satisfied with the effect of the new exhibition hall. Tourists want to add a distinctive form of cultural experience to the exhibition hall. They are willing to accept digital exhibition halls incorporating MR technology and are very happy to participate in the exhibition method of image interaction. This shows that the use of image interactive display based on MR technology in the layout of the exhibition hall is recognized by people and has strong feasibility. This article has reference significance for the digital upgrade of the exhibition hall and the development of the cultural tourism industry.

1. Introduction

In the display method of exhibits, the traditional exhibition hall mainly displays through entities, with the text explanation and picture appreciation. This kind of display is dull and dreary, and its appeal to contemporary young tourists is gradually diminishing [1]. With the development of the times, the role of various exhibition halls is not limited to a single collection, research, and exhibition but highlights the role of education and cultural dissemination [2]. The digital exhibition hall display method that combines the traditional physical exhibition hall display method and mixed reality (MR) technology has become the development trend of the current exhibition hall with the integration and development of modern science and technology and display design; it has been used in various fields such as medical treatment, construction, education, and entertainment. Artificial intelligence (AI) is a branch of computer science. It is an interdisciplinary field of information that encompasses multiple disciplines. AI can develop theories, methods, techniques, and application systems for simulating, extending, and expanding human intelligence. It seeks to understand the nature of intelligence and produce a new type of intelligent machine that responds similarly to human intelligence. Since the birth of AI, theory and technology have become increasingly mature, and the application field has expanded. AI can simulate the information process of human consciousness and thinking.

International Business Machines Corporation cooperated with the Forbidden City as early as more than ten years ago. Tourists can deeply interactively learn about the Forbidden City using computer technology and an interactive experience system to create a virtual Forbidden City [3]. The Dunhuang Museum is planning to build an experiential tourist service center. Visitors can experience the Mogao Grottoes and feel the background culture of the Mogao Grottoes using advanced display forms [4]. The Louvre, the Metropolitan Museum, and the Palace of Versailles put their collections and paintings on the website. Tourists can visit these museums without leaving home using three-dimensional (3D) virtual technology [5]. Lin et al. researched and summarized MR image reconstruction based on deep learning (DL) in medical imaging [6]. They argued that DL-based algorithms could rival traditional reconstruction methods regarding image quality and computational efficiency. Zaharchuk and Davidzon found that AI could achieve rapid image acquisition and reconstruction and address the problems of AI-driven approaches in performing MR-based attenuation correction [7].

Currently, the number of digital exhibition halls using this technology is still small. This has reduced the development of the cultural tourism industry and the spread of various regional cultures to a certain extent. The Shaanxi exhibition hall still uses the traditional mode of operation and publicity. This article studies the interactive display method of images in the digital exhibition hall to improve the historical and cultural propaganda ability of the Shaanxi exhibition hall. A display area related to Shaanxi history and culture is designed using AI combined with MR technology. First, video interaction and MR technology are introduced in detail. Second, the necessity of applying MR technology in digital exhibition halls is expounded by revealing the advantages of digital exhibition halls compared to traditional exhibition halls. Third, the theoretical scheme of the interactive display of images in the digital exhibition hall is systematically planned. Finally, the feasibility of the interactive display method of digital exhibition hall images based on AI and MR technology is proved through a questionnaire survey. This article has reference value for the digital upgrade of relevant exhibition halls and the development of the cultural tourism industry.

2. Interactive Display of Images in Digital Exhibition Halls Based on AI and MR Technology

2.1. Image Interaction. "Interaction" means to communicate and interact. With the continuous development of current technology, its definition combines theories and methods of various disciplines such as immersion theory, computer application, and technological innovation and has become a common part of current life [8]. Image interaction is to adopt images as a presentation method. Various sensors can read the viewer's reaction, and real-time rendering technology is used to interact with the viewer [9].

Image interaction integrates media based on traditional images and has the features of traditional media and interaction [10]. The features of image interaction are shown in Figure 1.

From Figure 1, the features of image interaction are divided into five points, namely, immersion, feedback, diversity, virtuality, and nonlinearity [11, 12]. The specific performance of each feature is shown in Table 1.

2.2. MR Technology. VR technology uses a computer to simulate a virtual environment, so the experiencer is immersed in a 3D virtual space. The experiencer seems to be in the real world by mobilizing the experiencer's vision, hearing, and other senses [18]. The biggest feature of VR is that it gives the experiencer a strong sense of "immersion" and "presence." Augmented reality (AR) is an extension of VR. It uses technologies such as computer vision to closely integrate virtual objects with the real world. The goal of AR is to enhance the experiencer's understanding of the real environment by combining virtual objects with the real environment [19]. MR is the combination of AR and VR, and the technology connects the user, the real world, and the virtual world in a way that enhances the user experience. The biggest feature of MR technology is that its processing and positioning of space are fine and accurate. Its calculations are performed in real-time. It combines virtual imagery with real space. It performs precise positioning according to its processing of images, thereby presenting a visual environment with both real and virtual images [20]. Table 2 reveals the features of MR technology.

2.3. The Necessity of Applying MR Technology in Digital Exhibition Halls. The traditional exhibition hall uses the display and placement transmission method to disseminate information through the physical exhibition, text explanation, picture appreciation, and model visit. The main visual design forms are exhibition boards, video broadcasts, and display props, which cannot break through the limitations of environmental space. The display method is rigid and boring, only staying in graphic design and lacking flexible design methods. The rich exhibition content cannot be presented [24]. In traditional exhibition halls, the uneven visual language and diverse information prevent visitors from making effective information choices. Too much emphasis is placed on the dominant position of the exhibits while ignoring the visitor's experience and feelings about the environment. In addition, the path that visitors choose to receive information is only static information presented by exhibits, pictures, and text. They do not form visual communication feedback with the display space. This single sensory communication interaction is mostly output in a one-dimensional pattern. There is no interaction between people and information, and visitors can only act as bystanders. Their aesthetic feeling is limited to static viewing and lacks a sense of communication experience [25].

According to the above shortcomings of traditional exhibition halls, the advantages of applying MR technology to digital exhibition halls are compared with them. This reflects the necessity of applying MR technology to digital exhibition halls. The specific comparison is demonstrated in Table 3.

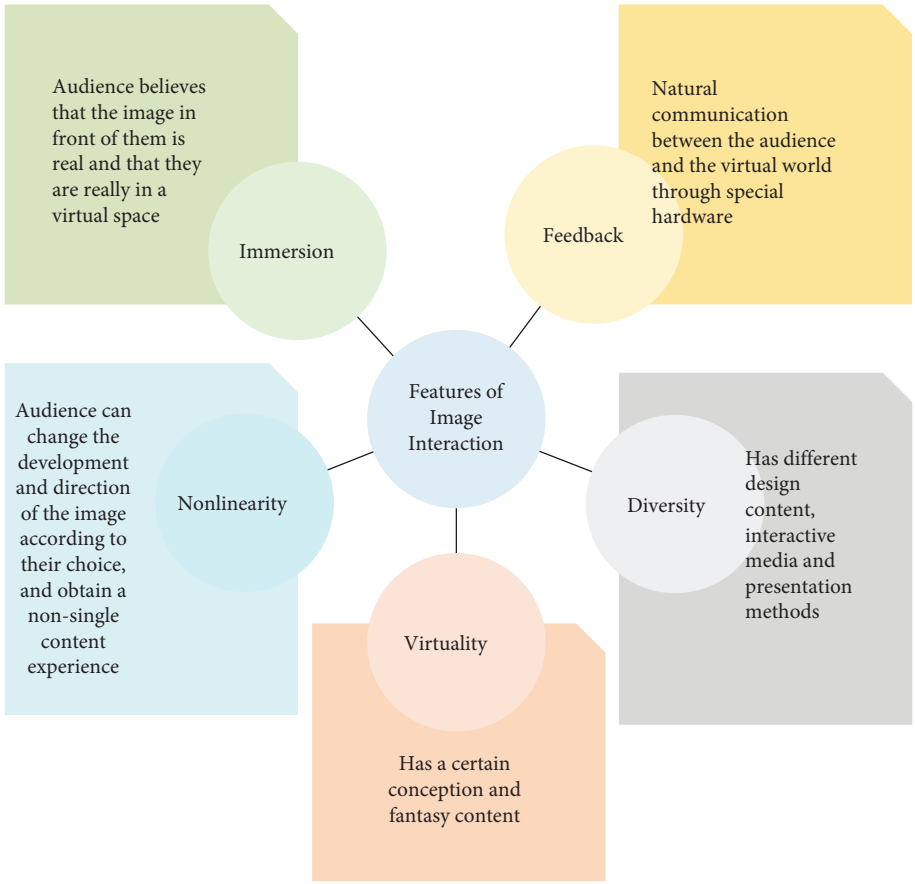


Figure 1: Features of image interaction.

Table 1: Feature performance of image interaction.

Feature	Specific performance
Immersion	Image interaction can bring the audience an immersive experience, and people enjoy sensory stimulation in the space expressed by images. This breaks through the traditional single visual sensory experience. Image interaction achieves full integration and immersion with the support of projection technology, sensing technology, image acquisition, sound technology, and other technologies and enters people's psychological space to gain a strong sense of participation and satisfaction [13].
Feedback	The participants get visual, auditory, and informational feedback for the first time to communicate and interact when they contact the works. In the virtual reality (VR) experience, people participate in the virtual environment through body movements, gesture changes, eye gaze, and voice control. Besides, interactions between scenes generate. People feel the grip, weight, and movement of real objects in the virtual environment, and the virtual images and participants can get interactive feedback in the experience [14].
Diversity	On one hand, the diversity of interactive images is reflected in the fusion of various media. The media produced by the development of science and technology and the interaction between media can form a whole. On the other hand, there are various forms of interaction. Image interaction can produce the interaction between the artist and the work, the audience and the work, the work and the environment, and the work and the work. Image interaction is an active process that presents various interactive modes [15].
Virtuality	Image interaction simulates the real environment through digital technology. The audience wears VR glasses to participate in the VR scene. People see digital virtual space through visual experience. The real action in the virtual environment can change development in the virtual environment and affect people's emotions to achieve the realism of virtual experiences. Virtuality is characterized by a sense of interlacing between virtual and real space-time, which escapes the constraints of time and space and increases the audience's perceptual experience [16].
Nonlinearity	The interactive experience is innovative in the traditional video linear narrative method. The spontaneous dissemination of the audience is based on the strong feeling of the heart. Image interaction is different from the single clue narrative of traditional works of art. It can extend various aspects of story clues and generate thinking through a single point. Everyone will have different spiritual perceptions and artistic results when they participate. The cognition of nontraditional art theory is a nonlinear occurrence to generate the possibility of multiple narratives [17].

Table 2: Features of MR technology.

Feature	Specific performance
Reality	MR can combine virtual and reality to create an immersive feeling. The experiencer feels a new environment that is half real and half virtual to produce a good sensory experience [21].
Conceptuality	In a realistic new environment, users use their imagination to construct an environment that does not exist objectively. MR technology has a very large room for development. MR will continue to enrich the interactive display technology of images in digital exhibition halls with the continuous development and improvement of technologies such as computers and VR [22].
Interactivity	MR technology combines VR with traditional physical displays, building a bridge between virtual objects and the real world. A connection between users and physical objects is established. Users can move freely between reality and VR, change the traditional static display state, and break the one-way nature of information transmission to realize real-time interaction and communication between people and objects [23].

Table 3: Comparison of traditional exhibition halls and digital exhibition halls.

Disadvantages of traditional exhibition halls	Advantages of digital exhibition halls	The necessity of MR technology in the digital exhibition hall
Single display method	The innovative form of popular science	At present, most of the exhibition halls in China still have many problems, such as the same and single display content and form. These problems seriously restrict the sustainable development and construction of Chinese museums. Many exhibition halls have adopted technologies such as multimedia audio, touch screen, video storytelling, interactive games, and hand-held interactive devices. However, the application of VR technology, which is very popular in recent years, is less in exhibition halls. In particular, the application of MR technology is almost nonexistent, so MR technology is applied to the digital exhibition hall. This move not only enriches the connotation of the exhibits but also innovates the display form of the exhibits [26].
Passive information reception	The audience's understanding and mastery of knowledge	MR uses features such as virtual-real integration, situational simulation, special sound effects, and real-time interaction to inject new vitality into the exhibition hall. The transmission of sound can enhance the appeal of the displayed content and reproduce the historical situation. The real objects are presented in front of the audience in all directions, and the interaction with the audience can be realized. The audience can zoom in and rotate the images of the exhibits according to their needs to understand the history of the exhibits or their connotations. This form is no longer limited to a single and distant display. The audience is subtly influenced by this form of presentation. The audience deepens the knowledge and understanding of the exhibits [26].
Insufficient communication and interaction	Audience's high interest	The MR display form uses new science and technology to let the audience enter the virtual situation based on the original cabinet display. The audience interacts with exhibits in many audio effects. The exhibits present not only the exquisite appearance but also the accumulation of cultural knowledge. The new display form of MR is helpful to highlight the details of the exhibits and enrich the content of the exhibits. The realization of MR technology makes up for the shortcomings of traditional display forms in terms of time and space transformation, emotional penetration, and multisensory participation. This can stimulate the audience's interest in visiting and enhance the educational effect of the museum [26].

2.4. Theoretical Scheme of Interactive Display of Images in Digital Exhibition Halls. The theoretical scheme of interactive images display in digital exhibition halls based on AI and MR technology is mainly divided into the material collection stage, 3D modeling stage, and interaction design stage [27], as shown in Figure 2.

The material collection stage includes the understanding of MR technology and the sorting out of the exhibition hall content. The understanding of MR technology includes understanding its definition and principle, the characteristics of MR technology, the development process, and the related hardware equipment and software support for realizing MR. It also delves into the performance methods and required technical conditions of MR technology. Sorting out the contents of the exhibition hall refers to the on-the-spot inspection of the exhibition hall and understanding the exhibition hall through photos, videos, and written records to form a complete system. According to the collected materials, they are classified and

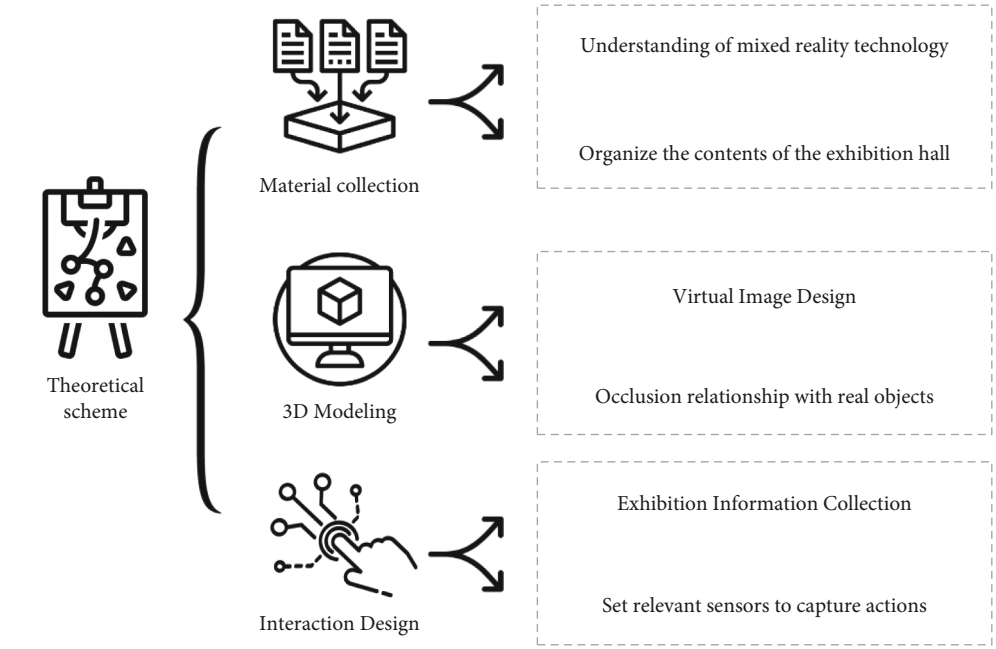


Figure 2: The theoretical scheme of interactive display of images in digital exhibition halls based on AI and MR technology.

integrated and finally expressed on paper through sketches and flowcharts [28].

The 3D modeling stage deals with the occlusion relationship between virtual images and real exhibits in the real space through 3D modeling. This step requires rough depth estimation and precise contour processing by the image interaction system [29]. Therefore, when the overall occlusion environment is constructed, it is mainly divided into the following steps:

Step 1: Model reconstruction of the target exhibit is performed. This step is conducive to accurately fitting the edge of the virtual image and the real object.

Step 2: The position of the virtual model corresponding to the position of the real exhibits is marked and placed. The virtual coordinates and the physical coordinates of the real space are accurately aligned by the processor of MR technology. The position of the virtual image is used to simulate the occlusion objects in the real space to effectively occlude the virtual information.

Step 3: The occlusion material is attached to the reconstructed digital model to make the digital model occluded.

Step 4: The adjustment of the rendering parameters and the positional relationship between the front and rear of the virtual target makes the virtual object approach the occlusion effect of the real object in the display. Virtual imagery is used for precise localization and depth estimation of objects. Besides, scenes are built for MR spaces with occlusion relationships. In the mixed space, the real collections and virtual images have similar real edge occlusion and good interaction relationships.

In the interaction design stage, the interactive process design of the digital exhibition hall image is mainly divided into two parts. In the first part, the characteristics of the exhibition hall and exhibits are collected through tracking and registration technology. The mixed virtual and real environment is effectively constructed and positioned. According to the system guidance, the target exhibits and the spatial relationship between the target exhibits and the entire environmental space are scanned, and data are loaded in advance, which helps the image interaction system input and read the coordinate information of the space. The image interaction system will construct a hybrid virtual space according to the obtained coordinate system. The occlusion relationship of the mixed space is constructed combined with vertex shaders in the mixed environment to improve the construction of the entire mixed environment [30]. Figure 3 displays the overall framework of the system.

In Figure 3, the data acquisition layer mainly collects image, speech, and human body coordinate data to provide basic data for subsequent panorama generation and display. The data processing layer is responsible for virtual image generation, gesture recognition, speech recognition, and target tracking. The user interface layer is responsible for displaying panoramic virtual images to realize human-computer interaction (HCI).

In the second part, the interaction mode of MR images is mainly based on user gestures, user gaze tracking, and visual locking. Gestures and sight lines are captured and acquired through the product's camera. A ray is emitted at the obtained coordinate position to determine the visual line of the user. The virtual information is encoded and set as a collidable material. The time is recorded when the simulated gaze line collides with the virtual information. The user's operation is made timely and effective. The specific operation steps of the instruction recognition program are shown in Figure 4.

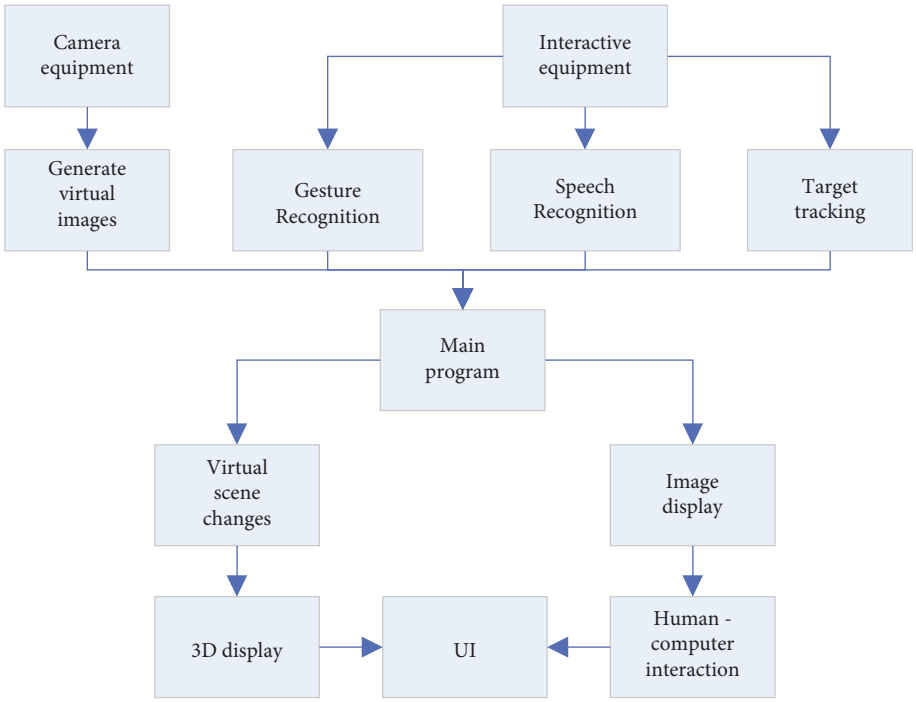


Figure 3: The overall framework of the system.

This theoretical solution for the interactive display of images in digital exhibition halls applies to not only digital exhibition halls but also various museums. It can even be extended to all kinds of scenic spots and historical sites. The application of AI and MR technology in the digital exhibition hall is novel, intuitive, and highly interactive compared to the traditional exhibition hall display method. The scheme enhances the fun of the visit and deepens the understanding of the exhibits.

2.5. Questionnaire Survey. This article investigates the positioning and views of digital exhibition halls in people's minds by issuing questionnaires to ensure that the interactive display of images in digital exhibition halls based on AI and MR technology can be recognized and loved by public visitors. In this survey, the place where the questionnaires are distributed is located near the Shaanxi History Museum, mainly for the visitors who come to the museum. Visitors to the museum are randomly selected and complete a questionnaire after consent. Preinvestigation is conducted on five visitors, and the problems in the questionnaires and the inadequacy of the investigation are corrected. The test is conducted for two days. The specific content of the questionnaire is demonstrated in Table 4.

A total of 130 questionnaires are distributed, and 123 valid questionnaires are recovered, with an effective recovery rate of 94.62%.

The Cronbach's coefficient is used to test the reliability of the questionnaire results to examine the internal consistency of the scale.

$$\alpha = \frac{k\bar{r}}{1 + (k-1)\bar{r}}. \tag{1}$$

In equation (1), k is the number of evaluation indicators, $\bar{r}$ is the mean of the correlation coefficients of the k evaluation indicators, and α is between zero and one. The survey results are transformed according to the corresponding scores. The Cronbach's alpha of each index factor is obtained by inputting the results into the computer for SPSS processing. The scores of each index factor are screened to form the final index factor of the questionnaire. After calculation, the consistency of each Cronbach's alpha is within the acceptable range (0.8), indicating that the questionnaire results are highly reliable.

The content validity of the questionnaire is evaluated by the content validity ratio (CVR).

$$CVR = \frac{n - 0.5N}{0.5N}. \tag{2}$$

In equation (2), n is the number of questionnaires that the researchers believe can better reflect the measured items in the relevant questionnaire scoring results. N is the total number of returned questionnaires. The statistical result by SPSS software is 0.94, indicating that the overall questionnaire has high validity.

3. Survey Results and Analysis

3.1. Survey Results. In this questionnaire survey, the age and education level distribution of the surveyed population is shown in Figure 5.

From Figure 5, the age distribution of the respondents is 5 people under 18 years old, 60 people between 18 and 28 years old, 28 people between 29 and 39 years old, 21 people between 40 and 55 years old, and 9 people over 55 years old. They account for 3%, 50%, 23%, 17%, and 7% of the respondents, respectively. The education level of the surveyed population is 9 people from high school and below, 32 people from high school and technical secondary school, 78

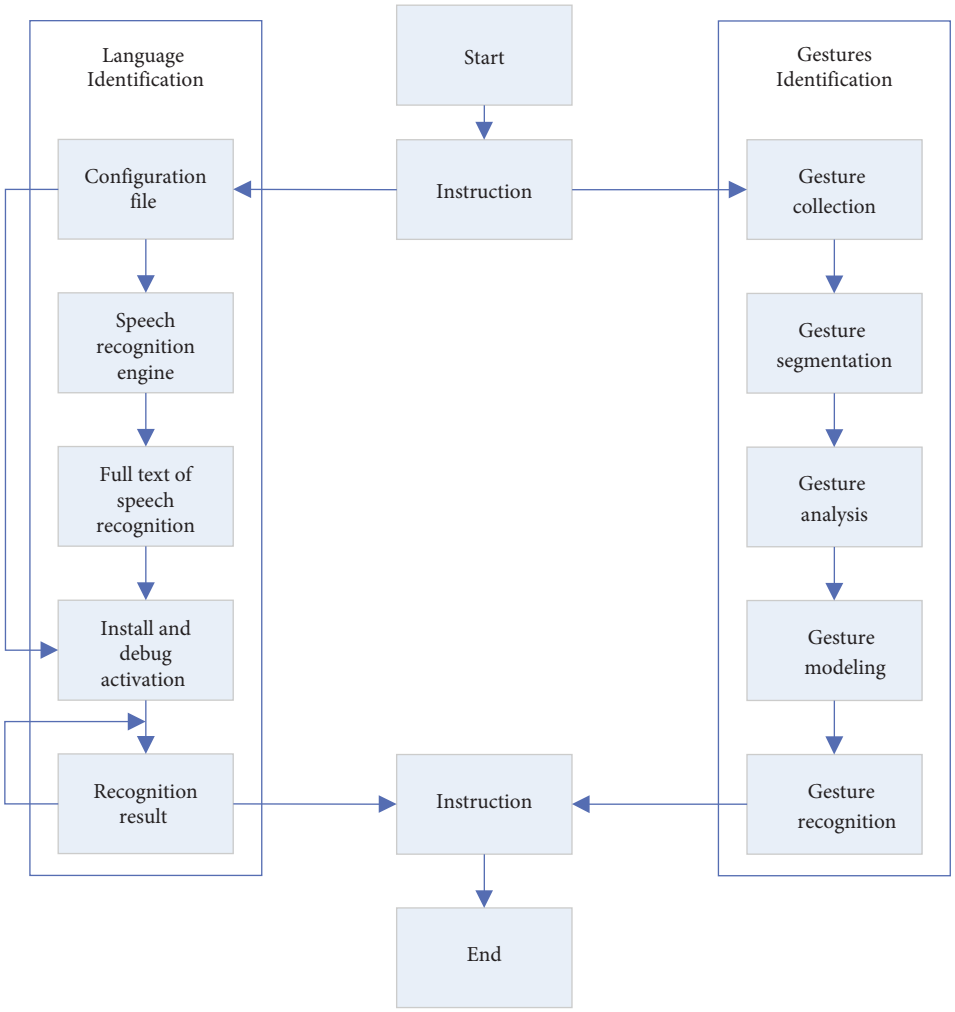


Figure 4: Instruction recognition program.

people from junior college and undergraduate, and 4 people with a master's degree or above. They account for 7%, 27%, 63%, and 3% of the respondents, respectively.

According to the questionnaire, the surveyed people's liking for Shaanxi's history and culture and their liking for cultural experience are shown in Figure 6.

From Figure 6, 97% of people like the history and culture of Shaanxi and only 3% dislike it. Besides, 96% are willing to experience Shaanxi history and culture through the digital exhibition hall, and only 4% are unwilling.

Figure 7 reveals the degree of understanding of the people surveyed on Shaanxi's history and culture and their satisfaction with the local cultural experience.

From Figure 7, regarding the level of understanding of Shaanxi's history and culture, 10% say they know very well, 3% say they relatively understand, 70% say they are general, 13% say they do not understand, and 4% say they do not understand at all. According to the survey of satisfaction with Shaanxi cultural experience, 27% say they are very satisfied, 60% say they are relatively satisfied, 13% say they are general, and the number of dissatisfied people is zero.

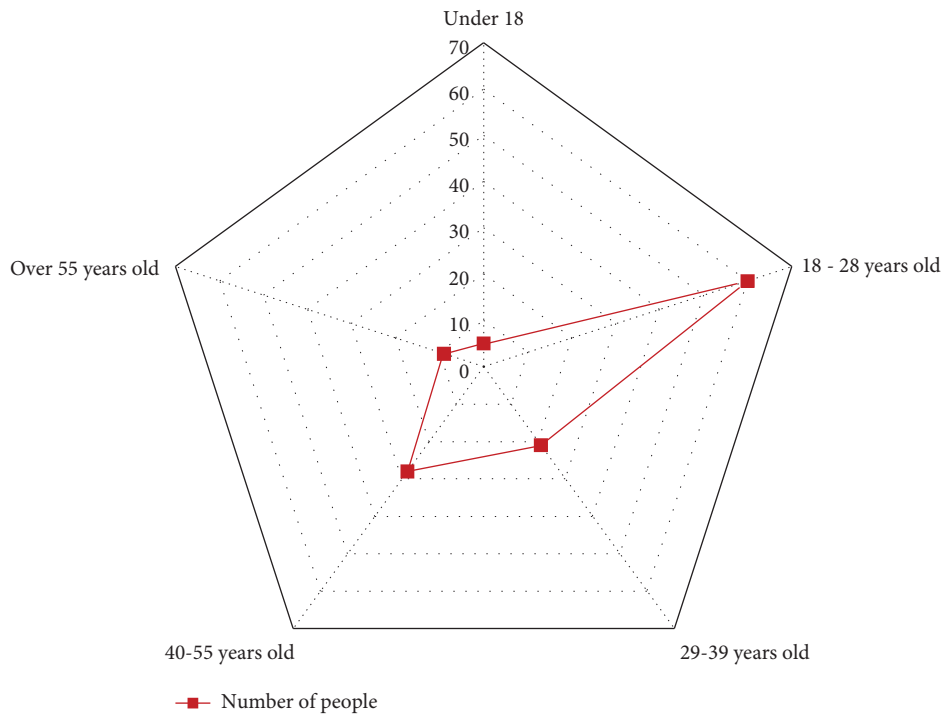
The proportion of respondents' willingness to participate in Shaanxi's historical and cultural experience and their approval of the development direction of cultural tourism is shown in Figure 8.

From Figure 8, 93% say they are willing to participate in cultural experience activities based on MR technology, and 7% say they are not. Among the expectations for the development of cultural tourism, 72% of the respondents express their agreement or strong agreement, and 28% say they are general or disagree.

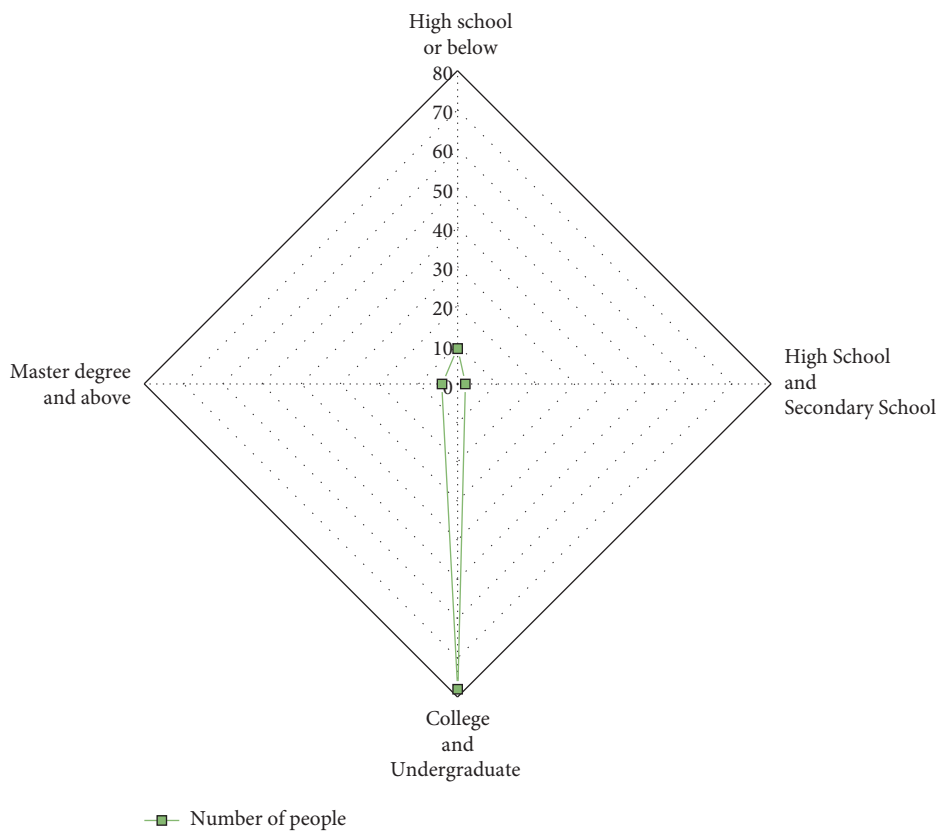
3.2. Analysis of Results. According to the analysis of the above information, the tourists who come to Xi'an are very interested in the history and culture of Shaanxi. They are very willing to participate in the Shaanxi history and culture experience project, but the tourists who have arrived have a general understanding of the history and culture of Shaanxi. There are some obstacles to the spread of culture. Their satisfaction with the Shaanxi cultural experience in Xi'an is not high. Therefore, there is still room for improvement in the cultural experience project. For the development of the cultural tourism industry, tourists hope to have a new and characteristic form of cultural experience. People are willing to accept digital exhibition halls incorporating MR technology and

Table 4: Visitor questionnaire.

Question number	Questions and options				
1	Gender	☐Male	☐Female		
2	Age				
	☐Under 18	☐18–28	☐29–39	☐40–55	☐Over 55
3	Education level				
	☐High school and below	☐High school and technical secondary school	☐Junior college and undergraduate	☐ Master's degree or above	
4	Identity				
	☐Staff of government agencies/institutions			☐Company boss/employee	
	☐Scientific research/education/medical workers		☐Student	☐Soldier	☐Retiree
	☐Others (please specify)				
5	Travel method				
	☐Travel organized by travel agencies		☐Free travel		
6	Do you like Shaanxi culture?		☐Yes	☐No	
7	Do you feel the historical and cultural experience of Shaanxi in Xi'an?			☐Yes	☐No
8	Understanding of Shaanxi history and culture				
	☐Not at all	☐Do not understand	☐Generally		☐Understand
	☐Understand well				
9	Evaluation of Xi'an tourism and cultural experience				
	☐Very satisfied	☐Satisfied	☐Generally	☐Dissatisfied	☐Very dissatisfied
10	Are you willing to participate in the Shaanxi historical and cultural experience?			☐Yes	☐No
11	Favorite form of cultural experience				
	☐Cultural exhibition hall	☐Cultural experience hall	☐Humanistic performance	☐Interactive experience	☐Cultural attractions
12	Reasons that affect your experience of Shaanxi history and culture				
	☐Like natural scenic spots		☐Poor publicity		
	☐Do not pay attention		☐Very concerned about and love the history and culture of Shaanxi		
	☐Cultural experience is traditional				
13	Are you satisfied with the form and content of the current cultural experience?			☐Yes	☐No
14	Would you like to participate in cultural experience activities using MR technology?				
	☐Yes			☐No	
15	Expectations for the development of the cultural tourism industry				
		Agree very much	Agree	Generally	Disagree	Very disagree
Hope there are many new forms					
Desire to develop a unique cultural experience					
Hope the new technology can be applied to cultural tourism					



(a)



(b)

Figure 5: Distribution of age and education level of the surveyed population. (a) Age distribution of the surveyed population; (b) educational level distribution of the surveyed population.

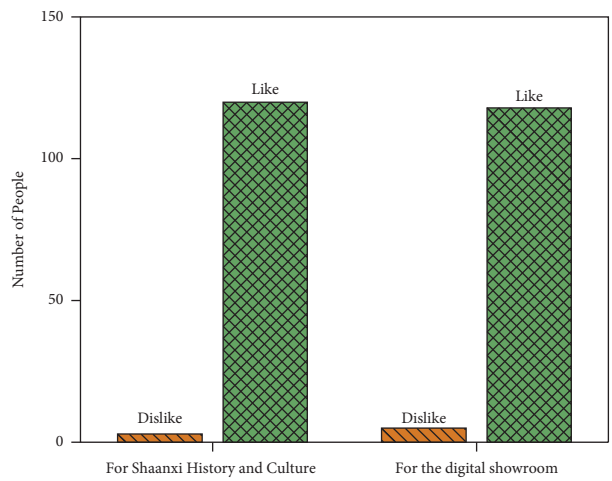


Figure 6: The degree of liking for Shaanxi's history and culture and the degree of liking for cultural experience among the surveyed people.

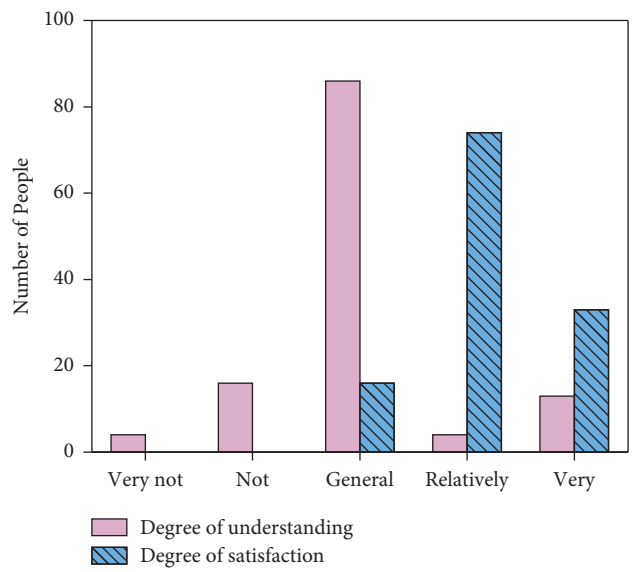


Figure 7: The surveyed population's understanding of Shaanxi's history and culture and their satisfaction with the local cultural experience.

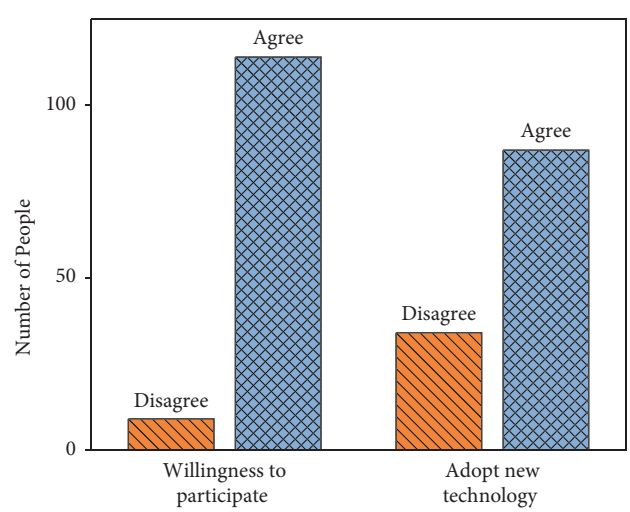


Figure 8: The willingness to participate in Shaanxi's historical and cultural experience and the percentage of respondents agreeing with the development direction of cultural tourism.

are very happy to accept the exhibition method of image interaction. Bec et al. introduced the concept of second chance tourism and the role of innovative conservation methods such as virtual and MR [30]. This could reduce the pressure on inherently fragile tourist destinations, decrease the deterioration of tourist destinations, and create new tourism experiences. Bae et al. proposed a theoretical model based on brand equity theory in the context of MR experiences [31]. They surveyed 251 tourists visiting Seoul, South Korea's cultural and arts tourist attractions. The results showed that the characteristics of MR (interactivity and liveliness) not only affected the emotional aspects of the visitor experience (perceived immersion and enjoyment) but also positively impacted brand awareness, brand association, and brand loyalty. It is feasible to use the image interactive display method based on MR technology in the exhibition hall layout [32–37]. A conscious cultural communication behavior can be generated through the participation of tourists. This can not only facilitate the development of the cultural tourism industry in Xi'an but also play an important role in promoting the history and culture of Shaanxi.

3.3. Impact Analysis of AI and MR Technology. The use of MR technology in the interaction between exhibits and the audience has indeed improved the interactive experience under the premise of satisfying the display of important information and content of Shaanxi's history and culture [38]. In the space built by MR technology, the display of the situation is the most intuitive feeling when the viewer enters the MR display space. The sense of presence brought by the corresponding situation mainly drives the viewer's emotions [39–45]. The shaping and construction of the situation take the viewer's emotions and perspective as the main reference, mobilizing the viewer to generate many associations in the constructed virtual scene. The authenticity of the virtual scene fully improves the overall sense of substitution and the presence of the experiencers to mobilize their subjective initiative.

The HCI experience of the digital exhibition hall based on artificial intelligence enables the interactive operation between the experiencer and the technical equipment to jointly realize the intuitive information presentation of the mixed display effect, which is immersive. The experiencer can actively participate in the display of the digital exhibition hall to increase the enthusiasm and the fun of participation of the experiencer [46–51]. Therefore, these two technologies effectively improve the effect of the interactive display of images in digital exhibition halls in the sense of interaction and education.

4. Conclusion

This article studies the interactive display of images in digital exhibition halls based on AI and MR technology to promote the development of the cultural tourism industry across the country and strengthen the publicity and promotion of local cultures. Video interaction and MR technology are introduced in detail. Then, the necessity of applying MR technology in digital exhibition halls is expounded by revealing the advantages of digital exhibition halls compared to traditional exhibition halls. Furthermore, the theoretical scheme of the interactive display of images in the digital exhibition hall is systematically planned. Finally, a survey is conducted on 123 tourists who come to Shaanxi History Museum in Xi'an. The survey results show that 120 people like Shaanxi's history and culture, accounting for 97% of the total, but only 17 people say they know or know very well about Shaanxi's history and culture, accounting for only 13% of the total. Most tourists say they are satisfied with the cultural experience of Shaanxi. Only 33 tourists are very satisfied, accounting for 27%. Besides, 94% of tourists are willing to experience Shaanxi history and culture through digital exhibition halls, and 93% of tourists are willing to participate in cultural experience activities based on MR technology. The survey results imply that there is still room for improvement in the Shaanxi historical and cultural experience project in Xi'an, and tourists hope to have a new and characteristic form of cultural experience. People are willing to accept digital exhibition halls incorporating MR technology and are very happy to participate in the exhibition method of image interaction. Therefore, it is feasible to use the image interactive display method based on MR technology in the exhibition hall layout. This article only proposes a theoretical scheme for the interactive display of images in digital exhibition halls due to limited capabilities. Subsequently, these theoretical schemes will be applied in practice, and feasible practical schemes will be made. This article also has the following shortcomings and deficiencies. The new exhibition hall proposed here is only a theoretical concept, which needs to be implemented to become a specific new display method. In the future, it is necessary to combine design materials and decoration works to improve the feasibility of this theory. This article has certain reference significance for developing interactive displays of images in digital exhibition halls.

Data Availability

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethical Approval

This article does not contain any studies with human participants or animals performed by any of the authors. Informed consent was obtained from all individual participants included in the article.

Conflicts of Interest

All authors declare that they have no conflicts of interest.

Acknowledgments

The authors acknowledge the help from the university colleagues. This research was supported by following fundings: (1) Scientific Research Project of Education Department of Jilin Province in 2022: Research on the Path of "Rural Revitalization Animation IP" to Reshape the Regional brand of Jilin Province (Grant No. JJKH20221195SK); (2) Ministry of Education Youth Fund Project "Research on Innovation of National Cultural Image Construction in

Overseas Communication of Chinese Animation (2004–2020)."

References

[1] M. A. Khan, S. Israr, A. S Almogren, I. U. Din, A. Almogren, and J. J. P. C. Rodrigues, "Using augmented reality and deep learning to enhance taxila museum experience," *Journal of Real-Time Image Processing*, vol. 18, no. 2, pp. 321–332, 2020.

[2] T. Krzywinska, T. Phillips, A. Parker, and M. J. Scott, "From immersion's bleeding edge to the augmented telegrapher: a method for creating mixed reality games for museum and heritage contexts," *Journal on Computing and Cultural Heritage*, vol. 13, no. 4, pp. 1–20, 2020.

[3] G. Margetis, K. C. Apostolakis, S. Ntoa, G. Papagiannakis, and C. Stephanidis, "X-reality museums: unifying the virtual and real world towards realistic virtual museums," *Applied Sciences*, vol. 11, no. 1, p. 338, 2020.

[4] H. Lee, T. H. Jung, M. tom Dieck, and N. Chung, "Experiencing immersive virtual reality in museums," *Information & Management*, vol. 57, no. 5, Article ID 103229, 2020.

[5] S. Vosinakis, V. Nikolakopoulou, M. Stavrakis, L. Fragkedis, P. Chatzigrigoriou, and P. Koutsabasis, "Co-design of a playful mixed reality installation: an interactive crane in the museum of marble crafts," *Heritage*, vol. 3, no. 4, pp. 1496–1519, 2020.

[6] D. J. Lin, P. M. Johnson, F. Knoll, and Y. W. Lui, "Artificial intelligence for MR image reconstruction: an overview for clinicians," *Journal of Magnetic Resonance Imaging*, vol. 53, no. 4, pp. 1015–1028, 2021.

[7] G. Zaharchuk and G. Davidzon, "Artificial intelligence for optimization and interpretation of PET/CT and PET/MR images," *Seminars in Nuclear Medicine*, vol. 51, no. 2, pp. 134–142, 2021.

[8] H. Elkadi, S. Al-Maiyah, K. Fielder, I. Kenawy, and D. B. Martinson, "The regulations and reality of indoor environmental standards for objects and visitors in museums," *Renewable and Sustainable Energy Reviews*, vol. 152, Article ID 111653, 2021.

[9] A. Sugiura, T. Kitama, M. Toyoura, and X. Mao, "The use of augmented reality technology in medical specimen museum tours," *Anatomical Sciences Education*, vol. 12, no. 5, pp. 561–571, 2019.

[10] C. Aguayo, C. Eames, and T. Cochrane, "A framework for mixed reality free-choice, self-determined learning," *Research in Learning Technology*, vol. 28, no. 0, pp. 1–19, 2020.

[11] F. Ercan, "An examination on the use of immersive reality technologies in the travel and tourism industry," *Business & Management Studies: International Journal*, vol. 8, no. 2, pp. 2348–2383, 2020.

[12] O. Emeafor and S. Onyemechalu, "Objectivity in museums: the nigerian civil war according to the national war museum, umuahia," *The International Journal of the Inclusive Museum*, vol. 14, no. 1, pp. 49–70, 2021.

[13] R. Hammady, M. Ma, C. Strathern, and M. Mohamad, "Design and development of a spatial mixed reality touring guide to the egyptian museum," *Multimedia Tools and Applications*, vol. 79, no. 5-6, pp. 3465–3494, 2020.

[14] S. Bertrand, M. Vassiliadi, P. Zikas, E. Geronikolakis, and G. Papagiannakis, "From readership to usership: communicating heritage digitally through presence, embodiment and aesthetic experience," *Frontiers in Communication*, vol. 6, Article ID 676446, 2021.

[15] P. Z. Suroto, M. H. Dewantara, and A. A. Wiradarmo, "The application of technology in museums," *International Journal of Applied Sciences in Tourism and Events*, vol. 4, no. 2, pp. 170–181, 2020.

[16] M. Shehade and T. Stylianou-Lambert, "Virtual reality in museums: exploring the experiences of museum professionals," *Applied Sciences*, vol. 10, no. 11, p. 4031, 2020.

[17] M. Torres-Ruiz, F. Mata, R. Zagal, G. Guzmán, R. Quintero, and M. Moreno-Ibarra, "A recommender system to generate museum itineraries applying augmented reality and social-sensor mining techniques," *Virtual Reality*, vol. 24, no. 1, pp. 175–189, 2020.

[18] A. Reading, J. Bjork, J. Hanlon, and N. Jakeman, "The labour of place: memory and extended reality (xr) in migration museums," *Memory Studies*, vol. 14, no. 3, pp. 606–621, 2021.

[19] I. Camps-Ortueta, L. Deltell-Escolar, and M. F. Blasco-López, "New technology in museums: ar and VR video games are coming," *Communications Society*, vol. 34, no. 2, pp. 193–210, 2021.

[20] P. J. S. Cardoso, J. M. F. Rodrigues, J. Pereira et al., "Cultural heritage visits supported on visitors' preferences and mobile devices," *Universal Access in the Information Society*, vol. 19, no. 3, pp. 499–513, 2020.

[21] A. Vokshi, E. Shehu, and S. Dervishi, "Military archaeology and contemporary reality in Albania," *City, Territory and Architecture*, vol. 8, no. 1, pp. 14-15, 2021.

[22] J. Condell, N. Mcshane, J. Avlarez, and A. Miller, "Virtual community heritage – an immersive approach to community heritage," *The Journal of Media Innovations*, vol. 7, no. 1, pp. 4–18, 2021.

[23] A. J. Fairchild, S. P. Campion, A. S. García, R. Wolff, T. Fernando, and D. J. Roberts, "A mixed reality telepresence system for collaborative space operation," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 27, no. 4, pp. 814–827, 2017.

[24] J. Chalhoub and S. K. Ayer, "Using mixed reality for electrical construction design communication," *Automation in Construction*, vol. 86, pp. 1–10, 2018.

[25] M. P. Strzys, S. Kapp, M. Thees et al., "Augmenting the thermal flux experiment: a mixed reality approach with the hololens," *The Physics Teacher*, vol. 55, no. 6, pp. 376-377, 2017.

[26] M. White, P. Petridis, F. Liarokapis, and D. Plecinckx, "Multimodal mixed reality interfaces for visualizing digital heritage," *International Journal of Architectural Computing*, vol. 5, no. 2, pp. 321–337, 2007.

[27] J. Birt, Z. Stromberga, M. Cowling, and C. Moro, "Mobile mixed reality for experiential learning and simulation in medical and health sciences education," *Information*, vol. 9, no. 2, p. 31, 2018.

[28] I. V. Osipov, "Cubios transreality puzzle as a mixed reality object," *International Journal of Virtual and Augmented Reality*, vol. 1, no. 2, pp. 1–17, 2017.

[29] L. Frajhof, J. Borges, E. Hoffmann, J. Lopes, and R. Haddad, "Virtual reality, mixed reality and augmented reality in surgical planning for video or robotically assisted thoracoscopic anatomic resections for treatment of lung cancer," *The Journal of Visualized Surgery*, vol. 4, p. 143, 2018.

[30] X. Wu, L. Rong, Y. Jie, X. Song, Y. Cao, and S. Yang, "Mixed reality technology launches in orthopedic surgery for comprehensive preoperative management of complicated cervical fractures," *Surgical Innovation*, vol. 25, no. 4, p. 421, 2018.

[31] A. Bec, B. Moyle, V. Schaffer, and K. Timms, "Virtual reality and mixed reality for second chance tourism," *Tourism Management*, vol. 83, Article ID 104256, 2021.

[32] S. Bae, T. H. Jung, N. Moorhouse, M. Suh, and O. Kwon, "The influence of mixed reality on satisfaction and brand loyalty in cultural heritage attractions: a brand equity perspective," *Sustainability*, vol. 12, no. 7, p. 2956, 2020.

[33] R. Liu, X. Wang, H. Lu et al., "SCCGAN: style and characters inpainting based on CGAN," *Mobile Networks and Applications*, vol. 26, no. 1, pp. 3–12, 2021.

[34] B. Cao, J. Zhao, Z. Lv, and P. Yang, "Diversified personalized recommendation optimization based on mobile data," *IEEE Transactions on Intelligent Transportation Systems*, vol. 22, no. 4, pp. 2133–2139, 2021.

[35] W. Zhou, L. Yu, Y. Zhou, W. Qiu, M. W. Wu, and T. Luo, "Local and global feature learning for blind quality evaluation of screen content and natural scene images," *IEEE Transactions on Image Processing*, vol. 27, no. 5, pp. 2086–2095, 2018.

[36] D. Li, S. S. Ge, and T. H. Lee, "Simultaneous arrival to origin convergence: sliding-mode control through the norm-normalized sign function," *IEEE Transactions on Automatic Control*, vol. 67, no. 4, pp. 1966–1972, 2022.

[37] B. Zhu, Q. Zhong, Y. Chen et al., "A novel reconstruction method for temperature distribution measurement based on ultrasonic tomography," *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 69, no. 7, pp. 2352–2370, 2022.

[38] X. Wu, W. Zheng, X. Chen, Y. Zhao, T. Yu, and D. Mu, "Improving high-impact bug report prediction with combination of interactive machine learning and active learning," *Information and Software Technology*, vol. 133, Article ID 106530, 2021.

[39] T. Cai, D. Yu, H. Liu, and F. Gao, "Computational analysis of variational inequalities using mean extra-gradient approach," *Mathematics*, vol. 10, no. 13, p. 2318, 2022.

[40] Z. Wang, L. Dai, J. Yao et al., "Improvement of Alcaligenes sp. TB performance by Fe-Pd/multi-walled carbon nanotubes: enriched denitrification pathways and accelerated electron transport," *Bioresource Technology*, vol. 327, Article ID 124785, 2021.

[41] Y. Lin, H. Song, F. Ke, W. Yan, Z. Liu, and F. Cai, "Optimal caching scheme in D2D networks with multiple robot helpers," *Computer Communications*, vol. 181, pp. 132–142, 2022.

[42] X. Zhang, Y. Wang, M. Yang, and G. Geng, "Toward concurrent video multicast orchestration for caching-assisted mobile networks," *IEEE Transactions on Vehicular Technology*, vol. 70, no. 12, pp. 13205–13220, 2021.

[43] W. Zhou, H. Wang, and Z. Wan, "Ore image classification based on improved CNN," *Computers & Electrical Engineering*, vol. 99, Article ID 107819, 2022.

[44] J. Yan, H. Jiao, W. Pu, C. Shi, J. Dai, and H. Liu, "Radar sensor network resource allocation for fused target tracking: a brief review," *Information Fusion*, vol. 86-87, pp. 104–115, 2022.

[45] Y. Ban, M. Liu, P. Wu et al., "Depth estimation method for monocular camera defocus images in microscopic scenes," *Electronics*, vol. 11, no. 13, p. 2012, 2022.

[46] Q. Wang, G. Zhou, R. Song, Y. Xie, M. Luo, and T. Yue, "Continuous space ant colony algorithm for automatic selection of orthophoto mosaic seamline network," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 186, pp. 201–217, 2022.

[47] W. Zheng, X. Liu, and L. Yin, "Research on image classification method based on improved multi-scale relational network," *PeerJ Computer Science*, vol. 7, p. e613, 2021.

[48] Z. Ma, W. Zheng, X. Chen, and L. Yin, "Joint embedding VQA model based on dynamic word vector," *PeerJ Computer Science*, vol. 7, p. e353, 2021.

[49] W. Zheng, L. Yin, X. Chen, Z. Ma, S. Liu, and B. Yang, "Knowledge base graph embedding module design for Visual question answering model," *Pattern Recognition*, vol. 120, Article ID 108153, 2021.

[50] M. Qi, S. Cui, X. Chang et al., "Multi-region nonuniform brightness correction algorithm based on L-channel gamma transform," *Security and Communication Networks*, vol. 2022, Article ID 2675950, , 2022.

[51] S. Jiang, J. Zhou, and S. Qiu, "Digital agriculture and urbanization: mechanism and empirical research," *Technological Forecasting and Social Change*, vol. 180, Article ID 121724, 2022.

Hindawi
Computational Intelligence and Neuroscience
Volume 2022, Article ID 9123922, 9 pages
https://doi.org/10.1155/2022/9123922

Research Article

Couplet Analysis of Linguistic Topology Using Deep Neural Networks in Cognitive Linguistics

Dongmei Zhu,[1] Nan Wang ,[2] and Fuqiang Yang[3]

[1]*School of Foreign Languages, Henan University, Kaifeng 475001, China*
[2]*College of Literature and Art, Shihezi University, Shihezi 832000, China*
[3]*School of Electrical & Information Engineering, Kunsan National University, Gunsan, Jeollabuk-do 54150, Republic of Korea*

Correspondence should be addressed to Nan Wang; zdm330511@henu.edu.cn

Received 14 July 2022; Revised 5 September 2022; Accepted 3 October 2022; Published 11 October 2022

Academic Editor: Ning Cao

The work reported here primarily aims to realize the automatic generation of couplets using the linguistic topology of deep neural network (DNN). First, the symmetry, topology, and cognitive linguistics of language are explored to lay a theoretical foundation for subsequent model establishment and analysis. Then, the recurrent neural network (RNN) is employed to build the Seq2Seq model, and Liweng's Guide to Rhyme (an ancient Chinese enlightenment reading material to poetry creation) is imported into the Seq2Seq model as a basic corpus. Eventually, the entire system is implemented automatically on TensorFlow. The system undergoes tests of the five-character quatrain, the seven-character quatrain, the couplet, and the part-of-speech detection. Results demonstrate that both the first and the second lines of the couplet present an excellent correspondence regarding sentences and words. After some famous verses are entered, the second line of the couplet obtained is quite vivid and appropriate. Meanwhile, the results can be generated quickly and meet the requirements on rhyme and couplet matching. This model can input verses according to users' own needs and generate the second line of the couplet quickly, showing good correspondence in words, part-of-speech, and sentence pattern. Because the couplet belongs to Chinese traditional culture, it has a strong local Chinese cultural flavor. The system designed based on computer technology can help people learn and experience the charm of couplets.

1. Introduction

In the early 20th century, the computer was just born, while scientists had already begun to use computers to analyze human language. The main purpose at that time was for strategic needs because it was in the age of war, and the decryption and transmission of information belonged to high-level secrets. Later, when it came to the Cold War Period, the United States used the language analysis function of computers to decipher Russian confidential information into English and finally achieved real-time monitoring of Soviet defense scientific and technological information [1]. In this process, the US research team got inspiration from the information deciphering process of the Soviet Union; they believed that they only used different coding expressions for the same semantics while deciphering different languages. This inspiration fulfilled the early machine translation idea and also created a new era of natural language [2].

From 1970 to 1980, because the international environment entered a period of slightly stable relaxation, a series of science and technology developed rapidly. Computer was one of the important technologies, and simultaneously, Udine's million-level corpus was processed. An independent field with natural language processing as the core was established in the general direction of artificial intelligence in the following decade [3]. At this time, two different branches emerged in natural language processing: the rule school and the statistics school. The branch based on linguistic theories was the rule school. After linguistic scientists explained language phenomena, scholars of this school described the language using explanations made by linguistics. The other school was based on data statistics of corpus. This school focused on mathematics and could find laws from large

numbers of language materials to represent natural language. Until around 1990, natural language processing greatly affected people's lives with the continuous development of Internet technology. In China, people can use Chinese search engines to edit Chinese words. Many problems have been solved with a series of breakthroughs of algorithms. These algorithms are summarized as "machine learning" [4]. These algorithms are basically based on the working principles of the brain and neurons and gradually mature by simulating some human cognition and behaviors. Because many algorithms in machine learning can efficiently handle multi-dimensional and nonlinear problems, they have attracted many research teams.

Comerica used simple complex topological structures to simulate political and cultural structures. They translated wedge, cone, and hanging operations into the language of political structures and showed how these structures correspond to merged structures and the introduction of intermediaries. They introduced the concepts of agent's viability and political system stability, examined their interaction with simple and complex topologies, and expressed the interaction between them in the language of topology theory as much as possible [5]. Kiran et al. proposed a framework for language restoration, which included a network-based view of the brain regions involved in language restoration. Besides, the mechanism model of language recovery was established to explain individual differences in behavior, network topology, and treatment response [6]. Shen and Ho combined information and communication technology to improve teaching and learning. Through latent semantic analysis of the topic, they summarized the state of knowledge accumulation and helped people quickly understand the articles [7].

This couplet analysis of language topology based on deep neural network (DNN) first discusses the symmetry and topology of language and then establishes the Seq2Seq model through the recurrent neural network (RNN). Besides, a corpus is imported into the Seq2Seq model. Finally, the TensorFlow framework is used to realize the automatic operation of the whole system. It is hoped to establish and match couplets quickly and efficiently through deep learning. The innovation lies in the optimization of the artificial neural network (ANN) and using the RNN to achieve the repeated transmission of text content, which minimizes the final matching between the second line of a couplet and the first line of a couplet.

2. Theoretical Analysis and Research Design

2.1. Symmetry of Language. To some extent, language signs can be regarded as symmetric operation signs. These signs can be combined through symmetric conversion to output the language. During language expression, each character is a conversion operation. The symmetry of language is an attribute of all languages [8, 9]. Symmetry is one of the inherent properties of language. Chinese is precisely one of the best languages in the world. Because Chinese is gradually transformed from hieroglyphics, the symmetry shown in Chinese is even more unique than other languages. Chinese has perfect symmetry in text, phonetics, grammar, and rhetoric, which is also a big difference from other languages [10]. The overall shape of Chinese characters is square, and each word has a special meaning and a single syllable, which once again highlights the superiority of Chinese.

Regarding dimensional analysis, Chinese sentences are more prominent in terms of symmetry. There are many representative Chinese writing styles, such as poetry and couplet, which are similar to parallel prose and antithetical sentences in ancient Chinese literature [11]. In the earlier Shang Dynasty (1600–1046 BC), there was already a rhetorical method of Chinese antithesis, and both verses and prose in that period possessed extremely high antithesis. Since the Spring and Autumn Period and the Warring States Period (770–221 BC), the majority of literati had begun to use the rhetoric of confrontation. It was not until the Wei and Jin Dynasties (220–420 AD) and the Southern and Northern Dynasties (420–589 AD) that the four-six-character couplet appeared. By the time of the Tang Dynasty (618–907 AD), rhythmical prose characterized by parallelism and ornateness was quite mature [12, 13]. The couplet has existed in China for a long time; however, it never reached its peak until the Song Dynasty (960–1279 AD) and Yuan Dynasty (1271–1368 AD). In the Ming Dynasty (1368–1644 AD) and Qing Dynasty (1636–1912 AD), the couplet reached its climax. For example, "Two golden orioles sing amid the willows green; A flock of white egrets flies into the blue sky" is a very antithetical poem proposed by Du Fu (712–770 AD), a famous poet in the Tang Dynasty. Although it has developed to modern times, Chinese still has the symmetrical structure [14, 15].

2.2. Linguistic Topology. Topology first appeared around 1900. Linguistic topology has a special meaning that can be analyzed from three levels. The first level of meaning is that linguistic topology is transformed from psychology and then classified based on the characteristics of language pronunciation, part-of-speech, and sentence pattern [16]. It belongs to the most widely used linguistics in the field of topology. According to the above statement, every language has its own topology. Although there may be different topologies in other aspects, the function of topology is to reasonably divide languages into different categories [17, 18]. In the 19th century, some scholars focused on the lexical classification of languages in early linguistics, and this classification was regarded as the topological classification. This definition introduces the basic meaning of topology to contemporary linguistics: topology is closely connected to the comparison between different languages. Topology is a method of forming linguistic theory [19]. This means that it is a method of language analysis that establishes a linguistic theory different from other methods. In the meantime, the ideas and propositions should be closely associated with functional linguistics. Therefore, topology in this sense is often referred to as a functional topology method. The mapping between topological spaces f: $X-Y$ is bijective, and f and $f-1$ are continuous; then, f is a homeomorphism [20]. If there is a homeomorphic mapping from X to Y, then X and Y will be

considered homeomorphic, and this process is also called topological transformation. The first line of the couplet is equivalent to X, and the second line of the couplet is equivalent to Y. The first line of the couplet is entered to generate the second line of the couplet. This process is a topological mapping from the first line to the second line of the couplet [21].

2.3. Cognitive Linguistics. Cognitive linguistics does not belong to a particular kind of linguistics because it is also mixed with theories of cognitive neurology. Hence, it is interdisciplinary. It was born under the second generation of cognitive and empirical science based on the opposing mainstream linguistic transformational grammar, which was gradually formed from 1980 to 1990 [22, 23]. The development of cognitive linguistics has now included subjects in many fields, such as psychology, artificial intelligence, and systems theory subjects. Meanwhile, cognitive linguistics has an important theoretical basis: human beings will promote language innovation, learning, and use during the cognition process of external things since the source of human knowledge is that human beings are born with cognitive abilities. Cognitive linguistics has had a significant impact on traditional linguistics once it came into being [24].

2.4. Corpus Establishment. The primary concern is data; in natural language processing, it is the corpus [25]. Second, corpus must be carefully screened in terms of quality and data scale to ensure that the results generated have sufficient accuracy [26]. Due to the large capacity of computer storage, authentic text data, and fast and accurate information extraction, linguists can use the electronic corpus to describe language from multiple angles and levels, verify various language theories and hypotheses, and even establish new language models and language concept. The basic corpus adopted by the couplet system in the present work is Liweng's Guide to Rhyme (an ancient Chinese enlightenment reading material to poetry creation). Liweng's Guide to Rhyme is a corpus adopted by the Intelligent Science Laboratory for a long time. Indeed, only using Liweng's Guide to Rhyme as the corpus is far from enough, and the effect of the couplet will be not good [27, 28]. Under machine learning theories, enough language data must be gathered to ensure the automatic operation of the couplet system. At this time, a huge corpus is required, which must contain most of China's excellent couplets and poetry information for thousands of years. Therefore, before the experiment, some extra corpus is collected on the Internet to improve the quality of couplet generation. In the present work, a large number of Chinese classic couplet sentences and antithetical sentences are selected to maximize the content of the corpus used. At this stage, the construction of corpus has been completed.

2.5. Seq2Seq Model Establishment

2.5.1. ANN. ANN is basically similar to topology, and both are cognitive models. In the present work, neural networks will constitute not only an important part of deep learning technology but also the core of the Seq2Seq couplet generation model [29]. Using the Seq2Seq model is inspired by previous literature, and it can mimic the human brain and thereby human thinking. The composition of the human brain is quite complex. According to statistics, the human brain is probably composed of nearly 100 billion neurons, and the relationship among them is intricate. Because of this complex relationship, ANN used this time also needs to build a large number of "neuron" structures, and each "neuron" is an activation function [30]. A connection is established between any two neurons, called a weighted value, which is a temporary memory of the input. The activation function will process the input accordingly. The output of the neural network is determined by the weight and the activation function. The simplest neural network consists of only one neuron [31], as shown in Figure 1.

Equation (1) describes the weighted input variables.

$$x^j = \sum_{i=1}^{n} x_i^{j-1} w_i^{j-1}. \tag{1}$$

Equation (1) indicates that the input variable of a layer is the sum of the product of all variables of the previous layer and the weight. The activation function is an output function, which is the secondary processing of intermediate results, as shown in the following equation:

$$y_j = f\left(x_j\right). \tag{2}$$

A complex neural network is composed of two interconnected neurons between two layers. Its structure is demonstrated in Figure 2.

Figure 2 displays a multi-level structure in neural networks. Its structure can be expressed by a vector as presented in the following equation:

$$\text{Net} = w^T x + b, \tag{3}$$

where w represents the weight vector of each node, x denotes the input value of each layer, and b refers to the paranoid vector that usually starts from 1. The function of ANNs is clearer and more intuitive by simplifying all values to vectors.

2.5.2. RNN. ANN is the prototype of RNN. RNN is an optimization over ANN. Its structure is more complex than ANN, but it is more efficient in processing similar sequential data [32]. Figure 3 displays the simplest RNN structure.

Through Figure 3, assuming a neural network is A, it is essential to obtain the input X_t of a certain state and then output a value H_t. Due to cyclic use, information can be passed from the current time step to the next time step. To be specific, the same network will undergo numerous cycles in different time states, and each neuron will send the updated results to the neuron at the next moment. Figure 4 illustrates the internal structure of an RNN.

In Figure 4, the arrow stands for the direction of signal movement towards the next step. The mathematical formulas of this process are as follows. The lateral movement can be written as

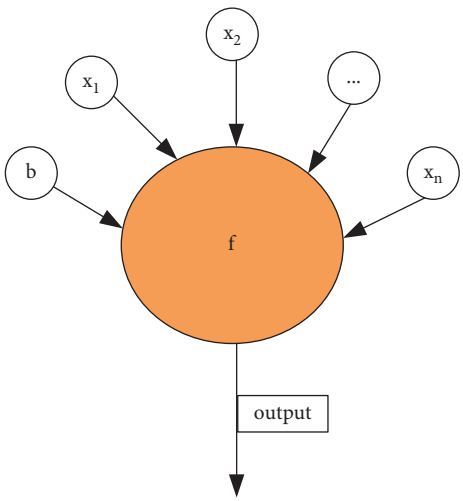


Figure 1: Neuron structure.

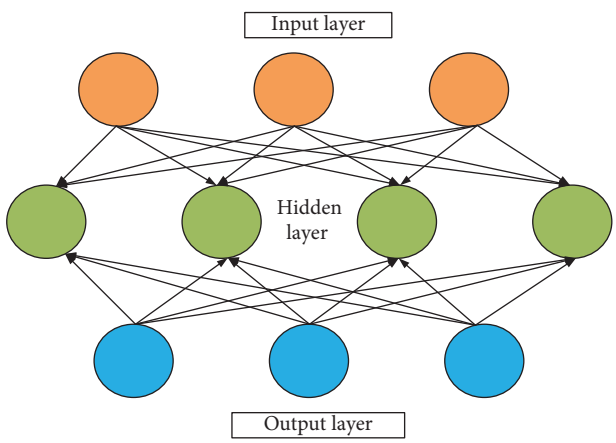


Figure 2: Neural network structure.

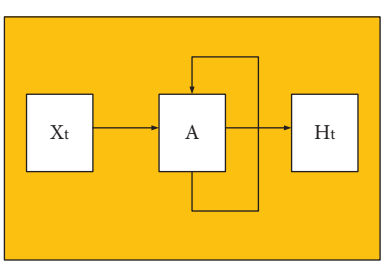


Figure 3: RNN structure.

$$h_1 = f\left(\mathrm{U}x_1 + \mathrm{W}h_0 + \mathrm{b}\right). \tag{4}$$

The subsequent states are the repetition of the first process, and the parameters are still U, W, and b. The rest can be deduced by analogy.

The equation for vertical transfer is

$$y_1 = \mathrm{soft\,max}\left(\mathrm{V}h_1 + \mathrm{c}\right). \tag{5}$$

Similarly, the subsequent output and the first output follow the same law. Each arrow is a transformation process,

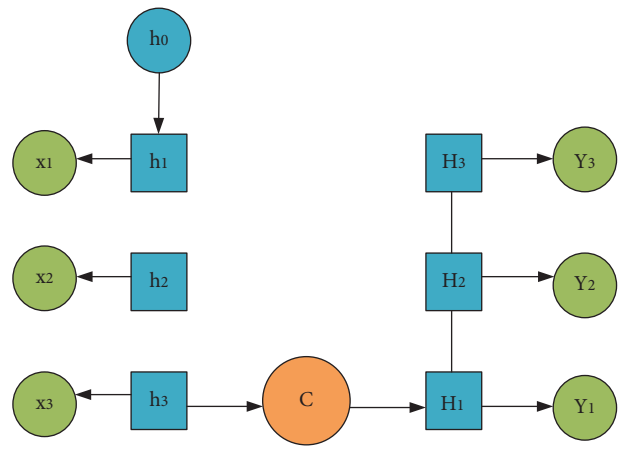


Figure 4: Complete RNN internal structure.

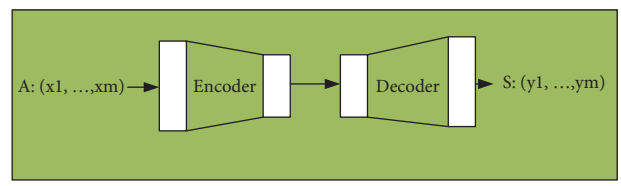


Figure 5: Sequence couplet automatic generation models.

and on this basis, the calculation continues [33]. In the field of natural language processing, c is defined as a context vector for communicating encoding and decoding. Through continuous iteration, all input sequences are ultimately converted into output sequences.

Softmax function is one of the activation functions used in multi-classification problems, and it can normalize the output value to a probability value [34]. Here, it is used to optimize the output of the RNN. Assume that C is the number of categories to be predicted and a is the output; then, C outputs are denoted as a_1, a_2..., a_c. Therefore, for a sample, the probability that it belongs to class i is the value of softmax function. Its mathematical expression is

$$\mathrm{softmax} = \frac{e^a}{\sum_{K=1}^{C} e^a k}. \tag{6}$$

2.5.3. Seq2Seq Model. The most basic Seq2Seq model consists of three parts: the encoder, the decoder, and the intermediate state vector connecting the former two.

First, the most basic model, the serial couplet system, is introduced. After receiving the signal of the input sentence, the sequence couplet model system performs data statistics on the special part-of-speech in the sentence by analyzing the characters in the first line of the couplet using RNN. In this way, a single vector representing the preceding clause is obtained. At this time, another RNN is used to decode the input vector into the subsequent clauses through intelligent character generation. Basically, this process generates a sequence through encoding and decoding, which is a global sentence. Figure 5 shows the process of continuous couplet generation.

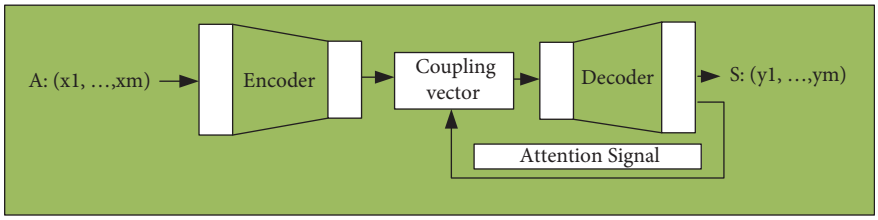


Figure 6: Couplet generation model under attention mechanism.

The second model is based on the couplet generation of the attention mechanism. The attention mechanism simulates the alignment between the input and output positions, which can be regarded as a local matching model. Besides, the pitch coding problem can also be solved by paired attention mechanisms. Figure 6 displays the extension of the attention mechanism to the sequence couplet generation model.

The couplet system reported here adopts the statistical method in the field of natural language processing. Based on the principle of machine learning, the model needs continuous training. Only when the model is optimized can the performance of the couplet system improve. The purpose of model training is to optimize the parameters and algorithms, and the evaluation criteria are responsible for testing the effectiveness of the model. The loss function is such an evaluation function. Here, cross entropy is selected as a loss function. The mathematical definition of cross entropy is shown in the following equation:

$$H(p,q) = \sum_i p(i)\log\frac{1}{q(i)}, \tag{7}$$

where p and q denote two different distributions, of which one is the real distribution, and the other is the predictive distribution. The similarity between the distribution of the test set and the real situation can be calculated by (7) to judge the accuracy of the model in the test set.

2.6. Framework Design of Machine Learning Algorithm. TensorFlow is the leader among many deep learning frameworks; thanks to its excellent performance and algorithm execution efficiency, it is preferred by many researchers and programmers [35–38]. TensorFlow has a crucial component, the client, which connects to the master and multiple workers via the session interface. Each worker is connected with other hardware devices, and simultaneously, a central processing unit or a graphics processor must be used instead to achieve the goal, thereby controlling each kind of hardware. The system flow is demonstrated in Figure 7.

Figure 7 reveals the complete development process of the algorithm. From the above flowchart, this is a typical algorithm development process including observation, experiment, model determination, test, model update, and result output, and every step is indispensable. The continuous iteration is a process of keeping improving and selecting the optimal. At the same time, the part-of-speech

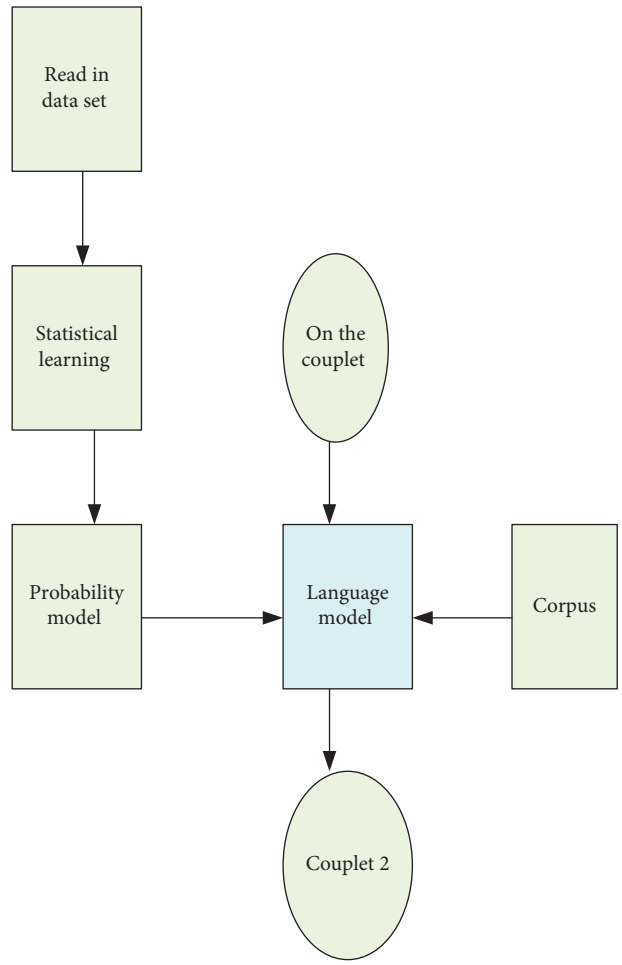


Figure 7: Flowchart of the couplet automatic generation system.

dimensions (16, 32, 64, and 128) of the second line of couplet output shall be tested to determine the corresponding relationship between the words in the first line and second line of couplets.

The development environment of this test is as follows:

Hardware: a notebook computer.

Operating system: Windows 10 system.

Development platform: PyCharm.

Development language: Python 3.6.

Calculation framework: TensorFlow 1.6.

The corpus used is saved in text format. One character is stored in each line, with a total of 9,131 words.

Table 1: Output results.

Topology	Input	Output
Five-character quatrain	Lush, lush grass on the plain	Red, red flower in the water
Seven-character quatrain	The morn rain o' Wei Town has laid the light dust clean	The spring wind o' Qinling Mountain has dyed the green lake blue

3. Test Results

3.1. Verse Test. Test of couplet output, to a large extent, is also the test of part-of-speech accuracy. Therefore, using poem verses for the part-of-speech test can be more conducive to evaluating the rationality of the system. In the present work, the classic five-character and seven-character quatrains in Tang poems are entered in the dialog box of the first line of the couplet. Through automatic generation, the output results are very close to the first lines of the couplet and are also very neat. However, they are also quite different from the original verses, showing a unique style. Hence, the Seq2Seq model can perform targeted learning and imitation after getting the language reminder. Nevertheless, different from the couplet, poetry is not particularly demanding for antithesis. The results are summarized in Table 1.

Based on the above two examples, the output second lines do not correspond well to the first lines of the couplet, and each word does not form a sharp contrast with each other in the corresponding position. One reason is that the first line of the couplet is a verse, which itself does not require a strict antithesis.

3.2. Couplet Test. Regarding the ancient couplet, a verse in Liweng's Guide to Rhyme is input: "Passenger in the post encounters the plum rain." The result given by the system is "People in the village get drunk with the wicker wind." In contrast, the second line of the couplet in the original text is "People in the pavilion pick the lotus wind." This example suggests that the automatically generated result does not exactly match with the first line of the couplet. Still, there is an internal connection between the part-of-speech comparison. Therefore, in the second test, the sentences in modern poems are selected. The verse "Knock the chess to find the way last night" is input as the first line of the couplet. The original text of the second line is "See my face in the mirror today." The result of the model system is "See my smiling face in the mirror today." Only one word is different, but the feeling is quite different; the "smiling face" is more vivid than the "face." Therefore, the designed couplet system shows good generation results whether for the ancient couplet or the modern couplet, proving its strong adaptability. Figure 8 presents the dimension detection of part-of-speech.

According to Figure 8, as the dimension of the part-of-speech vector becomes larger, the semantics of the word itself will change. However, regarding the score, the overall trend is stable, meaning that the first and the second lines of the couplet are better in terms of part-of-speech matching.

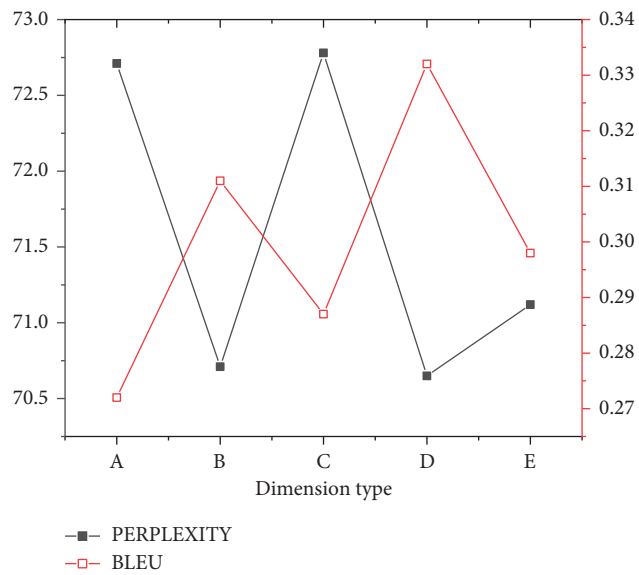


Figure 8: Dimension detection of part-of-speech. (a) 0 dimensions. (b) 128 dimensions. (c) 64 dimensions. (d) 32 dimensions. (e) 16 dimensions.

3.3. Reduplicated Word and Vocabulary Test. Figure 9 illustrates the results of the simulation experiment.

From the data curve in Figure 9, after a period of automatic generation training of couplets, the loss magnitude of the model gradually decreases, indicating that the generation results of the model have good accuracy. The art requirements of words in couplets are excessively high, so this experiment also tests reduplicated words. Through Figure 9, there is no obvious difference in the performance score of reduplicated words in different word frequencies, demonstrating that the model can reasonably understand and detect reduplicated words in the first line of couplets. When inputting "Fo Jiao Qing Quan Piao Piao Piao Piao Piao Xia Liang Tiao Yu Dai" in the dialog box of the first line of couplets of the system, the system outputs the second line of couplets as "Chan Xin Ming Yue Jiao Jiao Zhao Zhao Lai Lai Yi Dian Chan Xin," failing to reach a neat confrontation. The words "Jiao Jiao Zhao Lai" in the second line of couplets are obviously wrong in part-of-speech. It is obvious that there are two groups of reduplicated words plus a verb in "Piao Piao Piao Piao Piao" in the first line of couplets, but the system only gets three groups of reduplicated words in the second line of couplets, showing no correlation. Therefore, this system cannot precisely recognize reduplicated words in the first line of couplets for verse, and the accuracy of the algorithm needs to be improved. There are also two other

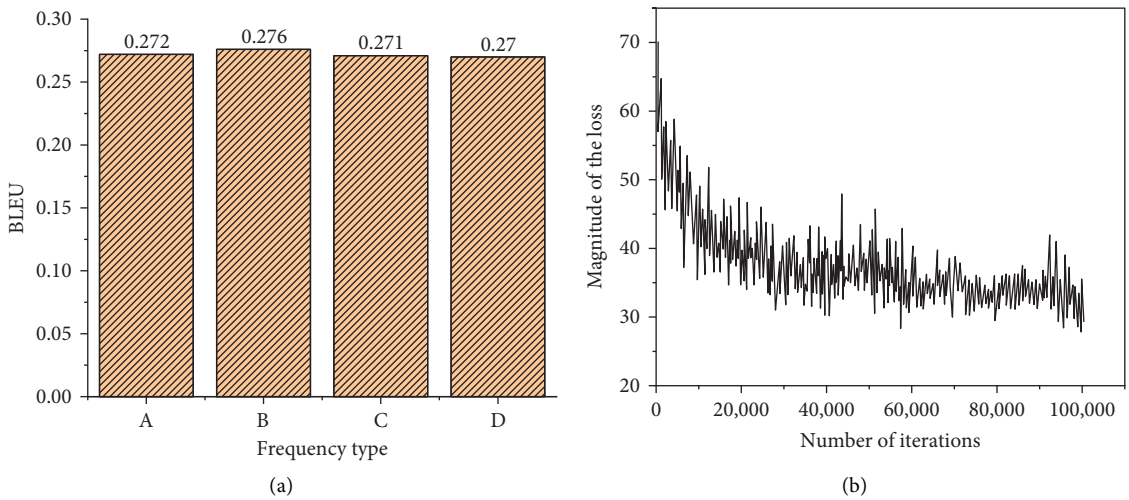


Figure 9: Results of the simulation experiment. (a) Performance testing of reduplicated words. (b) Result of the loss function (A: word frequency = 0; B: word frequency ≤ 1; C: word frequency ≤ 3; D: word frequency ≤ 5).

groups of experiments to detect reduplicated words, and the two pairs of couplets are "Shan Shan Shui Shui Chu Chu Ming Ming Xiu Xiu, Jia Jia Hu Hu Jia Jia Hao Hao Qi Qi" and "Xiu Xiu Ling Ling Yuan Yuan Jin Jin Cui Cui, Chun Zhen Pu Shi Chun Du Chun Pu Hou Hou Chun Zhen," respectively. Through the above three pairs of couplets, evidently, when dealing with a single word, the response of the system is often not satisfactory. The second line of couplets automatically generated by the system has low matching degree and relevance with the second line of couplets, and the output results are just a set of words similar in form but without particular meaning.

3.4. Boundary Test. According to actual application requirements, the maximum boundary value of the software presentation box must be checked. What is applied is a unique method in software engineering: black box testing. The test results are summarized in Table 2.

In the initial stage of the system, there is no input for the first line of the couplet, and the word count at this time is judged to be zero. If the first line of the couplet enters excessive words in the dialog box, the second line of the couplet will issue an early warning signal during the generation process, reminding "Too many words entered, please try again." This form of warning is similar to the reminder when the password does not match on the usual login page. If no word is entered in the first line of the couplet, no result will be generated in the second line of the couplet. When the number of words entered by the user is greater than 50 words, the system cannot continue. Because the number of words is strictly controlled when setting the system program: the above result will appear when the number of words input is greater than 50 words. Under normal circumstances, the number of the first line of the couplet input by the user will not exceed 30 words. When it exceeds 50 words, the couplet itself does not have any practical application value.

Table 2: Statistics on the output categories.

Input topology	Output topology
0 words	No output
0–50 words	Normal output
≥50 words	Too many words entered, please try again.

4. Discussion

Through experiments, the Seq2Seq model can handle the first line of the couplet in different categories reasonably, including the response to punctuation. In the meantime, the first and the second lines of the couplet present high connection and neatness. However, after many times of debugging, some weaknesses of the model are discovered. When the Seq2Seq model deals with single words, the words generated cannot be accurately matched with the words of the first line of the couplet. Only when the number of input words is large can the automatically generated second line of the couplet be improved.

In future works, improvement over corpus should be the first concern. Corpus is the most basic and most vital link to the entire system. If it cannot provide sufficient language content or has excessive errors, inaccurate results or high error rates will occur. In the entire couplet automatic generation system, the selection of corpus must not only have a strong pertinence but also allow the corpus to possess the generalization capability. Second, in terms of language model, couplet itself has many characteristics (the flatness of grammar, the correspondence of part-of-speech, and the degree of semantic relevance). These two aspects are not considered in the present work while establishing the language model. In this case, the next step is to optimize these problems. For example, at this stage, other algorithms can be added to the original model, and some words, phrases, and sentences that cannot meet the requirements can be deleted before the language model can be reestablished and executed. In the end, the designed model can be further

optimized regarding "matching degree" and "language expression habits" while meeting performance requirements. According to the test of the automatic couplet system launched by Microsoft Research Asia, many output results are very unsatisfactory. Meanwhile, as artificial intelligence develops quickly, various algorithms can be optimized continuously.

5. Conclusion

Under cognitive linguistics, the present work aims to automatically generate couplets using the linguistic topology of DNN and analyze the DNN outputs.

The symmetry, topology, and cognitive linguistics of language are explored to lay a theoretical foundation; next, RNN is employed to build the Seq2Seq model, which is then implemented automatically on TensorFlow. Results demonstrate that after the corpus is input, the output five-character quatrain, seven-character quatrain, part-of-speech, and couplet all present good correspondence. The couplets can be generated quickly and meet the requirements on rhyme and couplet matching. Still, this model has some limitations. The corpus used is not rich enough. Only one corpus is input, far from satisfying the requirements on couplet richness. Besides, the confrontation of couplet is not only the correspondence between words but also the association between phrases. Moreover, the requirements on words in modern couplets are higher, so that further optimization of the model system is required. Hopefully, the model can promote traditional Chinese culture by integrating artificial intelligence and other advanced technologies, in an effort to achieve better cultural inheritance and absorption.

Data Availability

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Authors' Contributions

All authors have made a substantial, direct, and intellectual contribution to the work and approved it for publication.

References

[1] R. K. Umarov, V. V. Solovyev, and I. B. Rogozin, "Recognition of prokaryotic and eukaryotic promoters using convolutional deep learning neural networks," *PLoS One*, vol. 12, no. 2, p. 0171410, 2017.

[2] T. P. Oliveira, J. S. Barbar, and A. S. Soares, "Computer network traffic prediction: a comparison between traditional and deep learning neural networks," *International Journal of Big Data Intelligence*, vol. 3, no. 1, pp. 28–37, 2016.

[3] S. Sundaravelpandian, S. Johan, and G. Philipp, "Deep-learning neural-network architectures and methods: using component-based models in building-design energy prediction," *Advanced Engineering Informatics*, vol. 38, no. 1, pp. 28–37, 2018.

[4] W. C. F. Mariel, S. Mariyah, and S. Pramana, "Sentiment analysis: a comparison of deep learning neural network algorithm with SVM and nave Bayes for Indonesian text," *Journal of Physics: Conference Series*, vol. 971, no. 25, p. 012049, Article ID 012049, 2018.

[5] S. T. Comerica, "Adaptive network tools and services network topology mapper," *Computing Reviews*, vol. 60, no. 10, p. 383, 2019.

[6] S. Kiran, E. L. Meier, and J. P. Johnson, "Neuroplasticity in aphasia: a proposed framework of language recovery," *Journal of Speech, Language, and Hearing Research*, vol. 62, no. 11, pp. 3973–3985, Oct. 2019.

[7] C. W. Shen and J. T. Ho, "Technology-enhanced learning in higher education: a bibliometric analysis with latent semantic approach," *Computers in Human Behavior*, vol. 104, no. 3, Article ID 106177, 2020.

[8] J. Wang, H. Cao, J. Z. H. Zhang, and Y. Qi, "Computational protein design with deep learning neural networks," *Scientific Reports*, vol. 8, no. 1, p. 6349, 2018.

[9] K. Ghosh, A. Stuke, M. Todorovic et al., "Machine learning: deep learning spectroscopy: neural networks for molecular excitation spectra (adv. Sci. 9/2019)," *Advancement of Science*, vol. 6, no. 9, Article ID 1970053, 2019.

[10] B. P. Nayoga, R. Adipradana, R. Suryadi, and D. Suhartono, "Hoax analyzer for Indonesian news using deep learning models," *Procedia Computer Science*, vol. 179, no. 2, pp. 704–712, 2021.

[11] D. Zhang and S. Lou, "The application research of neural network and BP algorithm in stock price pattern classification and prediction," *Future Generation Computer Systems*, vol. 115, no. 31, pp. 872–879, 2021.

[12] M. E. Karar, F. Alsunaydi, S. Albusaymi, and S. Alotaibi, "A new mobile application of agricultural pests recognition using deep learning in cloud computing system," *Alexandria Engineering Journal*, vol. 60, no. 5, pp. 4423–4432, 2021.

[13] Y. Weng, X. Wang, J. Hua, H. Wang, M. Kang, and F. Y. Wang, "Forecasting horticultural products price using ARIMA model and neural network based on a large-scale data set collected by web crawler," *IEEE Transactions on Computational Social Systems*, vol. 6, no. 3, pp. 547–553, 2019.

[14] A. A. Ojugo and R. E. Yoro, "Forging a deep learning neural network intrusion detection framework to curb the distributed denial of service attack," *International Journal of Electrical and Computer Engineering*, vol. 11, no. 2, p. 1498, 2021.

[15] S. J. Li, "Some developments of estimation procedures of mechanical model parameter based on deep learning neural network," *Artificial Intelligence and Robotics Research*, vol. 09, no. 02, pp. 100–109, 2020.

[16] S. Lee, J. Kim, H. Kang, D. Y. Kang, and J. Park, "Genetic algorithm based deep learning neural network structure and hyperparameter optimization," *Applied Sciences*, vol. 11, no. 2, p. 744, 2021.

[17] C. B. Gemirter and D. Goularas, "A Turkish question answering system based on deep learning neural networks," *Journal of Intelligent Systems: Theory and Applications*, vol. 4, no. 2, pp. 65–75, 2021.

[18] M. Gui and X. Xu, "Technology forecasting using deep learning neural network: taking the case of robotics," *IEEE Access*, vol. 9, no. 99, pp. 53306–53316, 2021.

[19] S. Singla and A. Patel, "Comparative study of the deep learning neural networks on the basis of the human activity

recognition," *International Journal on Computer Science and Engineering*, vol. 8, no. 11, pp. 27–32, 2020.

[20] J. Sarivougioukas and A. Vagelatos, "Modeling deep learning neural networks with denotational mathematics in UbiHealth environment," *International Journal of Software Science and Computational Intelligence*, vol. 12, no. 3, pp. 14–27, Sept. 2020.

[21] R. Y. Choi, A. S. Coyner, J. Kalpathy-Cramer, and P. Campbell, "Introduction to machine learning, neural networks, and deep learning," *Transl Vis Sci Techn*, vol. 9, no. 2, p. 14, 2020.

[22] S. Devunooru, A. Alsadoon, and P. Chandana P, "Deep learning neural networks for medical image segmentation of brain tumours for diagnosis: a recent review and taxonomy" J AMB," *INTEL HUM COMP*, vol. 12, no. 4, pp. 1–29, 2021.

[23] A. Alharbi and O. Sohaib, "Technology readiness and cryptocurrency adoption: PLS-SEM and deep learning neural network analysis," *IEEE Access*, vol. 9, no. 99, pp. 21388–21394, 2021.

[24] S. Saha, Z. Gan, L. Cheng et al., "Hierarchical deep learning neural network (HiDeNN): an artificial intelligence (AI) framework for computational science and engineering," *Computer Methods in Applied Mechanics and Engineering*, vol. 373, no. 113, Article ID 113452, 2021.

[25] V. Srividhya, K. Sujatha, R. Ponmagal, G. Durgadevi, and L. V Madheshwaran, "Vision based detection and categorization of skin lesions using deep learning neural networks," *Procedia Computer Science*, vol. 171, no. 1, pp. 1726–1735, 2020.

[26] S. Zhang, R. Chen, W. Du, Y. Yuan, and V. S. Vassiliadis, "A hessian-free gradient flow (HFGF) method for the optimisation of deep learning neural networks," *Computers & Chemical Engineering*, vol. 141, no. 1, Article ID 107008, 2020.

[27] S. Sumathi and N. Karthikeyan, "Detection of distributed denial of service using deep learning neural network," *Journal of Ambient Intelligence and Humanized Computing*, vol. 12, no. 1, pp. 1–11, 2020.

[28] Y. Qin, K. Li, Z. Liang et al., "Hybrid forecasting model based on long short term memory network and deep learning neural network for wind signal," *Applied Energy*, vol. 236, no. 1, pp. 262–272, 2019.

[29] Z. M. Yaseen, H. A. Afan, and M. T. Tran, "Beam-column joint shear prediction using hybridized deep learning neural network with genetic algorithm," *IOP Conference Series: Earth and Environmental Science*, vol. 143, no. 1, Article ID 012025, 2018.

[30] B. Rajalingam and R. Priya, "Multimodal medical image fusion based on deep learning neural network for clinical treatment analysis," *International Journal of Chemistry Research*, vol. 11, no. 6, pp. 160–176, 2018.

[31] Y. Feng, L. Li, and H. Liu, "SAR image target recognition algorithm based on deep learning neural network," *BOLETIN TECNICO*, vol. 55, no. 14, pp. 55–60, 2017.

[32] A. Chernoded, L. Dudko, I. Myagkov, and P. Volkov, "Deep learning neural networks and bayesian neural networks in data analysis," *EPJ Web of Conferences*, vol. 158, no. 1, Article ID 06008, 2017.

[33] W. Zheng and L. Yin, "Characterization inference based on joint-optimization of multi-layer semantics and deep fusion matching network," *PeerJ Computer Science*, vol. 8, p. 908.

[34] H. Zhu, M. Xue, Y. Wang, G. Yuan, and X. Li, "Fast visual tracking with siamese oriented region proposal network," *IEEE Signal Processing Letters*, vol. 29, pp. 1437–1441, 2022.

[35] X. Liang, L. Luo, S. Hu, and Y. Li, "Mapping the knowledge frontiers and evolution of decision making based on agent-based modeling," *Knowledge-Based Systems*, vol. 250, Article ID 108982, 2022.

[36] J. Yu, L. Lu, Y. Chen, Y. Zhu, and L. Kong, "An indirect eavesdropping attack of keystrokes on touch screen through acoustic sensing," *IEEE Transactions on Mobile Computing*, vol. 20, no. 2, pp. 337–351, 2021.

[37] Y. Zhang, X. Shi, H. Zhang, Y. Cao, and V. Terzija, "Review on deep learning applications in frequency analysis and control of modern power system," *International Journal of Electrical Power & Energy Systems*, vol. 136, Article ID 107744.

[38] Y. Choi, J. Wang, Y. Zhu, and W. F. Lai, "Students' perception and expectation towards pharmacy education: a qualitative study of pharmacy students in a developing country," *Indian Journal of Pharmaceutical Education and Research*, vol. 55, no. 1, pp. 63–69, 2021.

Hindawi
Computational Intelligence and Neuroscience
Volume 2022, Article ID 8336616, 10 pages
https://doi.org/10.1155/2022/8336616

Research Article

The Generation of Piano Music Using Deep Learning Aided by Robotic Technology

Jian Pan,[1] Shaode Yu,[2] Zi Zhang,[1] Zhen Hu,[3] and Mingliang Wei[4]

[1]*College of Music and Dance, Sichuan Film and Television University, Chengdu 611331, Sichuan, China*
[2]*School of Information and Communication Engineeringn, Communication University of China, Beijing 100024, China*
[3]*Information and Communication Engineeringn, Chengdu Medical College, Chengdu 610500, Sichuan, China*
[4]*Science and Technology Training School of Qimengweilai, Chengdu 611000, Sichuan, China*

Correspondence should be addressed to Mingliang Wei; 201428011132@chd.edu.cn

Received 2 August 2022; Revised 2 September 2022; Accepted 28 September 2022; Published 10 October 2022

Academic Editor: Ning Cao

In order to improve the accuracy and precision of music generation assisted by robotics, this study analyzes the application of deep learning in piano music generation. Firstly, based on the basic concepts of robotics and deep learning, the advantages of long short-term memory (LSTM) networks are introduced and applied to the piano music generation. Meanwhile, based on LSTM, dropout coefficients are used for optimization. Secondly, various parameters of the algorithm are determined, including the effects of the number of iterations and neurons in the hidden layer on the effect of piano music generation. Finally, the generated music sequence spectrograms are analyzed to illustrate the accuracy and rationality of the algorithm. The spectrograms are compared with the music sequence spectrograms generated by the traditional restricted Boltzmann machine (RBM) music generation algorithm. The results show that (1) when the dropout coefficient value is 0.7, the function converges faster, and the experimental results are better; (2) when the number of iterations is 6000, the error between the generated music sequence and the original music is the smallest; (3) the number of hidden layers of the network is set to 4. When the number of neurons in each hidden layer is set to 1024, the training result of the network is optimal; (4) compared with the traditional RBM piano music generation algorithm, the LSTM-based algorithm and the sampling frequency distribution tend to be consistent with the original sample. The results show that the network has good performance in music generation and can provide a certain reference for automatic music generation.

1. Introduction

In recent years, with the continuous development of theory and technology in the field of artificial intelligence (AI), its related research results have been widely used. For example, language recognition, image recognition, natural language processing, and other related achievements have brought a lot of convenience to all aspects of people's lives [1–4]. Robotics is one of the main technologies in the field of AI. Music robotics is a top-level applied discipline at the intersection of music and technology. Since the 21st century, music robot technology has been widely developed, including different technologies such as the principle of music robot pronunciation, expressive performance, bionic structure, and intelligence. As one of the symbols of "AI," the problem of music generation by music robots has become a research hotspot [5]. The most important thing in the field of music style recognition and generation is the extraction of relevant features and the selection of classifiers. Different music feature vectors are selected for music style recognition, and different classification effects will be produced. At present, the identification of musical styles still uses features such as pitch, timbre, and loudness. If this study only relies on manually extracted features, the speed and accuracy of classification will be greatly reduced. In order to obtain a more accurate recognition effect, it is also necessary to deeply mine the intrinsic correlation between the data [6].

With the deepening of research, deep learning models have been widely used in music generation. These deep models include recurrent neural network (RNN), generative adversarial network (GAN), restricted Boltzmann machine (RBM), convolutional neural network (CNN), and long short-term memory (LSTM) [7]. In addition, the mixed-use of networks in reinforcement learning (RL) and deep learning is also used in music generation. Each algorithm has certain drawbacks. For RNN, since the network does not have the effect of long-term memory, the effect of the generated music is not very good. For GAN, the network is not very good at dealing with variable content. A problem that is easy to arise is adding some rhythms to the generated music that are different from the original music. RBMs are slightly insufficient in controllability. Some scholars have used the forward feature selection algorithm to extract the underlying features of music. These features include spectral center value, linear prediction coefficient, zero-crossing rate, and more than ten musical features. Music features are combined with multimodal analysis and identification methods to finally achieve the effect of music genre classification. It mainly contains five music genres: classical, country, pop, jazz, and rock. In addition, some scholars also perform music recognition by extracting the rhythm and pitch from the music features to form a two-dimensional feature vector and finally achieve the effect of music recognition [5, 8–11]. Hizlisoys et al. proposed a method for music emotion recognition based on a deep neural network (DNN) architecture with convolutional LSTM. The CNN layer provides log Mel filter bank energy and Mel frequency cepstral coefficients to obtain features. The classification results show that the best performance can be obtained when the new feature set is combined with standard features using LSTM fusion DNN classifiers [12]. Li et al. transferred DNN models to music classification and used spectrograms to evaluate the performance of the models, as shown by extensive experimental evaluations on three music datasets. The balanced trust loss function model, Resnet50_trust, consistently outperforms other DNN models [13].

Music genre identification and classification have been deeply analyzed in music robotics research. Deep learning is also used to generate music and has a certain research basis. As a very important part of deep learning, RNN has also made great strides in music generation. The network is very powerful in dealing with long-time series problems. However, using an RNN model only for simple music is not ideal. This paper introduces the related technologies and theories that are needed for different genres of music generation in deep learning and summarizes some key technologies currently used in music feature extraction. Based on the LSTM network, the music style recognition and generation networks are redesigned. The dimensions of the input network matrix and output matrix are designed to make training easier. The comparison between different spectrograms illustrates the accuracy of the experiment, and genre classification predictions are also performed on the generated music to illustrate that the network can generate music of different genres.

2. Introduction of Related Algorithms and Model Establishment and Training

2.1. Music Generation and Robots. As a creative artistic expression, music is the unique crystallization of human wisdom and emotion. Therefore, automatically generating music with a clear style and in line with the aesthetics of listening, especially different from existing music, has become one of the criteria for AI evaluation. This is also a hot spot for deep learning applications [14].

Most methods use the musical instrument digital interface (MIDI) as the notation of music [15]. Its advantages are that it is straightforward, the result of the model is the musical score, and the training data are easy to obtain. Its downside is the lack of support for intervals or chords. In addition, data augmentation of the song are required. The MIDI file contains only one way of playing a melody for the model to learn all the possibilities of this melody in the entire note definition domain. This not only requires the network to have higher capacity but also exacerbates the consumption of training resources.

At present, most of the research focuses on the performance of musical robots. Therefore, music generation is the "performance basis" of music robots. The intelligent generation of music by the music robot will improve the service level of the music robot. Music robots recognize more beats and styles, including some music style recognition and prediction systems based on spectrum information and feature extraction of the support vector machine (SVM) algorithm. In the past, most of the effort was devoted to featuring extraction. The quality of the features largely determines the result of classification and recognition. Although many features have been extracted to solve audio problems, direct features related to the structure are always difficult to describe. Today, the renewed development in the field of deep learning has made neural networks widely used in the field of audio processing. Therefore, this study uses deep learning technology to study the intelligent music generation of robots and strives to lay a certain foundation for the development of music robots.

2.2. Advantages and Features of MIDI. Although the MIDI format has many shortcomings in terms of neural network data set acquisition and data modeling, this study prefers symbol sequences that can accurately express information such as rhythm, timing, pitch, and velocity, rather than audio signals. Therefore, MIDI files are a more suitable choice as the input dataset for the network.

The standard midi format (SMF) of the standard MIDI format is composed of a series of chunk data blocks. Its header data block is a file, and immediately after the file header block is a track block that stores track information [16]. The parts included in the standard MIDI format file are shown in Figure 1.

2.3. Concepts Related to Deep Learning. Deep learning has further improved based on machine learning. Since the deep learning architecture contains more layers of networks,

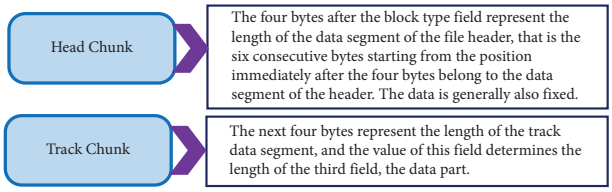


Figure 1: Standard MIDI format file.

more features can be obtained when analyzing features. This can improve the learning ability of the network. Compared with shallow networks, deep networks can use fewer neurons to perform the same function and be more efficient and accurate in the learning process. Deep learning has been used in various fields, such as finance, security, and manufacturing [17].

2.3.1. Neurons. Neurons are also called perceptrons. It is the most basic unit of neural networks. The neural network structure is shown in Figure 2.

In Figure 2, a complete neural network consists of an input layer, a hidden layer, and an output layer. The input layer is mainly used for input vectors. The hidden layer is used for vector analysis and parameter learning. The output layer is the output result. If the neural network has many hidden layers, it belongs to the category of deep learning [18]. Deep learning is mainly to study different DNNs, and then use DNNs to solve different problems.

The perceptron algorithm solves a lot of problems. The composition of a perceptron is shown in Figure 3.

The initial value of the weight W of the perceptron model is generally set randomly, which often fails to achieve a good fitting result. Therefore, it is necessary to calculate the output value and then make the difference between the actual output value and the theoretical output value to adjust each output. The learning rule is an algorithm used to calculate a new weight matrix W and a new bias b. The input layer input is a vector $(m_1, m_2, m_3, \cdots, m_n \mid m_i \in R)$. Each input corresponds to a weight w. Additionally, the other bias is b. The value of the offset is generally 1, denoted as w_0. The sum of weight values is calculated as shown in the following:

$$z = \sum_{i=1}^{n} w_i m_i + b. \tag{1}$$

The output value can be denoted as y. The value of y can be calculated by the activation function $g(z)$. There are many choices of activation functions, as shown in the following:

$$y = g(x). \tag{2}$$

2.3.2. Feedforward Neural Network. Multiple neurons are interconnected to form a neural network. The feedforward neural network (FNN), composed of a single layer of neurons, is shown in Figure 4.

In Figure 4, the single-layer neuron feedforward network has four input units and three neurons as output units. Each

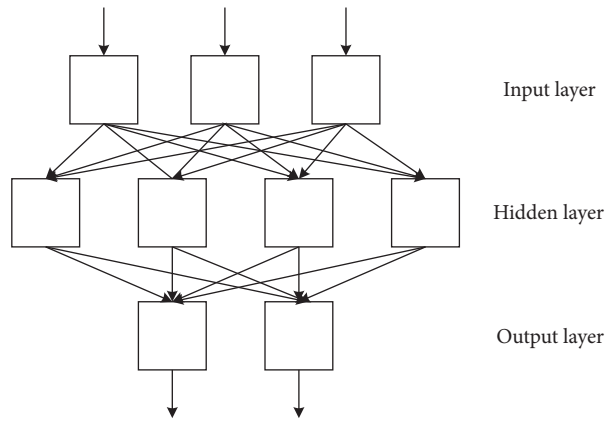


Figure 2: The structure diagram of the neural network.

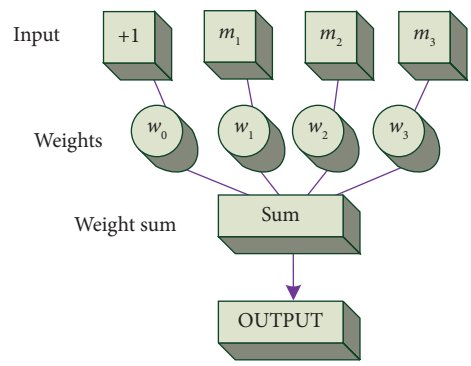


Figure 3: Perceptron diagram.

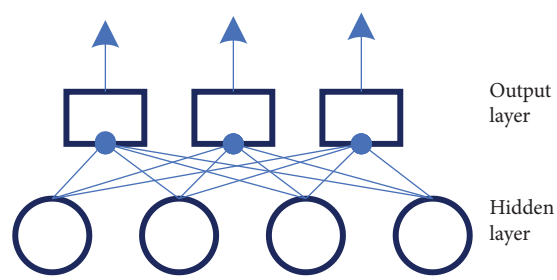


Figure 4: The network of single-layer neuron feedforward.

input unit feeds forward to the output layer of each neuron. This two-layer neuron structure is also called a perceptron. Another common network structure is a multilayer FNN. Its basic structure is like a single-layer feedforward network, but there are some additional hidden layers between the input and output layers. Neural networks can extract higher-order and global data based on these hidden layers. Neural network knowledge is acquired through the network's learning process on the dataset. Among them, the weights of the synapses are continuously adjusted in order to match the network output with the desired output. This approach is also known as supervised learning. The learning ability of multilayer networks is much stronger than that of single-layer perceptrons. Among them, the most commonly used method for training multilayer networks is the back-propagation algorithm. It is the most successful neural

network learning algorithm and can be used not only for multilayer FNN but also for other networks, such as RNN. The algorithm is based on the gradient descent strategy, and the difference between the expected value of the network and the actual output is sent back to the input, so the current learning situation of the network can be obtained. Usually, a loss function is used to express this criterion [19–21].

The most used loss function is the mean squared error. For the training data sample (x^n, t^n), assuming that the output of the neural network is $\bar{t}^{(n)}$, the mean square error of the neural network on this sample is calculated as shown in the following:

$$E_n = \frac{1}{2}\sum_{j=1}^{l}\left(\bar{t}_j^n - t_j^n\right)^2. \tag{3}$$

2.3.3. Gradient Descent. The gradient descent algorithm is relatively efficient. By using this algorithm, it is relatively easy to obtain the optimal solution for the function that needs to be trained, thereby improving the accuracy of the model. The method can be divided into three categories, namely batch gradient descent, mini-batch gradient descent, and stochastic gradient descent [22].

When this function is derivable, by calculating the inverse of the function, the optimal solution θ of the training function is regarded as a variable, and the optimal solution can be obtained. Additionally, the calculation of the weights also needs to be adjusted according to the learning rate η. The update of θ is shown in the following:

$$\theta = \theta - \eta\nabla_\theta J(\theta). \tag{4}$$

In the actual application process, the update of θ takes a long time. Therefore, in the process of training with large batches of data, it is difficult to complete the training online. For larger datasets, the training process may be harder because more memory is required.

One disadvantage of the stochastic gradient descent algorithm (SGDA) is that it is only suitable for training datasets. Each set of training data can calculate the gradient. For example, there is a set of training data $(x^i; y^i)$, the value of i belongs to $(0, n)$, and n represents the size of the entire data set. The update of θ is shown in the following:

$$\theta = \theta - \eta\nabla_\theta J\left(\theta; x^i; y^i\right). \tag{5}$$

Since this method is aimed at the training dataset, and only one sample is randomly selected to update the parameters at a time, the learning efficiency of this method is very high, and it can be updated online. The only downside is that there will be optimization fluctuations due to updates not going in the right direction. In general, the advantages of SGDA are still obvious, and it is relatively easy to find better local minimum points.

The mini-batch gradient descent algorithm is derived from the two algorithms mentioned above. The algorithm needs to find a balance point in the process of updating, and each update will select fewer samples than n from the training set. θ is updated as shown in the following:

$$\theta = \theta - \eta\nabla_\theta J\left(\theta; x^{(i:\ i+n)}; y^{(i:\ i+n)}\right). \tag{6}$$

Compared with the batch gradient descent algorithm, the mini-batch gradient descent algorithm greatly improves the learning efficiency of the mini-batch algorithm. It does not require a large amount of memory. It is also more efficient when performing matrix operations. Therefore, the mini-batch gradient algorithm is one of the most used algorithms in neural networks.

2.3.4. LSTM. LSTM is a special type of RNN. Its purpose is to solve the long-term dependency problem of traditional RNN. It is designed to avoid the rapid decay of back-propagated errors [23]. The storage unit structure of the LSTM network is shown in Figure 5.

In Figure 5, a single storage unit in a conventional RNN neuron is replaced with a storage block containing multiple storage units. They can pass memory cell values down multiple time steps along the time axis and can be memorized or forgotten at each time beat. It can capture the information and dependencies of long-distance steps, which is a very necessary function for the extraction of abstract musical features. Therefore, it is a more appropriate choice for the LSTM network to be applied to the generation of musical melodies [24].

The current state of a memory cell depends on its previous state, the network itself, the forget gate, and the input value of the input gate. The neuron state update is shown in the following:

$$S_c = S_c y^\varphi + g\left(\mathrm{net}_c\right) y^{\mathrm{in}}. \tag{7}$$

After the input vector enters the neuron, through the activation function, g squeezes the input value into the interval of 0 to 1 and then multiplies the input value obtained by the input gate. The process is shown in the following:

$$y^{\mathrm{in}} = \sigma\left(\mathrm{net}_{\mathrm{in}}\right). \tag{8}$$

Among them, σ represents the sigmoid activation function. When y^{in} is 0, by multiplying $g(\mathrm{net}_c)$ and y^{in}, the input gate can prevent the network input $\mathrm{net}_{\mathrm{in}}$ from updating the neuron storage unit.

Like in the calculation process of y^{in}, the output value y^c of the neuron is the storage unit state s_c. After being squeezed through the activation function h, it is multiplied by the value y^{out} of the output gate, as shown in the following:

$$y^{\mathrm{out}} = \sigma\left(\mathrm{net}_{\mathrm{out}}\right), \tag{9}$$

$$y^c = h\left(s_c\right) y^{\mathrm{out}}. \tag{10}$$

2.4. Implementation and Training of a Music Generation Model Based on Deep Learning

2.4.1. Training and Testing Data. The MIDI music website selects relevant data (musescore.com) and selects ten genres.

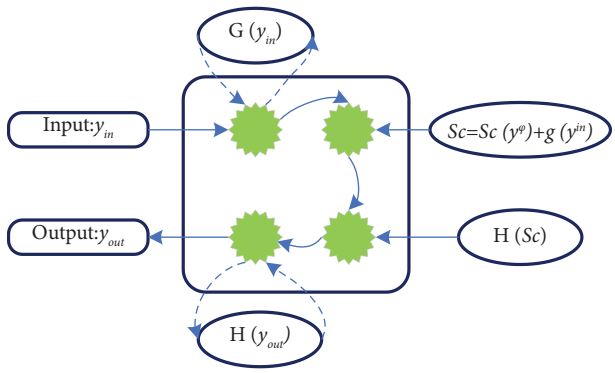


Figure 5: Storage unit structure of LSTM network.

One hundred music clips are selected for each music genre, for a total of 1000 music clips. The time of each song is different, generally between the 30 s and 300 s. All data are in MIDI music format. In the experiment, five groups of music were selected for analysis. The number of training and test sets for all groups is 180 and 20, respectively. The amount of data per genre is 200.

2.4.2. Network Training and Optimization. The neural network uses Python 3.10.6 to write the entire system, uses Tensorflow to implement the programming language framework, and uses Music21 for music classification. Training of the network: after all model parameters are determined, the network is trained. In the training process, the training data are used, and the updated parameters are selected repeatedly to optimize the model performance [25]. During the training process, the values of some hyper-parameters in the model are set, such as the number of iterations and the number of network layers and hidden layer units. Figure 6 shows the training process of the music generation algorithm.

In Figure 6, after the network is trained, it can generate music of different genres. The music generation network generates a new note, mainly by combining the previous input note with the current input note and then making predictions [26]. Here, all notes are converted into vectors, which are used to compare and predict. The detailed flow chart of music sequence generation is shown in Figure 7.

In Figure 7, first, the trained network and the set parameters are loaded. N represents the length of the sequence. If the sequence length is greater than 0, forward propagation is performed, and the value of the loss function is calculated. Here, there is a threshold. If the value of the loss function is already smaller than the threshold, it means that the music feature vector is valid at this moment, and the next moment's sequence value is predicted. Finally, the entire sequence of predicted music is output.

(1) Network Optimization. During training, when the dataset is small, a common problem is overfitting. The generalization ability is poor, and the consequence is that the effect of the model is not good and the accuracy is not enough. In order to prevent the overfitting problem and improve the accuracy of the network, the experiment adopts the dropout method. A dropout refers to randomly ignoring the weights of some hidden layer nodes when training a neural network [27]. During training, these nodes will not work. Their weights are also not updated. Dropout is added between hidden layers, as shown in the following:

$$r = (1 - p)xf\left(W_v + b\right). \tag{11}$$

Here $1 - p$ is a binary model that follows the Bernoulli distribution. When the probability value is p, the value is 1, and the rest is 0. In the experiment, all parameters are multiplied by p to achieve the purpose of changing the parameters [28, 29].

3. Experimental Results

3.1. Network Optimization Results. In this experiment, in order to optimize the performance of the model, the convergence effect of the loss function in the three cases of $p = 0.5$, $p = 0.6$, and $p = 0.7$ is analyzed separately. The convergence effect of the loss function is shown in Figure 8.

In Figure 8, when the probability value is $p = 0.5$, the convergence value is 0.0035, which is larger than when $p = 0.7$. At $p = 0.6$, the experimental effect is better than at $p = 0.5$ but not as good as at $p = 0.7$. Therefore, the experimental results are best when the probability value is $p = 0.7$. Finally, the dropout coefficient value of this model is determined to be 0.7.

3.2. Analysis of Algorithm Influencing Factors. According to the experience of neural network training and as the number of iterations of the LSTM network increases, the experimental error will become smaller and smaller [30–32]. The data means that the closer the actual output value is to the target value, the closer the training result is to the target value, as shown in Figure 9.

In Figures 9(a) and 9(b), after 3000 iterations, the frequency distribution of the generated music sequence is roughly the same as the original music sequence frequency distribution, but there are still some obvious differences. For example, the generated music sequence spectrogram contains many frequencies, not in the original music sequence spectrum. In Figure 9(c), after 6000

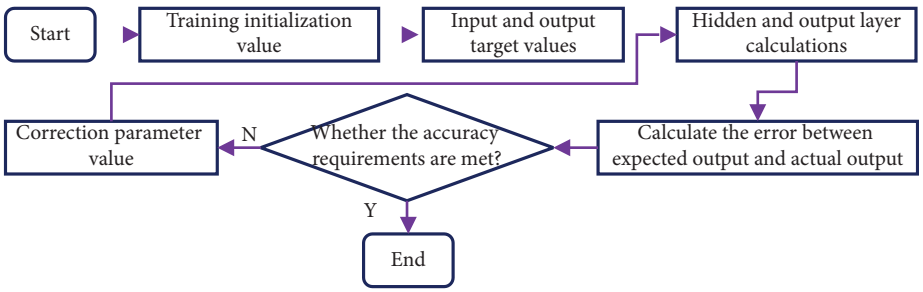


Figure 6: Flow chart of music generation training.

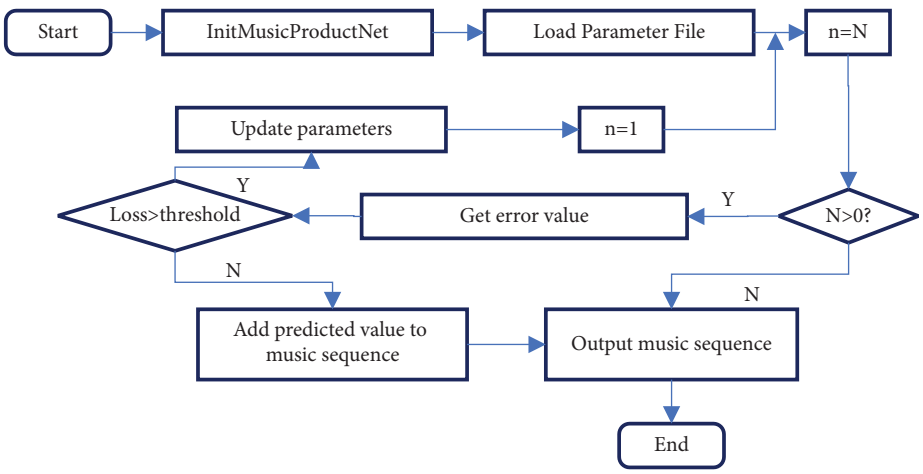


Figure 7: Flow chart of music sequence generation.

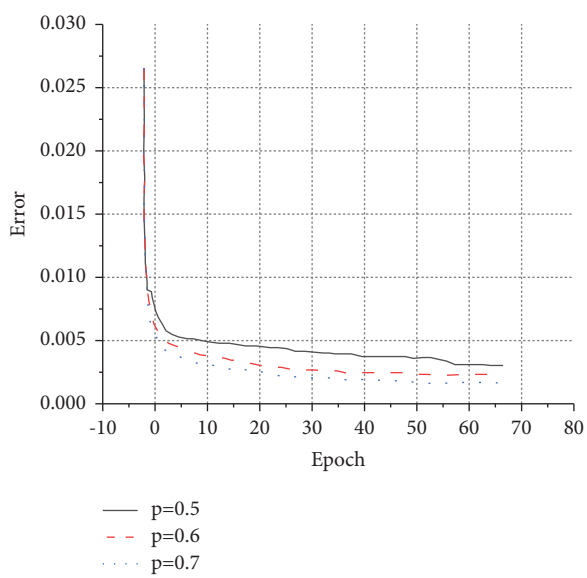


Figure 8: Function convergence diagram under different p values.

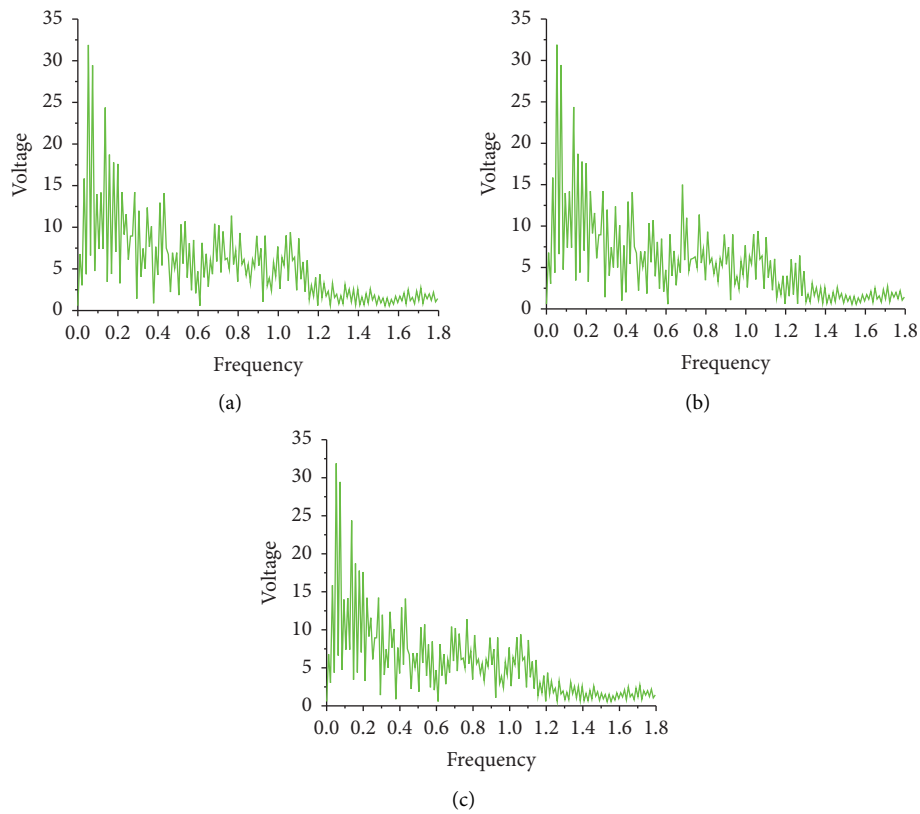


Figure 9: Music spectrogram under different iterations. (a): Spectrogram of original sample music sequence; (b): spectrogram of music sequence iterated 3000 times; and (c): Ssectrogram of music sequence iterated 6000 times.

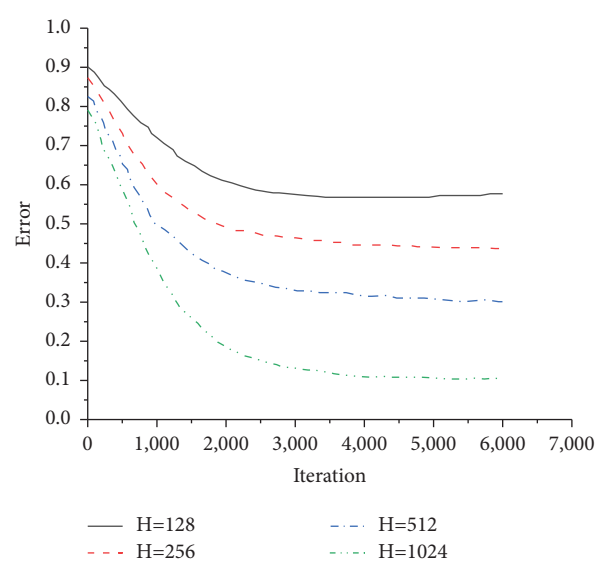


Figure 10: The effect of the number of neurons in the hidden layer on the error.

iterations, the generated music spectrogram is completely consistent with the original music spectrogram. The main reason is that with the increase in the number of iterations, the model parameters are also updated many times. Finally, the parameters of the model are optimized.

The effect of the number of neurons in the hidden layer of the neural network on the experiment is analyzed. The influence of the hidden layer neurons on the experimental error is shown in Figure 10.

In Figure 10, when the number of neurons in each layer is 128, 256, 512, and 1024, in turn, the training results become more and more accurate, and the error value becomes smaller and smaller. If the number of neurons is too large, the computer environment is very demanding, and the

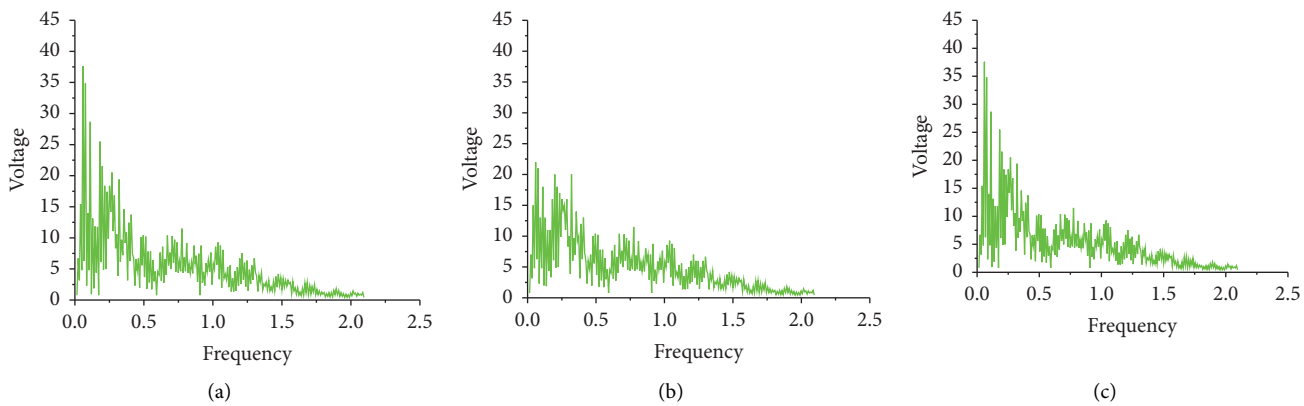


Figure 11: Spectrogram of music with different hidden layers. (a): Spectrogram of sample music sequence; (b): spectrogram of music when the hidden layer is 2; and (c): spectrogram of music when the hidden layer is 4.

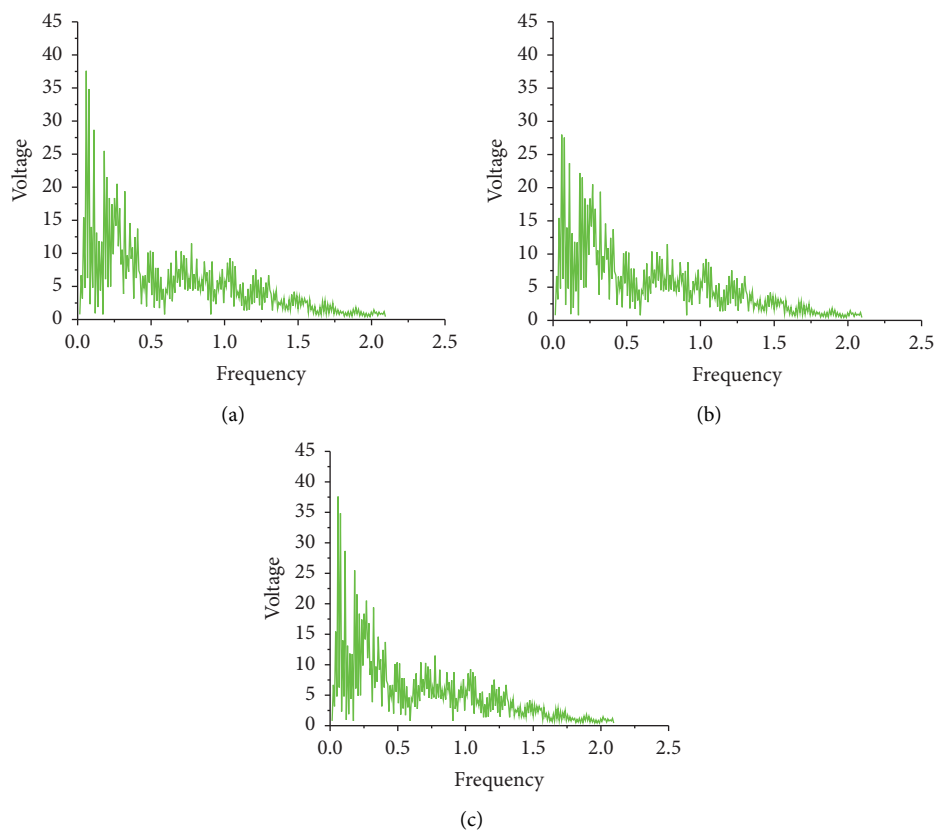


Figure 12: Music spectrograms of the two algorithms. (a): Original music spectrogram; (b): RBM-generated music spectrogram; and (c): LSTM-generated music spectrogram.

training time will increase geometrically, increasing the complexity by several degrees. Therefore, the number of neurons in the network's hidden layer is 1024, which can make the training result optimal.

During the experiment, the spectrum analysis is performed on the music sequences generated by the model under different hidden layers in turn. The generated music spectrograms and sample spectrograms are shown in Figure 11.

In Figure 11, the effect of LSTM on music analysis is still obvious. With the increased number of neural network layers, the trained music spectrogram is getting closer and closer to the original spectrogram, indicating that its accuracy is getting higher and higher. When there are two hidden layers, some frequencies do not appear. When the hidden layer has four layers, the difference between the generated music sequence and the original music sequence is very small, which shows that the

generated music is the most accurate when there are four hidden layers.

The LSTM-based music generation algorithm is compared with the traditional RBM music generation algorithm, and the spectrograms of the music generated by the two generative models are analyzed separately. The results of the music spectrogram generated by the two methods are shown in Figure 12.

In Figure 12, the trend of the spectrograms of the music generated by the traditional RBM and the original music is highly similar. However, RBM-generated music is not as accurate as LSTM-generated music. In Figure 12(a), when the frequency of the original sample music is around 1000–2000 Hz, the voltage value exceeds 35, and Figure 12(b) does not exceed. Compared with Figure 12(c), the music spectrogram generated by LSTM is more consistent with the original music spectrogram in both the overall frequency distribution and the sample frequency distribution.

4. Conclusion

This experiment mainly analyzes the experimental effect of deep learning in piano music generation under robotics technology. The music sequence spectrograms are analyzed to illustrate the accuracy and rationality of the algorithm. Firstly, based on introducing the basic concepts of deep learning, the advantages of the LSTM network in music generation are introduced. Meanwhile, dropout coefficients are used to optimize the neural network. Through experimental verification, the dropout coefficient value is 0.7. Secondly, this experiment analyzes the experimental effect of the algorithm, including the influence of the number of iterations and neurons in the hidden layer, on the effect of music generation. When the number of iterations is 6000, the error between the generated music sequence and the original music is the smallest. When the number of hidden layers of the network is set to 4 and the number of neurons in the hidden layer is set to 1024, the training results of the network are best. Spectrograms of sequences generated by music generation algorithms based on LSTM and traditional RBM show that neural networks perform well in music generation. The shortcomings of this study that can be improved in the future are (1) due to limited energy, the music training dataset selected is small, and the music styles are similar. The internal structure and logic of different styles of music are not similar. If it is mixed together to generate music, the accuracy will be much less. Future research will focus on how the more complex data can be separated from other multitrack data. The processing of the data of these tracks enables the neural network to process and analyze the multitrack data. (2) This experiment mainly extracts digital features of music, including pitch, timbre, and loudness. Other features can be extracted later, such as energy features and time domain features. Rich data features lead to better results. (3) The algorithm model finally generates a matrix containing music features, which also needs to be converted into playable music. Future research could be considered to include how to automatically generate music without the need to reverse the process.

Data Availability

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethical Approval

This article does not contain any studies with human participants or animals performed by any of the authors.

Consent

Informed consent was obtained from all individual participants included in the study.

Conflicts of Interest

All authors declare that they have no conflicts of interest.

Acknowledgments

The authors acknowledge the help from the university colleagues.

References

[1] B. Sisman, J. Yamagishi, S. King, and H. Li, "An overview of voice conversion and its challenges: from statistical modeling to deep learning," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 29, pp. 132–157, 2021.

[2] H. Talebi and P. Milanfar, "NIMA: neural image assessment," *IEEE Transactions on Image Processing*, vol. 27, no. 8, pp. 3998–4011, 2018.

[3] D. R. Rizvi, I. Nissar, S. Masood, and M. Ahmed, "An LSTM based Deep learning model for voice-based detection of Parkinson's disease," *Int. J. Adv. Sci. Technol*, vol. 29, no. 8, 2020.

[4] G. Demir, A. Çekmiş, V. B. Yeşilkaynak, and G. Unal, "Detecting visual design principles in art and architecture through deep convolutional neural networks," *Automation in Construction*, vol. 130, Article ID 103826, 2021.

[5] V. Jabade, V. Deshpande, and K. Aditya, "Music generation and song popularity prediction using artificial intelligence - an overview," *International Journal of Computer Application*, vol. 182, no. 50, pp. 33–39, 2019.

[6] J. Pérez-Marcos, D. M. Jiménez-Bravo, J. F. De Paz, G. Villarrubia, V. F. Lopez, and A. B. Gil, "Multi-agent system application for music features extraction, meta-classification and context analysis," *Knowledge and Information Systems*, vol. 62, no. 1, pp. 401–422, 2020.

[7] B. Ru, D. Li, Y. Hu, and L. Yao, "Serendipity—a machine-learning application for mining serendipitous drug usage from social media," *IEEE Transactions on NanoBioscience*, vol. 18, no. 3, pp. 324–334, 2019.

[8] J. P. Briot and F. Pachet, "Deep learning for music generation: challenges and directions," *Neural Computing & Applications*, vol. 32, no. 4, pp. 981–993, 2020.

[9] J. P. Briot, "From artificial neural networks to deep learning for music generation: history, concepts and trends," *Neural Computing & Applications*, vol. 33, no. 1, pp. 39–65, 2020.

[10] S. Y. Li and Y. Sung, "INCO-GAN: variable-length music generation method based on inception model-based conditional GAN," *Mathematics*, vol. 9, no. 4, p. 387, 2021.

[11] J. T. Xie, "A novel method of music generation based on three different recurrent neural networks," *Journal of Physics: Conference Series*, vol. 1549, no. 4, Article ID 042034, 2020.

[12] S. Hizlisoy, S. Yildirim, and Z. Tufekci, "Music emotion recognition using convolutional long short term memory deep neural networks," *Engineering Science and Technology, an International Journal*, vol. 24, no. 3, pp. 760–767, 2021.

[13] J. Li, L. Han, X. Li, J. Zhu, B. Yuan, and Z. Gou, "An evaluation of deep neural network models for music classification using spectrograms," *Multimedia Tools and Applications*, vol. 81, no. 4, pp. 4621–4647, 2022.

[14] D. C. Wu, C. Y. Hsiang, and M. Y. Chen, "Steganography via MIDI files by adjusting velocities of musical note sequences with monotonically non-increasing or non-decreasing pitches," *IEEE Access*, vol. 7, no. 1, Article ID 154056, 2019.

[15] Y. H. Liu and D. C. Wu, "A high-capacity performance-preserving blind technique for reversible information hiding via MIDI files using delta times," *Multimedia Tools and Applications*, vol. 79, no. 25-26, Article ID 17281, 2020.

[16] C. Gibson, "A sound track to ecological crisis: tracing guitars all the way back to the tree – ERRATUM," *Popular Music*, vol. 38, no. 3, p. 588, 2019.

[17] Y. Kim, H. J. Lee, and J. Shim, "Developing data-conscious deep learning models for product classification," *Applied Sciences*, vol. 11, no. 12, p. 5694, 2021.

[18] W. Li and W. Bian, "Projection neural network for a class of sparse regression problems with cardinality penalty," *Neurocomputing*, vol. 431, no. 1, pp. 188–200, 2021.

[19] J. Zuo, F. Feng, and Y. He, "Research and application of train online health status detection based on feedforward neural network," *Journal of Physics: Conference Series*, vol. 1828, no. 1, Article ID 012034, 2021.

[20] B. Li and X. Y. Zhuang, "Multiscale computation on feedforward neural network and recurrent neural network," *Frontiers of Structural and Civil Engineering*, vol. 14, no. 6, pp. 1285–1298, 2021.

[21] Y. Xiong and X. Tong, "Convergence of batch gradient method based on the entropy error function for feedforward neural networks," *Neural Processing Letters*, vol. 52, no. 3, pp. 2687–2695, 2020.

[22] T. Adachi, N. Hayashi, and S. Takai, "Distributed gradient descent method with edge-based event-driven communication for non-convex optimization," *IET Control Theory & Applications*, vol. 15, no. 12, pp. 1588–1598, 2021.

[23] T. Li and Y. Guan, "Dual memory LSTM with dual attention neural network for spatiotemporal prediction," *Sensors*, vol. 21, no. 12, p. 4248, 2021.

[24] M. Dua, R. Yadav, D. Mamgai, and S. Brodiya, "An improved RNN-LSTM based novel approach for sheet music generation," *Procedia Computer Science*, vol. 171, no. C, pp. 465–474, 2020.

[25] Z. Hu, M. Ivashchenko, L. Lyushenko, and D. Klyushnyk, "Artificial neural network training criterion formulation using error continuous domain," *International Journal of Modern Education and Computer Science*, vol. 13, no. 3, pp. 13–22, 2021.

[26] P. Q. Pfordresher and K. Chow, "A cost of musical training? Sensorimotor flexibility in musical sequence learning," *Psychonomic Bulletin & Review*, vol. 26, no. 3, pp. 967–973, 2019.

[27] I. D. Colley, M. Varlet, J. MacRitchie, and P. E. Keller, "The influence of visual cues on temporal anticipation and movement synchronization with musical sequences," *Acta Psychologica*, vol. 191, no. 1, pp. 190–200, 2018.

[28] Y. Chen and Z. Yi, "Adaptive sparse dropout: learning the certainty and uncertainty in deep neural networks," *Neurocomputing*, vol. 450, no. 1, pp. 354–361, 2021.

[29] H. Liu, "Design of neural network model for cross-media audio and video score recognition based on convolutional neural network model," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 4626867, 12 pages, 2022.

[30] G. Chen, P. Chen, W. Huang, and J. Zhai, "Continuance intention mechanism of middle school student users on online learning platform based on qualitative comparative analysis method," *Mathematical Problems in Engineering*, vol. 2022, Article ID 3215337, 12 pages, 2022.

[31] W. Zheng, X. Tian, B. Yang et al., "A few shot classification methods based on multiscale relational networks," *Applied Sciences*, vol. 12, no. 8, p. 4059, 2022.

[32] Y. Zhang, X. Shi, H. Zhang, Y. Cao, and V. Terzija, "Review on deep learning applications in frequency analysis and control of modern power system," *International Journal of Electrical Power & Energy Systems*, vol. 136, Article ID 107744, 2021.

Hindawi
Computational Intelligence and Neuroscience
Volume 2022, Article ID 2529372, 11 pages
https://doi.org/10.1155/2022/2529372

Research Article

A Dynamic Prediction Neural Network Model of Cross-Border e-Commerce Sales for Virtual Community Knowledge Sharing

Li Tian and Xiumei Wang

Business School, Nanfang College Guangzhou, Guangzhou 510970, China

Correspondence should be addressed to Li Tian; tianl@nfu.edu.cn

Received 25 July 2022; Revised 9 September 2022; Accepted 20 September 2022; Published 10 October 2022

Academic Editor: Ning Cao

In this paper, a neural network algorithm is used to conduct in-depth research and analysis on the sales dynamics prediction of virtual community knowledge sharing in cross-border e-commerce. Both the expected returns and the social network structure are analyzed, and both have positive effects on knowledge sharing in the actual development process, but the degree of them also possesses certain variability. A model of the factors influencing the quality of knowledge shared by users is constructed to explore the relationship between the dimensions in social capital and how they affect the community users' perceptions of knowledge quality. Exploring the strong influencing factors of product repurchase rate has key implications for promoting product sales and sales forecasting. The scale of this paper has undergone several minor revisions, and the content validity is very good. Criterion validity is generally reflected by Person correlation coefficient, and construct validity includes exploratory factor analysis and confirmatory factor analysis. First, the values were clustered, and the optimal variables were selected using stepwise regression and fitted with Poisson regression models to explore the relationship between the repurchase rate of different products and the factors that strongly influence the repurchase rate in the case. To predict the sales of goods, the advantages of the BP neural network, LSTM neural network, and Verhulst gray model are combined: BP neural network can predict sales by combining the current data corresponding to independent variables, LSTM neural network can explore the influence of historical data on sales, and Verhulst model can predict sales based on the growth trend of variables. A BP-LSTM-Verhulst nested neural network based on the AP algorithm is constructed to predict sales volume, and the accuracy of this method is proved by example. Finally, it is found that the proposed sales prediction method has higher accuracy than exponential regression and shallow neural networks. The deep learning prediction method combining unstructured data such as images proposed in the paper not only provides a more accurate sales prediction method for short life cycle products in e-commerce but also provides an effective deep learning method for management practices. The KMO value obtained after the test is less than 0.5, which is not suitable for factor analysis. Using SPSS22.0 to expand the KMO value and Bartlett's sphericity test, the KMO value is significantly greater than the minimum requirement of 0.5, the KMO value of self-efficacy is 0.826, and the KMO value of expected return is 0.870.

1. Introduction

In today's world, knowledge is gradually becoming an important force for enterprises to improve the core competitiveness of talents and continuously gain competitive advantages, and it plays an irreplaceable role in the sustainable and healthy development of organizations [1]. In recent years, knowledge management has gradually been paid attention to and studied by scholars in the field of management and has gradually become the focus of research. Knowledge has the characteristic of shared growth, and knowledge sharing refers to the stage in which knowledge providers in an organization summarize their knowledge and then transform it by using corresponding tools and channels, and finally knowledge gainers internalize and absorb it, which is a process of knowledge learning itself. During the development of an organization, knowledge sharing can effectively enhance the innovation ability of the organization and thus strengthen the core competitiveness of the organization and individuals. Therefore, a new model

structure of the recurrent neural network is introduced, which gives the neural network the function of memory and can obtain the influence of historical information on the current prediction. Contemporary Internet technology has shown vigorous development, and e-commerce is its representative product, which is a brand-new development direction of network technology application. Merchandise mix and sales forecasting have always been the core of e-commerce and an important prerequisite for the healthy development of enterprises and have a positive impact on market orientation and inventory control [2]. In recent years, the competition among e-commerce enterprises has become increasingly fierce. To occupy a favorable position in the market, it is necessary to make accurate predictions and judgments on the competitiveness of goods, and reasonable predictions will bring great economic benefits and maintain the stable development of the market, among which the research of effective methods of merchandise mix and sales forecasting is an urgent problem for current e-commerce enterprises.

With the enhancement of computer functions and the emergence of more machine learning techniques and statistical forecasting methods, the field of demand forecasting is booming. People are consuming increasingly, and the level of material consumption is gradually increasing, leading to an increase in the frequency of consumption and prompting a gradual shortening of product life cycles [3]. For example, the split seasons of clothing in fast fashion products have evolved from the traditional four seasons to multiple seasons, and the average shelf life of fashion apparel, satchels, shoes, and hats produced by fast fashion product companies is often only a few weeks. Short life cycle products cannot be active in the market for a long time and are easily replaced by homogeneous competitive products, showing the characteristics of a short sales life cycle and being easy to depreciate. For some fast fashion products, such as shoes and hats and apparel fashion, because the product life cycle is relatively short, its demand forecast is more challenging. Forecasting methods look for trends or seasonal patterns in historical sales data and sometimes correlate them with events from other sources. However, when a new product is introduced for launch, there is little historical data available other than known information such as attributes [4]. This study discusses fast fashion products with short life cycles and focuses on retail-oriented companies that forecast the demand as well as sales of fast fashion products. The most important product characteristics of fast fashion products are fast feedback on showroom design, short product life cycle, and higher frequency of updating product categories and types. It is mainly for products that are mainly sold in offline stores. Fast fashion products as an important part of the manufacturing and retail industry not only represent the soft power of the country but also bring important support and influence to the stability and development of the national economy.

In past studies, there is no lack of research on product sales forecasting methods, and the goal of scholars mainly focuses on how to use time series methods to discover the trend of sales data from historical sales data, to predict the trend of future sales. However, time series methods have high requirements for the smoothness of the data. Each consumer in the market is supposed to be random and disorderly; however, time series methods will fail when some factors appear to make the whole or part of consumers appear to have a certain purchasing tendency; for example, promotional tools and other factors that affect users' purchasing decisions appear. Each user's purchase decision behavior ultimately constitutes the overall sales volume, but in the traditional sales forecasting research methods, due to the difficulty of obtaining individual purchase decision influencing factors, therefore, individual purchase decision influencing factors are rarely used to study sales forecasting. Currently, in the Internet environment, users' browsing and purchasing processes are recorded in full, and companies can search for factors influencing users' purchasing decisions based on their behavioral records and use the factors influencing users' purchasing decisions to make product sales forecasts.

2. Related Works

For quantitative methods, companies tend to use time series methods for forecasting. Time series methods are operable for forecasting sales, but they lack consideration of factors that affect product sales. Many scholars have introduced factors affecting users' purchasing behavior as influencing factors of sales and have achieved better results [5]. In the study, Thomasson predicted the sales of apparel companies for one-quarter, but it is difficult to guide the ordering behavior of companies in the medium and long term compared to the short life cycle of apparel. Mumu and other scholars proposed a crown model for the short-term sales problem of online agricultural products, combined with deep learning algorithms, and used simulated data to analyze and obtain better results [6]. The study uses the self-coding function of neural networks to encode and reduce the dimensionality of category variables such as shipping place, which effectively reduces the dimensionality of high-dimensional data and uncovers the correlation of different categories [7]. The stacked autoencoding neural network can learn in real time and improve the shortcomings of random initialization of the neural network. The study collected egg sales data from Taobao and considered the influence of information such as discounts, packages, and other promotional tools on users' purchasing behavior. The study shows that the accuracy of phased sales prediction is higher than that of direct prediction using neural networks [8].

Cao et al. designed a hybrid learning algorithm based on multi-intelligent body theory and social networks to analyze the knowledge sharing of virtual community users with the help of NetLogo 5.0 [9]. Pananond et al. propose the demand for personalized services in information-sharing virtual spaces and the demand knowledge base construction based on multiagent technology and discuss the characteristics and implementation scheme of agent technology [10]. On the Internet, there is often a large amount of information, which makes it more difficult for network users to search for the information they need and consumes the energy of

community users to search for information, and the "relationship" chain communication mode of SNS provides a new way for people to obtain information. With its unique advantages of communication and knowledge sharing among community members, social network services have attracted extensive attention and research from a wide range of scholars [11]. He et al. put forward the concept of knowledge sharing itself, which pointed out that, among knowledge owners and demanders, information communication is used to help them complete the learning of knowledge, and the learning content mainly covers the knowledge and experience possessed during the learning period [12]. In the form of knowledge sharing, the process clarifies the importance of the channel itself, while pointing out that the use of certain tools can enhance the effectiveness of knowledge sharing.

When there is no consumer shopping, the vending machine must be connected to the cloud platform so that the cloud platform can read the vending machine's operational status data at any time. When consumers are shopping, it is important to ensure that the vending machine is connected to the cloud platform stably, as a failure of the connection can directly result in a failed order transaction. In addition, the order data transfer must be accurate, so a data verification service must be completed. From single models such as time series and artificial intelligence prediction methods to combined models, the empirical analysis proves that combined models can combine the advantages of single models and show high accuracy, and there is some inspiration for the method used in this paper. The predecessors have provided us with many valuable research methods, but the complexity of commodity forecasting itself predestines that there is more information value waiting for us to explore.

3. Analysis of the Factors Influencing Knowledge Sharing in Virtual Communities

The benefit-oriented rapid relationship includes three dimensions of mutual understanding, mutual benefit, and relationship harmony, among which, mutual benefit means that both parties feel that they will get positive benefits from establishing the relationship and can satisfy their respective interests, and this rapid relationship will prompt both parties to engage in continuous interactive behaviors. Relational commitment, interaction, and reciprocity play an important role in knowledge exchange, and long-term sustained behaviors are generated in relational virtual communities because they receive personal benefits during community participation and different experiential components influence sustained usage behaviors through perceived personal benefits. Knowledge sharing behaviors of users in virtual communities can improve their image and rank, as well as receiving favorable comments from other users, and psychological satisfaction and pleasure, and they are motivated to maintain long-term relationships to continuously obtain such benefits [13]. When the prediction generates a residual value, it starts to learn the residual value. There may be weak and long-term memory between the residual values. The corresponding accumulated residual series may not be stationary. Users sense that they will gain benefits from building relationships, and when the perceived benefits are strong, they are willing to maintain long-lasting relationships, which is an important motivation for sustained behavior. The relationships among users in academic virtual communities meet the definition of fast relationships and can explain users' willingness to sustain knowledge sharing.

About the content of the knowledge transformation perspective, some studies point out that knowledge sharing can occur not only in individuals and organizations but also corresponding to the mutual transformation process in both explicit and tacit knowledge, and the key to the analysis of this transformation process, which is also called the SECI model, is the internalization process. Tacit knowledge is highly personal and therefore cannot be easily grasped or expressed by others. Therefore, the externalization of the knowledge owner's knowledge behavior plays a key role in the process of knowledge sharing. The virtual community social network realizes the sharing of information exchange across regions. The network of interpersonal relationships is a real opportunity to facilitate or limit knowledge sharing behavior, which makes it possible for individuals to shift from no relationships to weak or even strong relationships [14]. Most existing studies are based on a social network theory perspective, and they argue that network size, for example, affects knowledge sharing behavior to some extent. Knowledge sharing behaviors start to arise when knowledge owners within an organization perceive that the organization has a high level of organizational trust; on the contrary, when there is a trust crisis within the organization, the quantity and quality level of knowledge sharing becomes seriously problematic. Therefore, the social network structure and organizational trust are the main choices used in this study in terms of opportunity elements, as shown in Figure 1. The other layers are very poor. To avoid training local optimality, pretraining and fine-tuning are adopted: training a network structure with only one hidden layer can learn the first-order features of the input residual data points.

Social network theory states that the actions of the subject depend to a large extent on the position of the community individuals in the structural system of social relations. The social network structure focuses on the analysis in terms of perceived centrality and other aspects, relying on the specific network association patterns to provide opportunities or restrictions to influence individual behavior. In virtual community social networks, members act as participating nodes with relatively fixed corresponding network sizes, so this paper chooses interactive relationships to achieve a replacement of network strength and network density, focusing on the fact that physical network technologies do not affect knowledge sharing.

Interaction relationship is mainly the frequency and depth of interaction among members in the virtual community social network, perceived centrality is the importance of members' status, and equivalence is the equal status of members in the virtual community social network. Based on Gravette's measurement scale, the scale was adjusted to

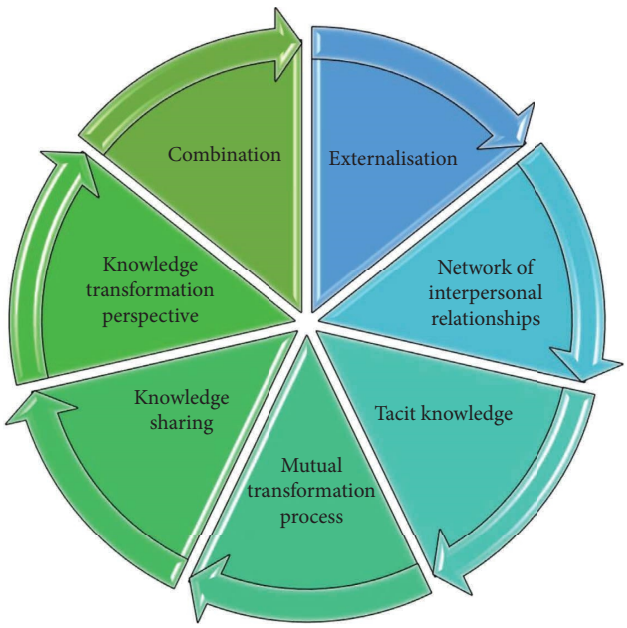


Figure 1: Knowledge sharing model of virtual community.

include 4 items, 3 items for the perceived centrality dimension, and 2 items for the equivalence dimension.

The validity test is specifically a test of the validity of the questionnaire to determine whether the designed items are reasonable and whether they can effectively reflect the researcher's research objectives. Validity can be divided into three categories: content validity, criterion validity, and structural validity. Content validity is to justify the design of the questionnaire from various perspectives, and validation factor analysis also includes convergent validity and discriminant validity [15]. This layer of network structure after reaching the optimization starts training the next one. The content validity of this paper is very good after several minor modifications. The validity of the scale is generally reflected by the Person correlation coefficient, and the structural validity includes exploratory factor analysis and validation factor analysis. Factor analysis can be a good way to test the structural validity. First, to determine whether the scale is suitable for factor analysis, the KMO and Bartlett's spherical test are used, which can largely show the distribution of the data, and the KMO can largely show the degree of bias correlation among the variables.

The KMO values obtained after the test are close to 1, in which case Bartlett's spherical test is judged to be significant and the factor analysis is very appropriate, and the KMO values obtained after the test are less than 0.5 and the factor analysis is not appropriate. Applying SPSS22.0 to develop KMO values and Bartlett's spherical test, the KMO values were significantly greater than the minimum requirement of 0.5, including KMO values of 0.826 for self-efficacy, 0.870 for expected reward, 0.848 for social network structure, 0.847 for organizational trust, and 0.840 for knowledge sharing. Bartlett's spherical test was significant and factor analysis was very appropriate as shown in Table 1. Content validity is to demonstrate the rationality of questionnaire design from various perspectives, and confirmatory factor analysis also includes convergent validity and discriminant validity.

Table 1: KMO and Bartlett's spherical test of variables.

Name	KMO	Bartlett's sphericity test		
		Approximate chi-square	df	Sig.
Self-efficacy	0.192	0.237	0.117	0.255
Expected return	0.221	0.35	0.222	0.304
Organizational trust	0.367	0.277	0.327	0.168
Social network structure	0.27	0.23	0.194	0.389
Knowledge sharing	0.264	0.377	0.112	0.142
General table	0.208	0.292	0.317	0.361

Among the above sales forecasting research methods, the traditional sales forecasting methods based on time series methods have a relatively simple calculation process, but the forecasting effect and the scope of application are relatively limited. Many scholars began to use machine learning methods to study sales forecasting, which improved the fitting ability of the forecasting model, and the combined forecasting model can further improve the forecasting effect. Many scholars have shown that there are factors that can be used for sales forecasting in addition to sales volume itself [16]. In addition to considering the change in historical sales, this topic also uses other marketing information affecting sales as independent variables for sales prediction, especially unstructured data such as pictures, which have been proven by many scholars' studies to be the most important marketing information online and have a huge impact on users' purchasing decision behavior.

4. Neural Networks for Cross-Border e-Commerce Sales Dynamic Prediction Model Design

BP neural network is essentially also a single-layer perceptron, which is also a multilayer feedforward neural network obtained by error backpropagation [17]. In this way, the influence of the early data in the sequence data on the later data can be further considered, so for the sequence data, the recurrent neural network often has a better effect. The core idea is to use the coordination mechanism of forwarding and backward propagation to continuously train the network and update the parameters. Forward propagation is like the normal flow of human blood circulation, inputting the neural network layer by layer to deeper neurons, and finally relying on the end neurons to output the results and do an error comparison with the actual output to judge whether there is an error and turn to the backward propagation stage of the error to update the parameter values when the results exist; reverse propagation is to propagate the output error back through the reverse network layer by layer to obtain the error information of each layer and use this as the basis for updating the weights of each layer.

The BP neural network uses the error function method to search for the optimal solution on the one hand and the

random number method to assign the initial weights on the other hand, which makes the BP neural network have no global search capability and may easily fall into local optimum and slow convergence speed. Therefore, the particle swarm mechanism is used to improve the parameters in the BP neural network, and the optimal model performance is obtained by optimizing the training of the network model and continuously updating the speed and position of the particle swarm. Some defects of the basic particle swarm are mentioned in the previous section, and the improved particle swarm algorithm IPSO is proposed based on two improved operations of inertia weights and adaptive variation of genetic algorithm, which results in not only improving the global optimization seeking ability of the model and accelerating the convergence speed but also improving the prediction accuracy of the model.

In the past, most neural network information transfer is one-way and the input and output dimensions are fixed, which cannot solve some complex problems. Therefore, a new model structure of the recurrent neural network is introduced, which gives the neural network the function of memory and can obtain the influence of historical information on current prediction.

$$a_l^t = \sum_{i=1}^{I} w_{il} x_i^t - \sum_{h=1}^{H} w_{hl} b_i^t + \sum_{h=1}^{C} w_{hl} c_i^t. \tag{1}$$

The processes of suitable data source selection, rigorous feasibility analysis, and fine data preprocessing greatly affect the final performance of the model. Therefore, to make the model accurately predict the sales of dishes, this part will introduce the three aspects of data sources, influencing factors processing, and data preprocessing.

The research object of this paper is the sales volume of the restaurant industry; however, the large differences in the dishes of stores in different business formats may seriously affect the generalizability of the model; at the same time, the daily sales volume of various dishes in most restaurant industries is proportional to the sales volume, and the sales volume of dishes can be further estimated from the daily dish sales; therefore, the sales volume of the restaurant industry is selected as the basic research object of this paper [18].

They jointly explained 48.9% of the variance of trust, and shared vision had the greatest impact; social interaction connection and shared vision had significant positive effects on reciprocity norm, and they jointly explained 43.6% of the variance of reciprocity norm. The direct relative returns of commodities and the relative returns of cross-selling have been mentioned earlier as two independent optimization objectives, and the direct relative returns of the rule corresponding to the combination of commodities are defined as the relative returns generated by the former set of commodities Z, i.e.,

$$Z_1 = \text{support}(Z) \sum_{i \in X} p\left(i^2\right). \tag{2}$$

The cross-selling relative return corresponding to this rule is defined as

$$Z_2 = \left(\frac{P(XY)}{P(X)} + \frac{P(\bar{X}Y)}{\bar{X}} \right) \sum_{i=1}^{Y} p\left(i^2\right). \tag{3}$$

A higher value of cross-selling relative return indicates that the correlation rule has a stronger effect on promoting the sales of the posterior item so that the candidate item combination not only has a high return of its own but also can promote the sales of other items.

Smoothing is performed on the nonstationary series. If the data series is not a random wandering trend, with a certain degree of regularity, the data is differenced, and the difference between the time series at a certain time interval is called delayed differencing. For the time series with periodic components, the delay may be periodic in period (width), the differencing process of the data is repeated, and the variance of the series is compared after all the several differencing, and the number of differencing with the smallest variance is selected as the result. At this time, part of the long-term trend and nonsmooth deterministic information of the original series is extracted, and all the time dependence of the original series will be eliminated [19]. The order of p and q is determined according to the AIC or BIC minimum criterion, and the estimates of the parameter coefficients are obtained by the least-squares estimation method. Least squares are a kind of estimation of parameters without first specifying the distribution of the series, with the principle of minimizing the sum of squares of the differences between the actual sales time series values and the model fitted values, with the help of an iterative method to make the residual sum of squares minimum as the goal. The value of a product drops rapidly over time, so even with strong promotions, sales of the product will still drop.

$$e_t = Y_t^{\text{new}} - {}^{\Gamma} Y_t. \tag{4}$$

According to the literature, in most cases, rolling step-by-step forecasting is more effective than multistep forecasting, while one-step rolling forward forecasting is used to predict one value at a time to the corresponding residual value, and the correlation information between multiple residual values after accumulation cannot be fully explored. The stacked self-coding neural network can learn in real time and improve the shortcomings of random initialization of the neural network, when the prediction produces a residual value, it starts to learn this residual value, and the residual value may have weak and long-lasting memory (correlation) between the residual values, the corresponding accumulated residual sequence may be nonsmooth, and the neural network is adaptable to the nonsmooth environment and can fit the nonlinearity well. There is a strong adaptive nature, which effectively expresses the characteristics of nonlinearity and randomness among the residual series data, as shown in Figure 2.

The first term is designed as the mean square error among the residual data, the second term is designed as a sum of weights that decreases with time, decreasing the magnitude of the weights, and a is the weight decay term. The third term is the sparse penalty term, β is the sparse penalty term coefficient, and p is the average activation of the

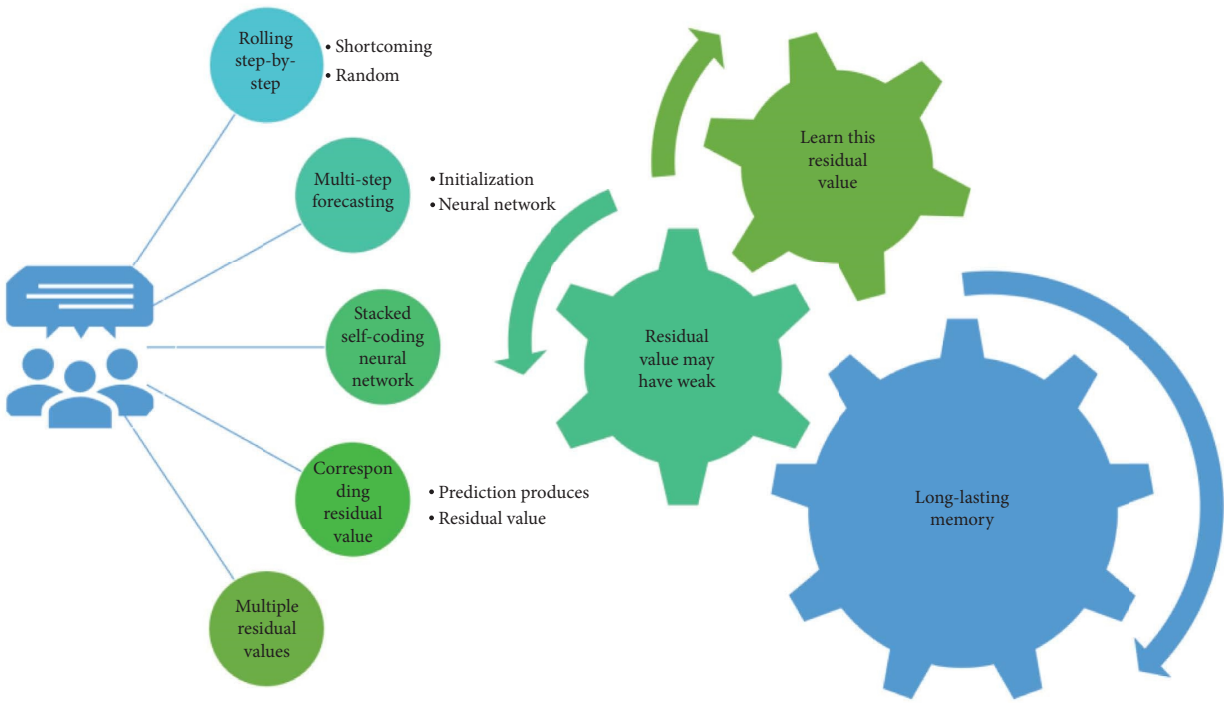


Figure 2: Sliding window calculation for input layer.

hidden layer node j, corresponding to the average activation of different subsequences. It is ensured that the distribution between the residuals learned by the computational model satisfies the approximately normal distribution, and the closer the distribution to the real white noise data, the better.

At this point, the neural network structure can make the linearly correlated or nonlinearly uncorrelated, relatively independent nodes in the hidden layer more sensitive to specific input features or residual data properties and explore the relevant linear region, which can be explored by learning the error of the objective function to obtain a neural network structure that can learn information.

Considering that the increase in the depth of the network structure can make the backpropagation fall into a local optimum, the deeper layers in the transfer of information lose some information, and the parameters of the layers near the output in the network are updated better, while the other layers are poor, to avoid training a local optimum. Training a network structure containing only one hidden layer that can learn the first-order features of the input residual data points, the network structure reaches optimization before starting to train the next one, and then these features are used as the input of the next network structure, using greedy learning ideas to complete the unsupervised pretraining of the overall stacked neural network structure.

Deep learning has gradually developed into an important branch of machine learning. This may be since the values in this time are all higher on the first few days of the forecast date, resulting in higher values predicted by the model inertia. Compared with traditional machine learning methods, a deep neural network of deep learning can project data tasks to a high-dimensional space for processing, thus handling prediction tasks of higher complexity and fitting a sales curve that more closely matches reality [20]. In addition, since the convolutional neural network in deep learning is insensitive to the absolute position of objects in the picture data when processing them, deep learning can also effectively solve unstructured data of picture type, as shown in Figure 3.

Furthermore, for sequential data like sales, traditional time series methods and machine learning methods treat each moment of data as an independent feature, while recurrent neural networks in deep learning can record several previous inputs and states, thus being able to further consider the influence of earlier data on later data in sequential data, so recurrent neural networks tend to have better results for sequential data. Then use the corresponding tools and channels to transform the knowledge, and finally the knowledge acquirer internalizes and absorbs the knowledge, which itself is a process of knowledge learning.

When goods are exchanged with money, the price is usually involved in the transaction as the amount of money per unit of goods, and the price is usually regarded as the reflection of the value of the goods themselves. When users shop online, whether they have a target product in mind or not, their knowledge will make them know or expect the performance of the product they are browsing, so the price of the product will have a significant impact on deepening or evoking consumer demand [21]. Price is also a major factor that consumers consider during the product information-gathering phase, which in turn influences the subsequent product utility perception and evaluation phase.

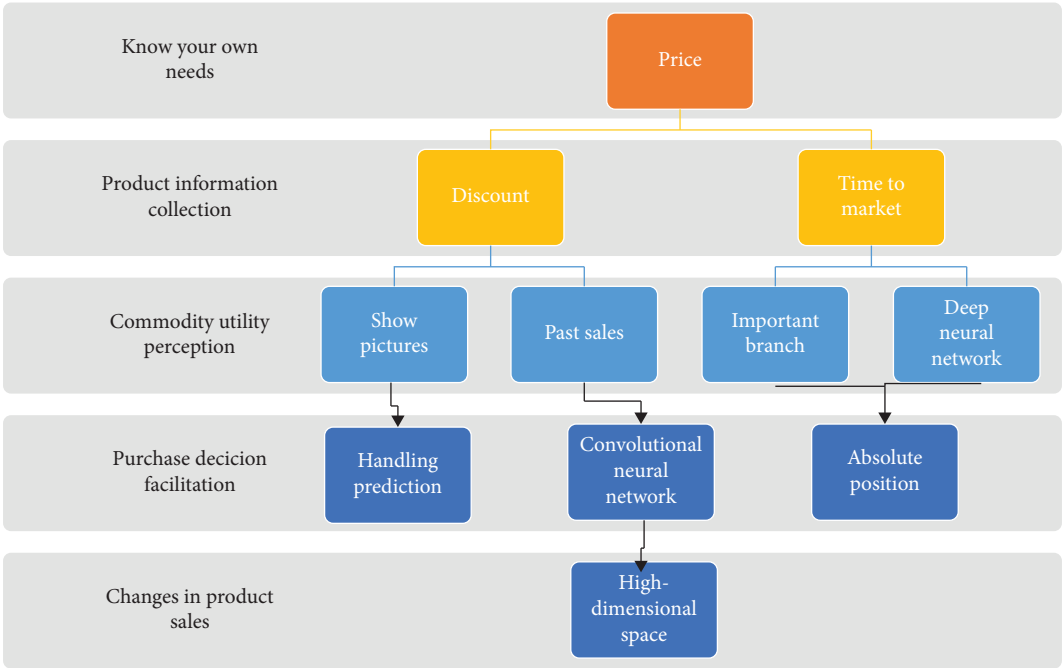


Figure 3: The formation stage of product sales and its influencing factors.

5. Analysis of Results

5.1. Results of the Analysis of Factors Influencing Knowledge Sharing in Virtual Communities. The results of the goodness-of-fit test and the regression matrix of the regression model show that the tolerance, VIF, and Durbin-Watson are within reasonable limits and there is no multicollinearity among the independent variables. Conversely, when there is a crisis of trust within an organization, the quantity and quality of knowledge sharing can become seriously problematic. Based on the regression coefficients and significance levels, the independent variables all positively affect knowledge sharing, all four basic hypotheses are valid.

This paper divides knowledge into explicit knowledge and tacit knowledge according to the characteristics of knowledge; i.e., the dependent variables are divided into two dimensions: explicit knowledge sharing behavior and tacit knowledge sharing behavior. To investigate whether there are differences in the effects of the respective variables under different knowledge characteristics, regression analysis is done separately in this paper. Figure 4 shows the regression analysis of independent variables for explicit knowledge sharing. It can be found that self-efficacy and expected return do not enter the fitted model and have a positive impact on market orientation and inventory control. In recent years, the competition among e-commerce enterprises has become increasingly fierce. Therefore, it can be concluded that the positive effects of self-efficacy and expected returns on explicit knowledge sharing are not significant, while the effects of social network structure and organizational trust on explicit knowledge sharing are significant, and social network structure has a relatively greater degree of influence than organizational trust.

In this paper, the regression analysis was done separately for explicit and tacit knowledge sharing by extracting self-efficacy and expectation reward, and the results showed that whether it is explicit or tacit knowledge and whether it is self-efficacy or expectation reward of individuals in the knowledge sharing process, they all positively affect the knowledge sharing behavior, but the degree of explanation for explicit knowledge sharing alone is low. To occupy a favorable position in the market, it is necessary to accurately predict and judge the competitiveness of commodities.

The reliability analysis of each research variable was conducted using SPSS, and Cronbach's α reliability coefficient was used to test the internal consistency of the questionnaire content. Validation factor analysis using AMOS showed that the fitted indicators (chi-squared degrees of freedom ratio = 2.061, GFI = 0.955, AGFI = 0.938, NFI = 0.954, IFI = 0.976, TLI = 0.970, CFI = 0.976, and RMSEA = 0.042) were all better than the standard values, indicating that the model was ideal. The standard factor loadings of each indicator were obtained to be above 0.7 and significant at the 0.001 level, with CR values greater than 0.7 and AVE values greater than 0.5, indicating that the scale has good convergent validity. The more rigorous AVE method was used to test the discriminant validity of the scale, and the results showed that the square root of the AVE values of each variable was greater than the corresponding correlation coefficients, indicating that the scale has good discriminant validity, as shown in Figure 5.

Social interaction bonding and shared vision have significant positive effects on trust and together explain 48.9% of the variance of trust, with shared vision having the largest effect; social interaction bonding and shared vision have significant positive effects on reciprocity norms and together explain 43.6% of the variance of reciprocity norms, with

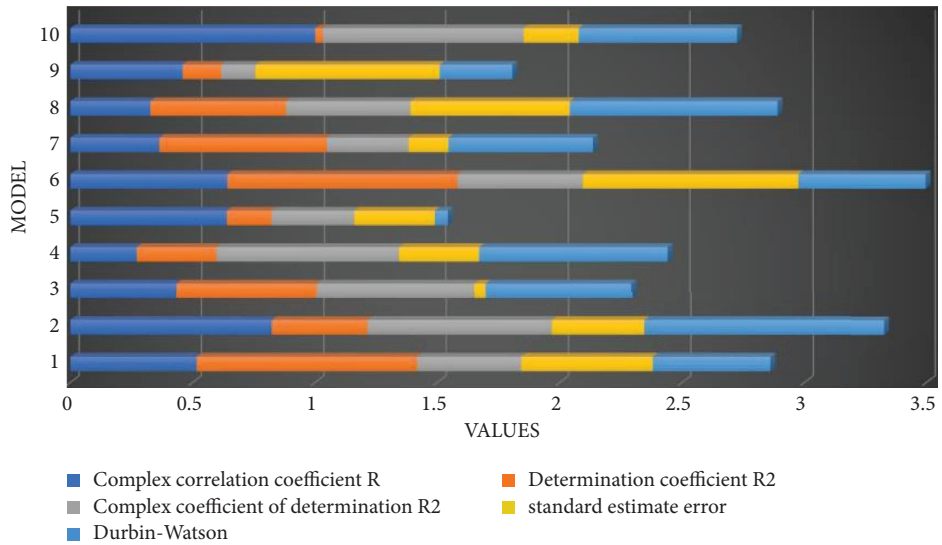


Figure 4: Goodness-of-fit test of the regression model of independent variables on explicit knowledge sharing.

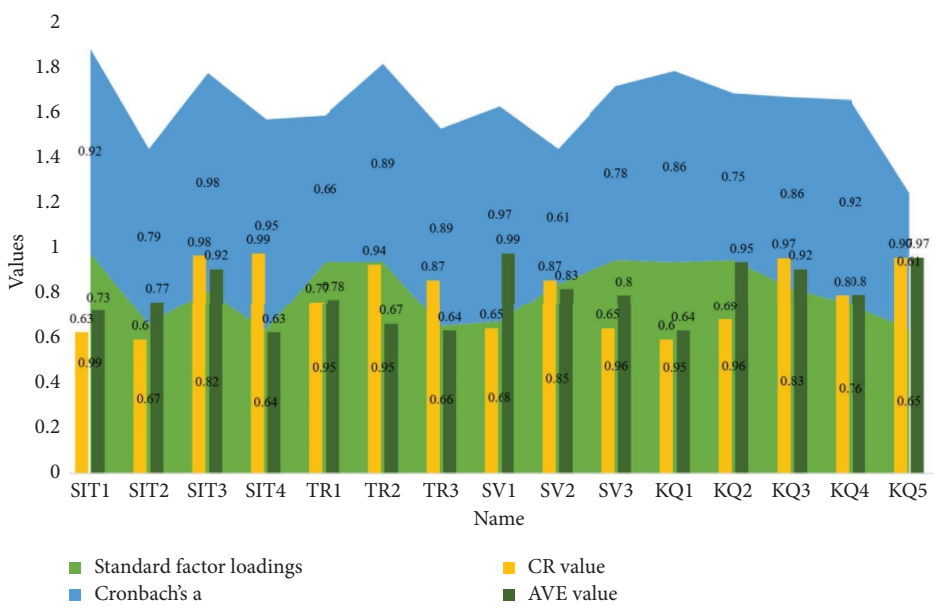


Figure 5: Standard factor loadings, Cronbach's alpha coefficient, CR values, and AVE values.

social interaction bonding having the largest effect; social interaction bonding has significant positive effects on a shared vision, and together they explain 27.9% of the variance of a shared vision. Social interaction connection has a significant positive effect on shared vision and jointly explains 27.6% of the variance of shared vision; trust, reciprocity norms, and shared vision all have a significant positive effect on knowledge quality and jointly explain 54.9% of the variance of knowledge quality, with the effect of reciprocity norms being the largest and the effect of social interaction connection on knowledge quality not being significant.

Based on this research objective, the relationship between social interaction linkages, shared vision and trust, and reciprocity norms, the relationship between social interaction linkages and shared vision, the relationship between social interaction linkages, trust, reciprocity norms, shared vision, and knowledge quality, and the mediating role of trust, reciprocity norms, and shared vision in academic virtual communities were explored with due consideration of the characteristics of academic virtual communities. A reasonable prediction will bring huge economic benefits and maintain the stable development of the market. After identifying the research questions, a questionnaire was designed using the question items that have been used in the existing domestic and international literature, combined with the results of a small-scale pretest. Based on this, an empirical study was conducted with questionnaires from

users who had participated in academic virtual communities. SPSS and AMOS statistical analysis software were used to analyze the data of the collected valid samples, including descriptive statistics, reliability analysis, hypothesis testing of the model, and mediating effect testing.

5.2. Analysis of Dynamic Prediction Results of Cross-Border e-Commerce Sales by Neural Networks. Most machines learning training is based on the batch gradient descent method. The batch gradient descent method will input all the training data into the model for each training, and for the overall data, there may be saddle points, and the training of the model may fall into the saddle points and stop training, which is the local optimum problem. An effective method is to randomly select a batch of data from the data for each training. Based on this training method, although it is almost impossible for the training process to directly descend in the direction of a large gradient, it makes it more likely to jump out of the saddle point, i.e., the local optimal point, during training. Moreover, for neural networks, the use of random batch gradient descent can reduce the training overhead due to its large number of weights, many computational steps, and the time required to bring in all the data for each optimization.

To avoid overfitting the training data, the data set is first divided into training and test data, and the loss values of both training and test data are observed. It shows the characteristics of short sales life cycle and easy depreciation. When the loss value of the test data no longer decreases, training is stopped. However, since the optimization process of stochastic batch gradient descent does not always optimize along the overall optimal direction, it is necessary to record the optimal loss value of the test data and stop training when n training sessions do not result in a better loss value and the loss value of the training data keeps decreasing.

In this paper, according to the different types of products of this enterprise, the sales of its products are counted by month, and four products with similar annual sales among the main products it sells are selected to study their sales changes and make a sales change trend graph, as shown in Figure 6. Different types of products showed different sales trends. The sales of down jackets, as cold-weather clothing, gradually increased after September when the weather became colder.

We can see that small promotion can lead to an increase in sales, but as the discounts increase, the sales decrease. The reason is that although price reduction will lower the economic cost of users and motivate them to make purchases, too much price reduction will also reduce the perceived quality of the product and increase the perceived risk of users, thus preventing them from making purchase decisions. In addition, large promotions are mostly seasonal clearance for companies, when price cuts are made to address the buildup of inventory caused by inaccurate sales forecasts in the past period. Apparel products are short life cycle products, and the value of products declines rapidly with time, so even with strong promotional actions, product sales will still decline. However, the trend and scatter plots show that the correlation between product sales and discounts is not that significant, so the relationship between product discounts and sales needs further consideration of other factors.

After the initialization of the neural network, we plot the distribution of the output of all hidden layers to prevent the Dead ReLu phenomenon. For some fast fashion products, such as shoes and hats, apparel, and fashion, it is more challenging to forecast the demand because the product life cycle is relatively short. According to the distribution of the network output, we can see that the initialization of this paper is effective, and many neuron nodes are in the active state. Taking the distribution of three convolutional layers of image data processing as an example, some neurons in the first convolutional layer are activated while some neurons are inhibited, which shows that the first convolutional layer maintains the nonlinearity of the neural network output, and the later higher convolutional layers do not have many zero outputs.

A reasonable initialization method can help avoid the gradient disappearance problem of neural networks with ReLu as the activation function but cannot avoid the gradient explosion problem caused by more layers of the neural network. In this paper, we use the gradient cropping method to avoid the problem of gradient explosion, which makes the weights of the deep neural network repeatedly wobble and fail to converge. Gradient clipping is to limit the gradient deinterval after the gradient calculation of each layer to avoid excessive gradients. After several experiments, it is found that setting the maximum and minimum values of the gradient to $[-0.01, 0.01]$ can avoid the gradient explosion while keeping the training speed as much as possible, as shown in Figure 7.

From the overall situation, there are two trends in the performance effect of the model. The first one is that the model using particle swarm optimization performs better than the single model in each index, and in turn, the model using improved particle swarm optimization performs better than the base particle swarm optimization model in each index; the second one is that the model of LSTM series outperforms the model of BP series in general. In terms of specific models, the IPSO-LSTM algorithm has the best results, corresponding to the smallest MAPE, RMSE, and MSE metrics and the largest 2R goodness-of-fit metric, which also largely indicates that our improved particle swarm approach is effective and can improve the prediction accuracy.

All six models can fit the true value trend, but the predicted values are higher than the actual values during the peak hours of 6:00 p.m. to 7:00 p.m. This may be because the values are all higher during this period in the days before the prediction date, resulting in higher values predicted by the model inertia. In terms of the individual models, the IPSO-LSTM algorithm model is better able to fit the trend of the true values and can accurately portray the predicted changes in sales, while the single model BP neural network is the worst fit and deviates far from the actual values.

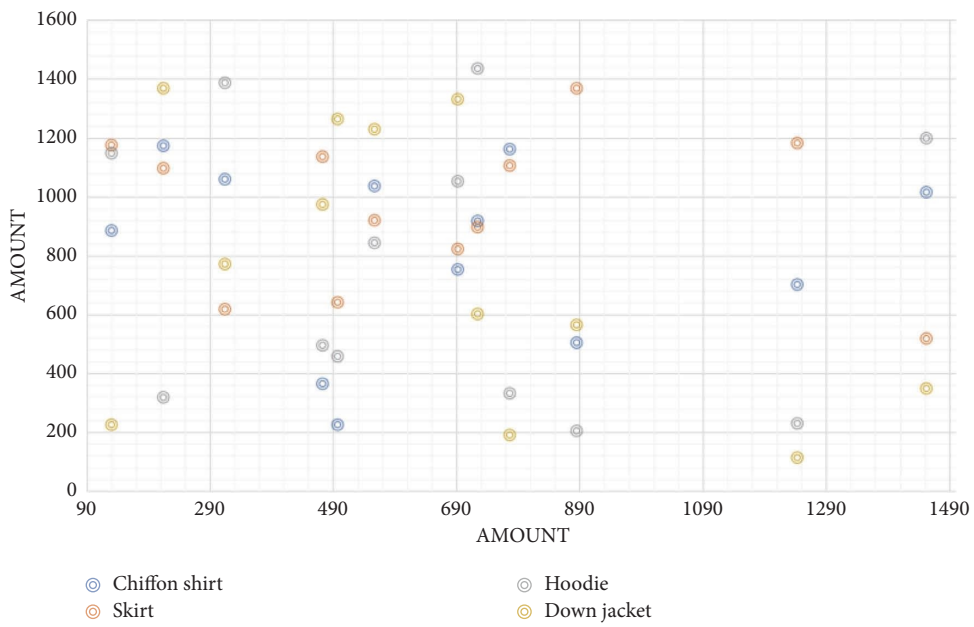


Figure 6: Sales trends of different products.

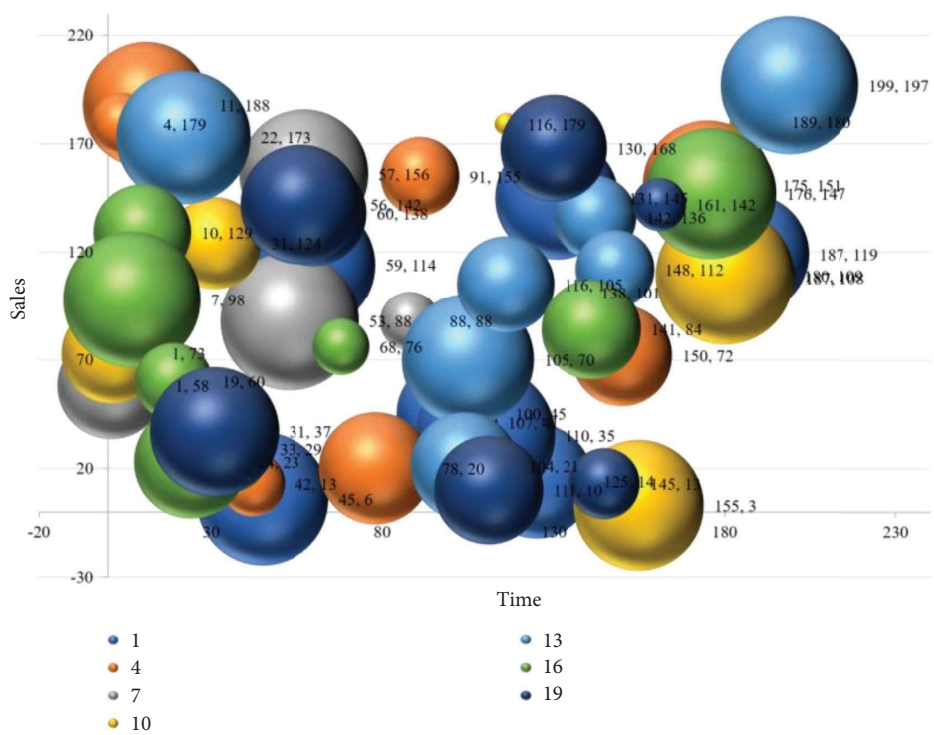


Figure 7: Price prediction results.

6. Conclusion

Under nontransactional virtual community codependency, users continuously participate to maintain dependency, increase loyalty to the virtual community, and continue knowledge sharing. To maintain identifiable identity, public self-awareness, group norms, and community participation to gain recognition and affirmation, users will invest more energy and thus enhance their loyalty to virtual communities; similarly, they will be more willing to sustain knowledge sharing in virtual communities. Forecast the demand and sales of fast fashion products. The biggest product features of fast fashion products are quick feedback on show design, shorter product life cycle, and higher frequency of updating product categories and types. In contrast, the higher the loyalty of nontransactional virtual

community users, the more they want the virtual community to operate more successfully and thus continue to share knowledge in the virtual community. Given the importance of product reorder rate, a Poisson regression model is used to calculate the reorder rate of different products in the case and the factors that play a strong influence on the reorder rate according to the formula, which lays the data foundation for subsequent product combination recommendation and product sales prediction. A nested neural network with BP, LSTM, and Verhulst models based on the AP algorithm is constructed. The BP neural network combines the current independent variables to predict sales, the LSTM neural network explores the influence of historical data on sales, and the Verhulst model can predict sales based on the growth trend of the dependent variable. The AP algorithm is used to forecast and analyze the sales of goods, and the error comparison proves that nested neural networks are more accurate in predicting the sales of goods.

Data Availability

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

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work was supported by Business School, Nanfang College Guangzhou.

References

[1] C. Wu, B. Yan, R. Yu et al., "Retracted article: digital forensics system based on dynamic path prediction and the competitiveness of cross-border e-commerce SMEs," *Personal and Ubiquitous Computing*, vol. 25, no. 1, p. 11, 2021.

[2] G. Li and N. Li, "Customs classification for cross-border e-commerce based on text-image adaptive convolutional neural network," *Electronic Commerce Research*, vol. 19, no. 4, pp. 779–800, 2019.

[3] H. Xia, J. Weng, J. Z. Zhang, and Y. Gao, "Rural E-commerce model with attention mechanism: role of Li ziqi's short videos from the perspective of heterogeneous knowledge management," *Journal of Global Information Technology Management*, vol. 25, no. 2, pp. 118–136, 2022.

[4] A. Rosário and R. Raimundo, "Consumer marketing strategy and E-commerce in the last decade: a literature review," *Journal of Theoretical and Applied Electronic Commerce Research*, vol. 16, no. 7, pp. 3003–3024, 2021.

[5] S. Guan, "Smart E-commerce logistics construction model based on big data analytics," *Journal of Intelligent and Fuzzy Systems*, vol. 40, no. 2, pp. 3015–3023, 2021.

[6] J. R. Mumu, P. Saona, M. A. A. Mamun, and M. A. K. Azad, "Is trust gender biased? A bibliometric review of trust in E-commerce," *Journal of Internet Commerce*, vol. 21, no. 2, pp. 217–245, 2022.

[7] J. Mou, Y. Cui, and K. Kurcz, "Bibliometric and visualized analysis of research on major e-commerce journals using Citespace," *Journal of Electronic Commerce Research*, vol. 20, no. 4, pp. 219–237, 2019.

[8] D. Ge, Y. Pan, Z. J. M. Shen, D. Wu, R. Yuan, and C. Zhang, "Retail supply chain management: a review of theories and practices," *Journal of Digital Information Management*, vol. 1, no. 1-2, pp. 45–64, 2019.

[9] L. Cao, Q. Yang, and P. S. Yu, "Data science and AI in FinTech: an overview," *International Journal of Data Science and Analytics*, vol. 12, no. 2, pp. 81–99, 2021.

[10] P. Pananond, G. Gereffi, and T. Pedersen, "An integrative typology of global strategy and global value chains: the management and organization of cross-border activities," *Global Strategy Journal*, vol. 10, no. 3, pp. 421–443, 2020.

[11] N. Pouti, M. T. Taghavifard, M. R. Taghva, and M. Fathian, "A comprehensive literature review of acceptance and usage studies in the social commerce field," *International Journal of Electronic Commerce Studies*, vol. 11, no. 2, pp. 119–166, 2020.

[12] M. He, W. Li, B. K. Via, and Y. Zhang, "Nowcasting of lumber futures price with google trends index using machine learning and deep learning models," *Forest Products Journal*, vol. 72, no. 1, pp. 11–20, 2022.

[13] J. Shaikh, "E-commerce business models in ethiopian market: challenges and scope," *INFORMATION TECHNOLOGY IN INDUSTRY*, vol. 9, no. 3, pp. 17–25, 2021.

[14] Z. Wang and H. Zhu, "Optimization of e-commerce logistics of marine economy by fuzzy algorithms," *Journal of Intelligent and Fuzzy Systems*, vol. 38, no. 4, pp. 3813–3821, 2020.

[15] G. Wang, "Blockchain technology helps the development of meteorological informatization," *Journal of Business and Management Sciences*, vol. 7, no. 3, pp. 112–120, 2019.

[16] X. Xu, Y. He, and Q. Ji, "Collaborative logistics network: a new business mode in the platform economy," *International Journal of Logistics Research and Applications*, vol. 25, no. 4-5, pp. 791–813, 2022.

[17] Y. Luo, J. Ma, and C. K. Yeo, "Identification of rumour stances by considering network topology and social media comments," *Journal of Information Science*, vol. 48, no. 1, pp. 118–130, 2022.

[18] D. Pérez-Campuzano, P. M. Ortega, L. R. Andrada, and A. López-Lázaro, "Artificial Intelligence potential within airlines: a review on how AI can enhance strategic decision-making in times of COVID-19," *Journal of Airline and Airport Management*, vol. 11, no. 2, pp. 53–72, 2021.

[19] G. Lopes, "The wisdom of crowds in forecasting at high-frequency for multiple time horizons: a case study of the Brazilian retail sales," *Brazilian Review of Finance*, vol. 20, no. 2, pp. 77–115, 2022.

[20] G. F. Watson, S. Weaven, H. Perkins, D. Sardana, and R. W. Palmatier, "International market entry strategies: relational, digital, and hybrid approaches," *Journal of International Marketing*, vol. 26, no. 1, pp. 30–60, 2018.

[21] F. Herzallah, M. M. Ayyash, and K. Ahmad, "The impact of language on customer intentions to use localized E-commerce websites in Arabic countries: the mediating role of perceived risk and trust," *The Journal of Asian Finance, Economics and Business*, vol. 9, no. 1, pp. 273–290, 2022.

Hindawi
Computational Intelligence and Neuroscience
Volume 2022, Article ID 7655001, 8 pages
https://doi.org/10.1155/2022/7655001

Research Article

Influence of Voice Interactive Educational Robot Combined with Artificial Intelligence for the Development of Adolescents

Yadong Zhang[1] and Hongkai Wang[2,3]

[1]*School of Art and Design, Shandong Women's University, Jinan 250000, China*
[2]*Academy of Arts and Design, Tsinghua University, Beijing, China*
[3]*College of Journalism and Communications, Shih Hsin University, Taipei, China*

Correspondence should be addressed to Hongkai Wang; d108810005@mail.shu.edu.tw

Received 8 August 2022; Revised 9 September 2022; Accepted 16 September 2022; Published 6 October 2022

Academic Editor: Ning Cao

In the context of multicultural information, to explore and analyze the use effect of voice interactive educational robot in the classroom of adolescent students, and the physical and mental impact of movie characters on adolescent students, and to lay the foundation for studying the positive development of adolescents, under the guidance of positive psychology theory, the relationship between positive psychology and adolescent mental health is analyzed, the application of adolescent educational robot is discussed, and the relationship between adolescent educational robot combined with movies characters and positive development of adolescent is analyzed. The questionnaire is used to collect data, including the questionnaire on the influence of movie characters on the positive development of adolescents, the questionnaire of pre- and postpopularization of artificial intelligence in primary and secondary schools, and the satisfaction questionnaire of voice interactive educational robot. The reliability and validity of the questionnaire are analyzed. The results show that students who watch more than 20 movies are 20% in grade one and that is only 10% in grade three. 79% of the students think that the movie characters have an impact on themselves. The distribution of the number of the grade two students exposed to the movie is relatively uniform, 40% of them watch 10-20 movies, and 82% of them think that the movie characters have an impact on themselves; the Cronbach coefficient of classroom satisfaction questionnaire is $0.929 > 0.9$, the average value of the correlation of total correction items of the corresponding item is $0.612 > 0.5$, the Kaiser–Meyer–Olkin measurement value is $0.812 > 0.6$, and the Sig value is $0.000 < 0.05$, indicating that the reliability and validity of the questionnaire are very high and 92.3% of the students are very satisfied with the classroom. This shows that voice interactive educational robot combined with movie characters can promote the positive development of adolescents.

1. Introduction

The development of human society is becoming more and more complicated. In the case of rapid multiparty flow of information, the speed of cultural renewal is also greatly improved. Meanwhile, different cultures need to serve the development of the society, which gradually forms a multiculture [1, 2]. Adolescents are the hope of national development, and they should grow and develop in a multicultural background. The common psychological problem of teenagers is the inferiority complex [3]. Most people have had inferiority complex in a certain period of growth or under certain circumstances. A slight sense of inferiority complex can promote people's discovery and improvement of their shortcomings, but excessive inferiority complex will make people lose hope for life, produce negative self-definition, and ultimately have adverse consequences [4]. Effective psychological intervention can help schools and families to pay attention to the harm of adolescents' inferiority complex, which can help adolescents establish self-confidence and make them in a healthy state of development. At present, foreign countries have attached great importance to the

psychological problems caused by the sense of inferiority of adolescents. Most of the domestic studies in related fields focus on the theoretical aspect. In fact, only a few of them have taken intervention measures against the sense of inferiority of adolescents, and they have not intervened according to the theory of positive psychology [5].

Information diversity makes society in the information explosion period, and people are surrounded by information all the time. Because of the particularity of the physical and mental environment, adolescents belong to the special group in the media information environment and are vulnerable to the impact of media information, which has an impact on the positive personality development of adolescents [6]. The movies are the core product of the media environment, which has a great impact on the thinking and behavior of people and adolescents. Adolescents are in a psychological state of trust for movie information, but they are also in a passive state. In the face of complex movie types, how should adolescents choose movies that are beneficial to their development, and under what conditions do they internalize the information conveyed by the movie characters? These are the problems that need to be solved when movies are selected for adolescents [7]. With the progress of science and technology, while media information is used to develop education on adolescents, the cultivation of artificial intelligence talents has occupied the core part of China's education. Adolescents are the reserve force of national science and technology progress. The artificial intelligence enlightenment education of adolescents is highly valued by the education department. The incorporation of artificial intelligence education into the classroom of adolescents will provide hope for the improvement of the national education level [8]. Artificial intelligence education in foreign countries starts earlier, and the artificial intelligence education in the United States has entered the national strategic stage, which strongly supports the relevant funds, policies, and talent introduction. Adolescent artificial intelligence education in Australia is a compulsory course. Adolescents will study the theoretical knowledge of artificial intelligence in the way of discussion and experimental operation [9, 10]. In the current adolescent education in China, the use of multimedia classrooms and the information-based training of educators show that artificial intelligence education has gradually replaced the traditional teaching mode and promoted the leap from theory to practice stage of adolescent education [11].

Based on this, under the guidance of positive psychology theory, questionnaire survey and empirical analysis are used to study the use of voice interactive robot in the classroom of adolescents and the influence of movie characters on the physical and mental development of adolescents, which can provide a practical basis for promoting the positive development of adolescents.

2. Method

2.1. Positive Psychology and the Development of Adolescents. Adolescents are in a critical period of physical and mental development. Positive psychology is used to help adolescents overcome inferiority complex and form a positive self-awareness concept, which can greatly promote the mental health level and positive development of adolescents [12].

2.1.1. Theoretical Basis of Positive Psychology. The research content of positive psychology includes positive emotional feelings such as full of love and subjective well-being, and how to obtain positive feelings and healthy mental state [13]. In positive psychology, it is proposed that positive emotional feeling can make an individual's work and life full of positive energy. According to the theory of "Broaden-and-Build" [14], even the unsystematic positive emotion can change the individual's instantaneous behavior ability and also improve the individual's intelligence quotient, social adaptability, and physical strength. Positive emotional experience can also actively prevent diseases and can promote the body's immunity to diseases. In positive psychology, it is also pointed out that the external and cultural environment will change the mental health level of individuals in varying degrees. When the positive personality characteristics are established, the individual also establishes a positive environment system. It means that positive psychology starts from the factors of individual growth environment and promotes individual to form positive personality by establishing positive external environment, family environment, and campus environment.

2.1.2. The Relationship between Positive Psychology and Mental Health of Adolescents. In mental health education, by observing the age and physical and mental development characteristics of adolescents and using various methods and measures, the mental potential of adolescents has been continuously explored, and their ability to resist pressure has been improved, which makes the physical and mental health of adolescents develop harmoniously [15]. As the main field of learning and life of adolescents, school mental health education belongs to the practical influence range of psychology. Under the influence of positive psychology, the teaching methods of school mental health education are constantly changing. In the past, school mental health education focused on the prevention and treatment of students' mental problems, and the measures taken were mainly mental education courses and psychological counseling, which can relieve the pressure and distress of adolescent students [16]. The positive mental health education starts from the positive elements of the students and guides the students to pay attention to their advantages, to be good at paying attention to the positive energy around them and to cultivate their self-confidence, which can produce a sense of self-efficacy, and make them develop in an all-round way while paying attention to their spiritual needs.

2.2. Artificial Intelligence and Active Development of Adolescents. In artificial intelligence, machines are used as the intermediate carrier to complete the simulation of human work and life [17]. Its core part is human intelligence. First, the brain is used as the intermediary to generate the ability of memory, research, and exploration; second, the perception organ carries out the perception of external stimulation;

finally, the mouth, arm, and other parts are used for quick information feedback. As a part of computer science, artificial intelligence has a great impact on the application and development of information technology [18].

2.2.1. Theoretical Basis of Educational Robot. The important purpose of the educational robot is to assist classroom education. The aim of artificial intelligence educational robot is people-oriented. Developmental toys can be used as an auxiliary tool for the teaching of basic education courses in schools. It can be used to educate students from various directions and finally realize "education through entertainment." Educational robots have the immeasurable potential in the process of education, which can make the educational cause have subversive changes. The object of the educational robot is students, and it can program students' behavior to achieve the purpose of assisting teaching. The robot-assisted teaching is the activity of teaching and learning with robot as a teaching aid tool. From the perspective of sociology [19], the educational robot can play the role of teacher and classmate and complete corresponding tasks; from the perspective of communication studies [20], it can expand the dissemination of teaching information, improve educators' knowledge reserve, and reduce students' learning pressure; based on the perspective of educational psychology, the educational robot can promote students' learning behavior and make students' knowledge perception and comprehensive utilization ability be improved. At present, the teaching forms of the educational robot are various, including the auxiliary of giving and listening to lessons, answering questions, teaching, and learning games. Table 1 shows the classification of educational robots.

2.2.2. The Relationship between Voice Interactive Educational Robot and the Development of Adolescents. The voice interactive robot can use a computer carrier to transform speech content into text. The goal is to realize the language conversation of the machine, so that the machine can understand the meaning of human speech [21]. Hidden Markov model is used; feature extraction, pattern matching, and model training technologies are combined; and the theoretical model of digital signal processing is mixed with many disciplines such as phonetics [22]. For the voice recognition of the current voice interactive robot, a great breakthrough has been made in the defects of accent, speech speed, and noise influence, which greatly promotes the education of adolescent students. The main application objects of the voice interactive educational robots are college students and adolescent primary and secondary school students. Its goal is to improve the scientific literacy of college students and adolescents. For adolescent students, improving their innovative thinking is the key purpose, which can greatly stimulate students' interest in learning. The application principle of voice interactive educational robot is the educational concept of "education through entertainment." It is used in language and behavior programming to achieve the purpose of educational companion, which is widely favored by adolescent learners. Many domestic digital learning equipment manufacturing enterprises, such as Sang Technology company, are committed to the research and development of voice interactive educational robots and corresponding teaching platforms, instilling the concept of robot education into primary and secondary school curriculum education. Through the synthesis of artificial intelligence thought and artificial intelligence knowledge, the voice interactive educational robot realizes the auxiliary function of artificial intelligence teaching, completes the perfect integration of technology and knowledge, creates a real artificial intelligence environment for the adolescent students, and helps them to explore the artificial intelligence world preliminarily by applying colorful activities to develop the students' intelligence to the greatest extent. Figure 1 shows the structure of the voice interactive educational robot.

2.3. Movie Characters and the Positive Development of Adolescents

2.3.1. Theoretical Basis. The core concern of positive psychology is the research on the individual's development potential and positive personality traits, which is a newly developed discipline [23]. Because of its special psychological characteristics, adolescents are in a special period of mental health education. Therefore, the combination of positive psychology and mental health education of adolescents has become a new direction of the development of education in China. At the same time, the corresponding research on the positive development of adolescents has gradually attracted attention. The positive development of adolescents was first put forward by Little in 1993. After continuous theoretical and practical improvement, it has formed a psychological theory with the development of adolescents based on the positive attitude as the core concern. In American encyclopedia, it is proposed that positive development provides a theoretical basis for schools to improve students' interest in learning and learning ability. Its core concern is to provide a supportive environmental atmosphere for adolescents and to provide practical exercises for adolescents through different channels, which make adolescents have the courage to contribute to the society. In addition, a positive attitude is adopted to pay attention to the development of adolescents to help them find their advantages and disadvantages as well as their learning ability, so that their advantages can be developed to the maximum extent. Effective methods are explored to develop the advantages of adolescents and encourage them to actively participate in meaningful activities. Therefore, the emergence of positive psychology can make adolescents focus on their positive aspects and not pay too much attention to their negative aspects and turn the adolescent education view to the education that promotes the potential of adolescents [24].

Bandura proposed that the ability to learn through observation can harvest complex and unified holistic behaviors, and it is not necessary to get correct conclusions in the process of constantly trying and making mistakes [25].

Table 1: Classification of adolescent educational robots.

Classification	Main function
STEAM teaching aids (Kaizhi KP3, Fischertechnik, and mBot)	Scientific, free to build, physics science popularization, visualization
Classroom-assisted teaching (Keeko intelligent robot)	Language teaching, teaching according to age characteristics, strong sports ability

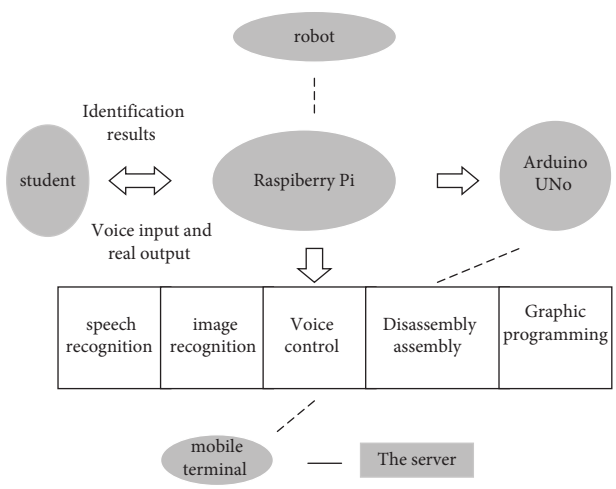


Figure 1: Structural framework of voice interactive educational robot.

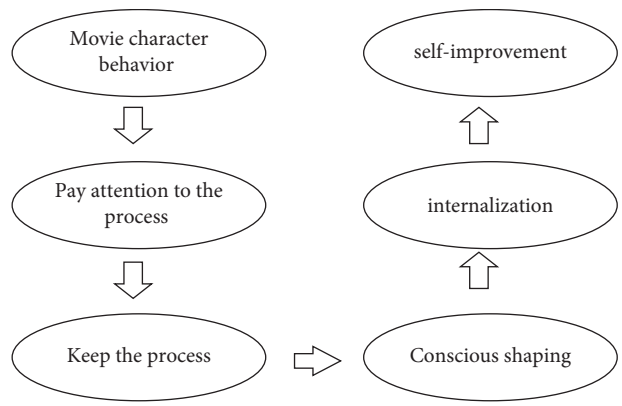


Figure 2: Model of movie's influence on adolescents.

Individuals can internalize and learn by observing other people's behaviors in solving problems. People learn only by trying and making mistakes, which is very hard and time-consuming. If their behavior and the corresponding results are observed, learners can correct their behavior without their practice. According to the guidance of observational learning theory, the influence of characters in contemporary movies on the positive quality of adolescents is analyzed.

2.3.2. The Influence of Movie Characters on the Positive Development of Adolescents. The main function of the movie is to create vivid characters and set their personality charm and show the rich cultural and social characteristics of the modern society in many performance forms to exert its subtle influence [26]. Under the current era background and social trend, the way of expression of the outlook on life, values, and behavior characteristics expressed by movie characters conforms to the characteristics of the times. It can help adolescents to fully understand the beautiful life and society, learn the behavior norms in daily life, and help adolescents release their emotions and learn to release their pressure with entertainment. In the process of watching movies, adolescents do not need to spend too much time adapting to the movie environment and can easily feel the information expressed by the movie characters. At the same time, some information and concepts are transformed into their internal parts, which influence the behavior of adolescents. Therefore, from a certain point of view, movies and movie characters become educational tools outside the campus, which can promote the growth and positive development of adolescents. Figure 2 shows a model of the impact of movies on adolescents.

At the same time, because of the wide-spread advantages of movies, the positive educational effect of movie characters on adolescents will even exceed the influence of campus education. This can lead to educators' thinking and attention, that is, to introduce movies and movie characters into the campus teaching environment, to stimulate students' learning enthusiasm through the movie's liveliness, and to carry out cultural education under the condition of physical and mental relaxation of adolescents, which plays an important auxiliary role in the positive development of adolescents.

2.4. Survey Method. Questionnaires are used to collect data, including the questionnaire on the influence of movie characters on the positive development of adolescents, the questionnaire of pre- and postpopularization of artificial intelligence in primary and secondary schools, and the satisfaction questionnaire of voice interactive educational robot. Then, the information in the questionnaire is statistically analyzed.

(1) The questionnaire on the influence of movie characters on the positive development of adolescents consists of two parts. One is the types and attention of adolescents to movie characters; the other is the influence of movie characters on adolescents. XX middle school students are randomly selected as the research object. According to the proportion of boys and girls in the school, a total of 260 questionnaires are distributed. A total of 200 valid questionnaires are selected based on the influence of the movie characters proposed by the students. From the

effective questionnaire, 80 students (52 boys and 28 girls) are randomly selected to conduct in-depth interviews on the influence of movie characters. The results of the interview are consistent with the hypothesis of the valid questionnaire.

(2) Class 3 students (72 students) in grade two of XX middle school are investigated. According to the Learning and Study Strategies Inventory (LASSI), evaluation of classroom satisfaction of the voice interactive robot (artificial intelligence educational robot car with Raspberry Pi and Arduino as the carrier) is carried out in four directions: teaching method, teaching attention, teaching atmosphere, and teaching effect feedback. For the evaluation of teachers, there are teaching methods and teaching organization. In terms of the students' attention, learning awareness, and behavior, there are 15 questions in total. The questionnaire imitates Likert five-point scale method [27], including five parts: very satisfied, satisfied, general, dissatisfied, and very dissatisfied. The scores can be divided into 5, 4, 3, 2, and 1 in turn, as shown in Tables 2 and 3.

Table 2: Classroom evaluation methods.

Object	Evaluation direction
Teachers	Teaching organization; teaching method
Students	Attention level; learning awareness and behavior

Table 3: Setting of classroom satisfaction questionnaire.

Degree	Score
Very satisfied	5 points
Satisfied	4 points
General	3 points
Dissatisfied	2 points
Very dissatisfied	1 points

3. Results

3.1. The Influence of Movie Characters on the Positive Development of Adolescents. According to the effective questionnaire, in-depth communication on whether the movie characters have an impact on them is conducted. Figures 3(a) and 3(b) show the results.

Figure 3 shows that the grade one students of middle school can contact with movies, but the distribution of contact quantity is uneven. In grade one, the number of students who watch less than 10 movies accounts for 46%, only 20% of them watch more than 20 movies, and only 30% of them think that the movie characters have an impact on themselves; the distribution of the number of students in grade two exposed to the movie is relatively uniform, 40% of them watch 10-20 movies, and 82% of them think that the movie characters have an impact on themselves; the distribution of the number of students exposed to movies in grade three is also uneven, 55% of them watch less than 10 movies, only 10% of them watch more than 20 movies, and 79% of them think that the movie characters have an impact on themselves.

3.2. The Influence of Voice Interactive Robot on the Positive Development of Adolescents

3.2.1. Reliability of the Questionnaire. First, the reliability of the questionnaire is analyzed for 72 students who participate in the class, as shown in Figure 4.

Figure 4 shows that the Cronbach coefficient of the classroom satisfaction questionnaire is $0.929 > 0.9$, which indicates that the quality of the collected information data is high. Compared with the Cronbach coefficient, there is no obvious change in the α coefficient value of the deleted item, so all 15 items in the questionnaire can be used. The average value of the total correlation of the corresponding correction items is $0.612 > 0.5$, indicating that the correlation between the analysis items is very good. Therefore, the reliability level of information data is high, and further study can be carried out.

3.2.2. Validity of the Questionnaire. According to the above results, the validity of the questionnaire is analyzed by factor analysis. Table 4 shows the analysis results.

Figure 4 shows that the Kaiser–Meyer–Olkin measurement value is $0.812 > 0.6$, and the Sig value is $0.000 < 0.05$, indicating that the classroom satisfaction questionnaire can be used to carry out factor analysis.

Finally, according to the classroom satisfaction questionnaire, score results are counted, as shown in Figure 5.

Figure 5 shows that there are 48 students with 5 points and 4 students with 4 points. The number of students who get 1, 2, and 3 points is 0, which means that 92.3% of students are very satisfied with the classroom, that is to say, voice interactive robots are welcomed by most students.

4. Discussion

Grade one students have just entered middle school, and they are in a strange and curious stage about the school environment. Therefore, they will not put too much energy into the movie. Therefore, the proportion of students who watch more movies is smaller, and the influence of movie characters on them is not strong. Grade two students are in the critical period of adolescents' ideological and moral transformation [28]. Being familiar with the campus environment and learning, they will have a strong curiosity about other things and have time and interest in watching movies. Therefore, the proportion of grade two students who watch more movies is larger because of the particularity of students' physical and mental characteristics [29, 30], and the influence of movie characters on them is also very strong. The grade three students face the pressure of entering school and have no time and energy to watch movies, so the proportion of students who watch more movies is smaller. However, grade three students tend to be mature physically and mentally, so they can distinguish the behavioral characteristics in the movie characters and are easily influenced

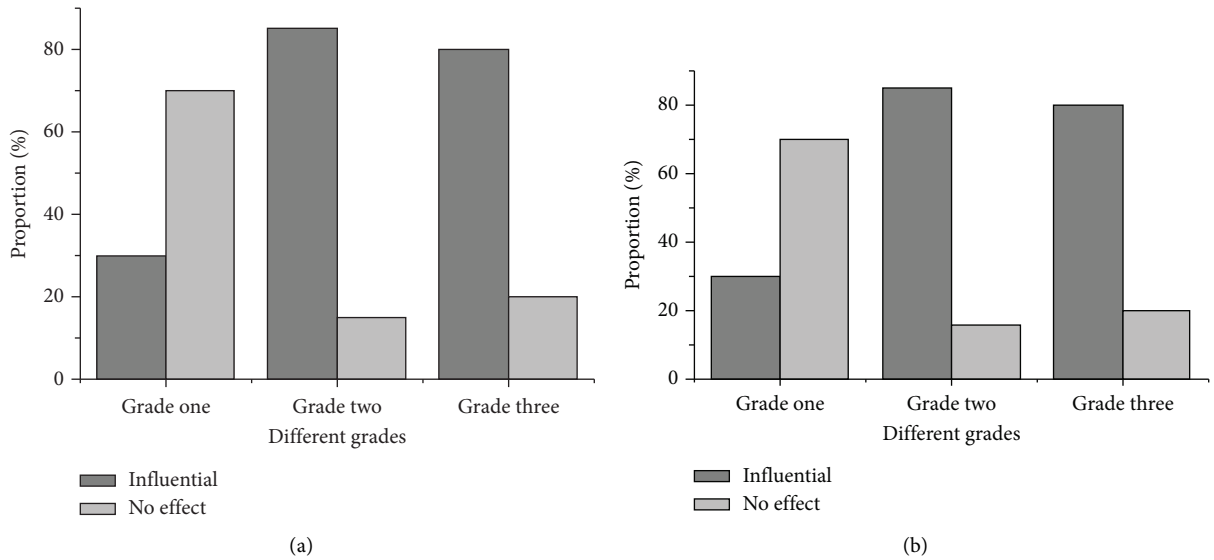


Figure 3: Influence of movie characters on middle school students ((a) the proportion of students in three grades of middle school watching movies; (b) the proportion of students affected by movie characters).

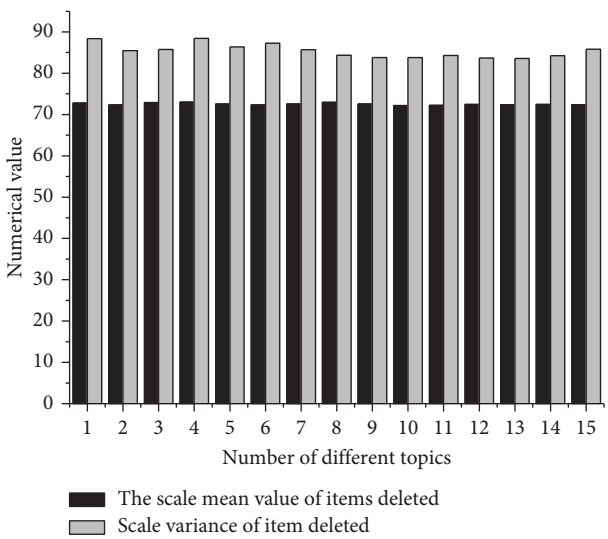


Figure 4: Analysis results of classroom satisfaction questionnaire (the Cronbach coefficient of each item is 0.929, the average value of the correlation of total correction items is 0.612, and the average value of α coefficient deleted items is 0.927).

Table 4: Results of validity analysis of classroom satisfaction questionnaire.

Different parameters	Numerical value
Kaiser–Meyer–Olkin measurement	0.812
Approximate chi square	1050.135
D*f*	14
Sig	0.000

by the characters, which is consistent with the research results of Hsia et al. [31].

Students are very interested in the human-computer interaction classroom learning mode, and most students

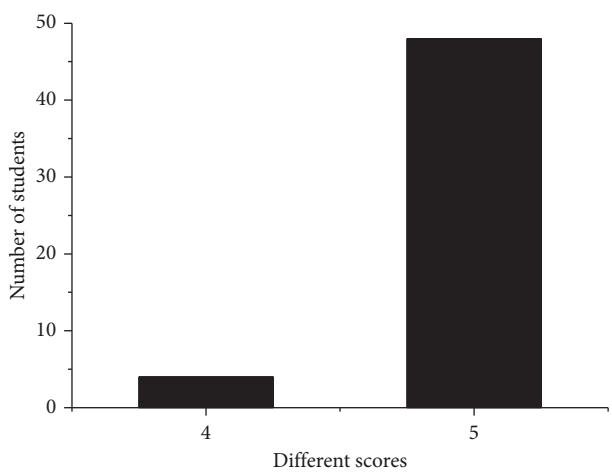


Figure 5: Score of classroom satisfaction questionnaire.

show interest in the teacher's teaching method, which is consistent with the research results of Li et al. [32]. Students hope to learn more extensive artificial intelligence knowledge with their classmates on the school, and then use their spare time to understand other knowledge in the field of artificial intelligence [33–38]. Some students hope to carry out the work in the field of artificial intelligence in the future [39–46]. It shows that the voice interactive educational robot under artificial intelligence can promote the positive development of adolescents.

5. Conclusion

Under the background of multicultural information, based on the theory of positive psychology, the voice interactive educational robot using artificial intelligence works as the assistant education tool of adolescents. More than 95% of the students are very satisfied with the classroom, and they have

a great interest in learning in the classroom. For middle school students, the number of students watching movies in grade two is higher. Even though the proportion of students watching more movies in grade one and grade three is relatively small, the influence of movie characters on students of three grades is relatively great. Therefore, voice interactive educational robot combined with movie characters will promote the positive development of adolescents. However, there are also some deficiencies. The area and number of participants in the questionnaire survey are limited, and the data is lack of certain universality, which needs to be improved in the follow-up study.

Data Availability

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethical Approval

This article does not contain any studies with human participants or animals performed by any of the authors. Informed consent was obtained from all individual participants included in the study.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

The authors acknowledge the help from the university colleagues.

References

[1] H. Korzilius, J. J. Bücker, S. Beerlage, and S. Beerlage, "Multiculturalism and innovative work behavior: the mediating role of cultural intelligence," *International Journal of Intercultural Relations*, vol. 56, no. Complete, pp. 13–24, 2017.

[2] M. Verkuyten and J. A. M. Maykel, "Immigration discourses and their impact on multiculturalism: a discursive and experimental study," *British Journal of Social Psychology*, vol. 44, no. 2, pp. 223–240, 2005.

[3] A. Kamper, "Psychological and psychosomatic problems in adolescence," *Monatsschrift Kinderheilkunde*, vol. 163, no. 9, pp. 900–910, 2015.

[4] M. Rubin, N. Denson, S. Kilpatrick, K. E. Matthews, T. Stehlik, and D. Zyngier, "I am working-class: subjective self-definition as a missing measure of social class and socioeconomic status in higher education research," *Educational Researcher*, vol. 43, no. 4, pp. 196–200, 2014.

[5] Y. Wu, X. Zhu, and Q. Yue, "Optimal few-weight codes from simplicial complexes," *IEEE Transactions on Information Theory*, vol. 66, no. 6, pp. 3657–3663, 2020.

[6] E. P. M. Brouwers, M. C. W. Joosen, C. van Zelst, and J. Van Weeghel, "To disclose or not to disclose: a multi-stakeholder focus group study on mental health issues in the work environment," *Journal of Occupational Rehabilitation*, vol. 30, no. 1, pp. 84–92, 2020.

[7] J. H. M. Evers, S. C. Hille, and A. Muntean, "Mild solutions to a measure-valued mass evolution problem with flux boundary conditions," *Journal of Differential Equations*, vol. 259, no. 3, pp. 1068–1097, 2015.

[8] C. Pelau and A. C. Chinie, "Econometric model for measuring the impact of the education level of the population on the recycling rate in a circular economy," *www.amfiteatreconomic.ro*, vol. 20, no. 48, pp. 340–355, 2018.

[9] N. Beadie, "Toward a history of education markets in the United States: an introduction," *Social Science History*, vol. 32, no. 1, pp. 47–73, 2008.

[10] C. Delany, L. Doughney, L. Bandler et al., "Exploring learning goals and assessment approaches for Indigenous health education: a qualitative study in Australia and New Zealand," *Higher Education*, vol. 75, no. 2, pp. 255–270, 2017.

[11] D. A. Hashimoto, G. Rosman, M. Volkov, D. L. Rus, and O. R. Meireles, "Artificial intelligence for intraoperative video analysis: machine learning's role in surgical education," *Journal of the American College of Surgeons*, vol. 225, no. 4, p. S171, 2017.

[12] H. Lankiewicz, E. F. Wasikiewicz, and A. Szczepaniak-Kozak, "Insights into teacher language awareness with reference to the concept of self-marginalization and empowerment in the use of a foreign language," *Porta Linguarum*, vol. 25, no. 25, pp. 147–161, 2016.

[13] C. L. Park and L. Crystal, "Integrating positive psychology into health-related quality of life research," *Quality of Life Research*, vol. 24, no. 7, pp. 1645–1651, 2015.

[14] R. B. King and A. B. I. Bernardo, "Positive emotions predict students' well-being and academic motivation: the broaden-and-build approach," *The Psychology of Asian Learners*, pp. 485–501, Springer, Berlin, Germany, 2016.

[15] D. C. Shin, "Development and application of an in-house health care program to improve the physical and mental health of working mothers: a pilot study," *Health Care for Women International*, vol. 41, no. 3, pp. 284–292, 2020.

[16] H. M. Dang, B. Weiss, C. M. Nguyen, N. Tran, and A. Pollack, "Vietnam as a case example of school-based mental health services in low and middle income countries: efficacy and effects of risk status," *School Psychology International*, vol. 38, no. 1, pp. 22–41, 2017.

[17] C. Staff, "Artificial intelligence," *Communications of the ACM*, vol. 60, no. 2, pp. 10-11, 2017.

[18] M. J. Lamberti, M. Wilkinson, B. A. Donzanti et al., "A study on the application and use of artificial intelligence to support drug development," *Clinical Therapeutics*, vol. 41, no. 8, pp. 1414–1426, 2019.

[19] H. W. Melvin and C. Campbell, "Toward a sociology of irreligion," *Review of Religious Research*, vol. 16, no. 2, p. 147, 1975.

[20] C. María, "Intercultural communication: a reader," *Estudios Filologicos*, vol. 25, no. 39, pp. 270–272, 2015.

[21] F. H. Chang, N. K. Latham, R. H. Friedman, and A. M Jette, "Interactive voice response version of the late-life function and disability instrument," *Journal of the American Geriatrics Society*, vol. 63, no. 4, pp. 770–775, 2015.

[22] A. Sasou, "An evaluation of the automatically topology-generated auto-regressive hidden Markov model with regard to an esophageal voice enhancement task," *Journal of the Acoustical Society of America*, vol. 140, no. 4, p. 2964, 2016.

[23] S. M. Schueller and A. C. Parks, "The science of self-help: translating positive psychology research into increased individual happiness," *European Psychologist*, vol. 19, no. 2, pp. 145–155, 2014.

[24] A. Nwoye, "African psychology and the emergence of the Madiban tradition," *Theory & Psychology*, vol. 28, no. 1, pp. 38–64, 2018.

[25] L. S. Aloia and D. H. Solomon, "The physiology of argumentative skill deficiency: cognitive ability, emotional competence, communication qualities, and responses to conflict," *Communication Monographs*, vol. 82, no. 3, pp. 315–338, 2015.

[26] R. R. Yager and R. Mesiar, "On the transformation of fuzzy measures to the power set and its role in determining the measure of a measure," *IEEE Transactions on Fuzzy Systems*, vol. 23, no. 4, pp. 842–849, 2015.

[27] V. Hermelingmeier and K. A. Nicholas, "Identifying five different perspectives on the ecosystem services concept using Q methodology," *Ecological Economics*, vol. 136, no. JUN, pp. 255–265, 2017.

[28] P. M. Wannier, R. Günter, and P. Sonderegger, "Expression of neuroserpin in the visual cortex of the mouse during the developmental critical period," *European Journal of Neuroence*, vol. 17, no. 9, pp. 1853–1860, 2015.

[29] A. Pfahlberg, K. F. Kölmel, and O. Gefeller For The Febim Study Group, "Timing of excessive ultraviolet radiation and melanoma: epidemiology does not support the existence of a critical period of high susceptibility to solar ultraviolet radiation- induced melanoma," *British Journal of Dermatology*, vol. 144, no. 3, pp. 471–475, 2001.

[30] J. Barth, K. Andrieu, E. Fotiadis, G. Hannink, R. Barthelemy, and M. Saffarini, "Critical period and risk factors for retear following arthroscopic repair of the rotator cuff," *Knee Surgery, Sports Traumatology, Arthroscopy*, vol. 25, no. 7, pp. 2196–2204, 2017.

[31] L. H. Hsia, I. Huang, and G. J. Hwang, "A web-based peer-assessment approach to improving junior high school students' performance, self-efficacy and motivation in performing arts courses," *British Journal of Educational Technology*, vol. 47, no. 4, pp. 618–632, 2015.

[32] Y. Li, Q. Lin, C. Qian, and S. Zhao, "Research on evaluation model for interactive classroom enabled with mobile terminals," *Interactive Learning Environments*, vol. 27, no. 2, pp. 163–180, 2019.

[33] W. Zheng, L. Yin, X. Chen, Z. Ma, S. Liu, and B. Yang, "Knowledge base graph embedding module design for Visual question answering model," *Pattern Recognition*, vol. 120, Article ID 108153, 2021.

[34] J. Li, K. Xu, S. Chaudhuri, E. Yumer, H. Zhang, and L. Guibas, "Grass: generative recursive autoencoders for shape structures," *ACM Transactions on Graphics*, vol. 36, no. 4, pp. 1–14, 2017.

[35] H. Tian, Y. Qin, Z. Niu, L. Wang, and S. Ge, "Summer maize mapping by compositing time series sentinel-1A imagery based on crop growth cycles," *Journal of the Indian Society of Remote Sensing*, vol. 49, no. 11, pp. 2863–2874, 2021.

[36] J. Zhang, C. Zhu, L. Zheng, and K. Xu, "ROSEFusion: random optimization for online dense reconstruction under fast camera motion," *ACM Transactions on Graphics*, vol. 40, no. 4, pp. 1–17, 2021.

[37] W. Zheng and L. Yin, "Characterization inference based on joint-optimization of multi-layer semantics and deep fusion matching network," *PeerJ Computer Science*, vol. 8, Article ID e908, 2022.

[38] H. Zhao, C. Zhu, X. Xu, H. Huang, and K. Xu, "Learning practically feasible policies for online 3D bin packing," *Science China Information Sciences*, vol. 65, no. 1, pp. 112105–112117, 2022.

[39] W. Zheng, X. Tian, B. Yang et al., "A few shot classification methods based on multiscale relational networks," *Applied Sciences*, vol. 12, no. 8, Article ID 4059, 2022.

[40] Y. Zhang, X. Shi, H. Zhang, Y. Cao, and V. Terzija, "Review on deep learning applications in frequency analysis and control of modern power system," *International Journal of Electrical Power & Energy Systems*, vol. 136, Article ID 107744, 2022.

[41] X. Xu, D. Niu, B. Xiao, X. Guo, L. Zhang, and K. Wang, "Policy analysis for grid parity of wind power generation in China," *Energy Policy*, vol. 138, Article ID 111225, 2020.

[42] G. Liu, "Data collection in mi-assisted wireless powered underground sensor networks: directions, recent advances, and challenges," *IEEE Communications Magazine*, vol. 59, no. 4, pp. 132–138, 2021.

[43] Y. Shen, N. Ding, H. T. Zheng, Y. Li, and M. Yang, "Modeling relation paths for knowledge graph completion," *IEEE Transactions on Knowledge and Data Engineering*, vol. 33, no. 11, pp. 3607–3617, 2021.

[44] W. Zheng, X. Liu, and L. Yin, "Sentence representation method based on multi-layer semantic network," *Applied Sciences*, vol. 11, no. 3, Article ID 1316, 2021.

[45] Z. Lv, D. Chen, H. Feng, W. Wei, and H. Lv, "Artificial intelligence in underwater digital twins sensor networks," *ACM Transactions on Sensor Networks*, vol. 18, no. 3, pp. 1–27, 2022.

[46] W. Zheng, X. Liu, X. Ni, L. Yin, and B. Yang, "Improving visual reasoning through semantic representation," *IEEE Access*, vol. 9, pp. 91476–91486, 2021.

Hindawi
Computational Intelligence and Neuroscience
Volume 2022, Article ID 5508365, 11 pages
https://doi.org/10.1155/2022/5508365

Research Article

A Deep Neural Network-Based Model for Quantitative Evaluation of the Effects of Swimming Training

Jun-Jie Hou,[1,2] **Hui-Li Tian**,[1] **and Biao Lu**[3]

[1]*School of Physical Education, Suzhou University, Suzhou 234000, China*
[2]*Philosophy of Sports, University of Perpetual Help System DALTA, Manila 0900, Philippines*
[3]*Information Engineering Department, Suzhou University, Suzhou 234000, China*

Correspondence should be addressed to Hui-Li Tian; tianhuili@ahszu.edu.cn

Received 7 August 2022; Revised 6 September 2022; Accepted 16 September 2022; Published 30 September 2022

Academic Editor: Ning Cao

This paper analyzes the quantitative assessment model of the swimming training effect based on the deep neural network by constructing a deep neural network model and designing a quantitative assessment model of the swimming training effect. This paper addresses the problem of not considering the influence of the uncertainties existing in the virtual environment when evaluating swimming training and adds the power of the delays in the actual training operation environment, which is used to improve the objectivity and usability of swimming training evaluation results. To better measure the degree of influence of uncertainties, a training evaluation software module is developed to validate the usability of the simulated training evaluation method using simulated case data and compare it with the data after training evaluation using the unimproved evaluation method to verify the correctness and objectivity of the evaluation method in this paper. In the experiments, the feature extractor is a deep neural network, and the classifier is a gradient-boosting decision tree with integrated learning advantages. In the experimental comparison, we can achieve more than 60% accuracy and no more than a 1.00% decrease in recognition rate on DBPNN + GBDT, 78.5% parameter reduction, and 54.5% floating-point reduction on DPBNN. We can effectively reduce 32.1% of video memory occupation. It can be concluded from the experiments that deep neural network models are more effective and easier to obtain relatively accurate experimental results than shallow learning when facing high-dimensional sparse features. At the same time, deep neural networks can also improve the prediction results of external learning models. Therefore, the experimental results of this model are most intuitively accurate when combining deep neural networks with gradient boosting decision trees.

1. Introduction

Swimming and fitness are mass sports that are gaining more attention in the general environment of national fitness. Competitive swimming is the first sport to be officially included in the Olympic Games. At the present stage, the teaching and training of competitive swimming are still mainly based on education by example, that is, through the demonstration of the coach's movements and the self-correction of the students to achieve teaching and training [1]. The teaching model is not real time, the student's observation of the coach's standard movements is not comprehensive, and it is difficult for the coach to judge the correctness of the students' body movements scientifically, so the training efficiency is low. It enables athletes to conduct strength training and posture simulation on land, and students and coaches can observe and communicate technical movements in real time [2]. However, when athletes use swimming simulator training, the correctness of their sports posture still relies on the coach's manual judgment, and there is a lack of natural swimming sensation in the visual and physical aspects. The above status quo requires competitive swimming training to introduce advanced technology to scientifically analyze and optimize athletes' technical movements and provide a complete real swimming experience to keep athletes' training motivated.

Deep learning has become an important research direction of machine learning in recent years. Compared with traditional machine learning algorithms, deep learning algorithms do not need to go through feature engineering to obtain good enough data features, its deep structure can complete the learning process of components by itself, and

there are enough trainable parameters, which makes deep learning able to train many samples with high accuracy [3]. Convolutional neural network (CNN) is one of the representative algorithms of deep understanding, whose basic structure consists of a convolutional layer, pooling layer, activation layer, and fully connected layer. Finally, the probability of each category is calculated by the fully connected layer and the classifier to realize the recognition of images [4]. The complete convolutional neural network is based on the CNN with the removal of the fully connected layer, using the deconvolution to enlarge the reduced feature map, recovering the detailed information by fusing the feature map obtained from the convolutional layer, and finally calculating the probability of each pixel belonging to each category on the feature map recovered to the input size to realize the semantic segmentation of the image [5]. Although deep neural networks have higher accuracy, they require a much larger number of labeled samples as support, and the training cost is enormous. Production of remote sensing image datasets involves a series of operations, such as correcting remote sensing images and then cropping and manual labeling, which consumes a lot of time and labor. And the use of small datasets for training the network is prone to over-fitting or under-fitting.

Deep learning methods have achieved good performance in many application domains. Still, until now, people have not been able to theoretically explain its effectiveness and consider it a black box model. On the other hand, the design and application of deep neural network structures are in urgent need of solid theories to guide them, but none of them has been accepted by most researchers yet. Despite the strong model representation capability of deep neural networks, the training optimization of deep neural network models is generally considered difficult [6]. The training optimization problem of deep neural networks is a high-dimensional nonconvex optimization problem, which is very different from the traditional machine learning optimization and the general mathematically studied nonconvex optimization problems [7]. Deep neural network training optimization is characterized by large scale, extensive data, and high-dimensional nonconvexity. Practical deep neural network models are often over-parameterized, which is very different from traditional machine learning optimization. Deep neural network models are large in scale, even tens of millions of dimensions, and mathematically studied nonconvex optimization theories generally do not address such high dimensions [8]. Note that the training optimization problem of deep neural networks includes many related issues, not only referring to optimization algorithm design and convergence analysis. In this paper, the relationship between swimming training and deep neural networks is explored to identify the connection between the two to determine the effectiveness of swimming training [9]. It is investigated using comparative experiments to analyze the influence of swimming training methods on the training effect. According to the research of this paper, it is beneficial to guide swimming teaching activities and provide technical reference for swimming coaches and has important practical significance for the development of swimming activities making full use of the sensitive period of physical quality development of swimmers as a window of opportunity for physical quality development, combine with the characteristics of swimming sport itself, introduce the concept of functional movement training in the process of physical training of swimmers, consolidate the material foundation, and optimize the training content. Through the modular design of training content, the physical training of swimmers can be scientific, systematic, and sustainable.

2. Related Works

Swimming training research gradually formed with the emergence of competitive swimming in the world competition in the late 1950s, which made a bare theoretical pavement for the development of swimming and made significant contributions to the development of swimming [10]. In recent years, academic research on swimming training has stagnated and developed during the development of swimming. Swimming training is in continuous product and improvement, both in theoretical research and practical activities [11]. Therefore, many experts, scholars, and front-line coaches are interested in exploring this field. The historical evolution of the hotspots of swimming training research as the direction of this study is a rational choice based on the unique charm of swimming training research and my strong interest in swimming training research.

The integration between disciplines is influenced by the development of information science and technology, and many scientific and technological developments also affect the outcome of swimming training. Still, the overall effect is always on the path of scientific, systematic, and comprehensive. The study attempts to summarize the development path and laws of current swimming training research and form a periodic theoretical system to provide an academic reference for future swimming training research [12]. At the same time, theory guides practice, and the theoretical development of swimming training has important guiding significance for swimming training activities. Based on the evolution law of the hotspots of swimming training theory and practice research, the future development direction of swimming training research is inferred by combining the characteristic properties of the swimming sport itself, the content system of swimming training, and the structural characteristics of swimming training [13, 14]. At the same time, to improve the training level of swimming training practice activities, it is necessary to master the theoretical development of swimming training. Gao et al. and Zhang et al. studied how swimming improved the nation's quality and mentioned that the water environment in which swimming exercise occurs is very different from the environment in which other types of exercise occur. Swimming, as a unique and attractive sport, not only has muscular fitness and fitness but also has the function of exercising people's will quality compared with other sports [15, 16]. Gao concluded from two months of swimming exercise that swimming could effectively change the body shape and various functions of college students under certain intensity [15]. Pawan et al. and Dewangan and Sahu tell us that swimming has specific fitness

effects for different social groups by talking about the impact of swimming on human health [17, 18].

Deep neural networks (DNNs) have become one of the most popular research directions due to the excellence of deep learning methods, which is a further enhancement of artificial neural networks (ANNs) in the field of machine learning [19, 20]. This paper demonstrates that a neural network with multiple implicit layers can learn features better and obtain better results. Still, the learned features must have great typicality to help visualize and classify data clearly [21, 22]. Since then, the status of BP neural networks has been high, and many experts and researchers have taken deep learning methods as one of the leading research objects. Universities have also set up special research institutes to investigate deep neural networks' characteristics, trial scenarios, and advantages and disadvantages. Up to now, the theoretical research about deep BP neural networks is still in its initial stage except for the corresponding explanation from the bionic perspective. Still, it has achieved excellent results in many applications to perform deep learning of functions [23, 24]. At the same time, its use in image recognition is relatively early and more mature; Lin et al. explains referring to a deep neural network with a convolutional structure inspired by a biological vision model called a convolutional neural network [25]. In an image recognition competition, Hinton and his student Alex Krzyzewski and others used CNNs that benefited from rich data and improved GPU computing power to get the best competitive results. Since then, BP network models have been widely used in graphic images to improve accuracy and reduce the timelines involved in human eigenvalue processing, resulting in a convenient way of image recognition [26, 27].

3. Quantitative Evaluation Model Design of Swimming Training Effect Based on Depth Neural Network

3.1. Depth Neural Network Model Construction. Deep convolutional neural network CNN includes a convolutional layer and pooling layer, a practical recognition method that has become popular in recent years. As shown in Figure 1, the convolutional layer neural network contains two layers: one is the feature extraction layer, and the other is the feature mapping layer. The feature extraction layer includes the convolutional layer and the pooling layer, its unique network layer. The input layer of each neuron is connected to the local domain of the previous layer, which can complete the extraction of local features, and the connection between local features and other features is fixed after extraction. A feature mapping layer is a two-dimensional plane; the neurons on each aircraft can share the weights, and the input data and the convolution layer data are set by themselves; when the convolution operation moves one position at a time, the actual process of its implementation is to complete the first convolution from the initial position, and the convolution kernel and the gray area of the input values in the graph are multiplied bit by bit and then summed up; that is, $\times 1 + 0 \times 2 + 1 \times 3 + 0 \times 2 + 1 \times 3 + 0 \times 4 + 1 \times 3 + 0 \times 4 + 1 \times 5 = 15$, 15 is the first value in the matrix after the convolution operation. The main difference is the different approach to layer definition and depth processing in multilayer perceptron. Deep neural networks mimic how the human brain thinks by first building single-layer neurons layer by layer so that one single-layer network is trained at a time. After all, layers are trained, and a wake-sleep algorithm is used for tuning. Convolutional neural networks are mediated by "convolutional kernels." The same convolution kernel is shared across all images, and the image retains its original position relationship after the convolution operation.

A CNN is a feed-forward neural network with a deep structure that contains many convolutional computational layers, which can provide an end-to-end approach to task analysis. CNN structures typically have multiple sets of learnable parameters and components. These components usually consist of an alternating cascade of four essential network layers: a convolutional layer, a pooling layer, an activation layer, and a regularization layer.

3.1.1. Convolution Layer. The convolution layer further samples the output feature map of the previous layer, so generally, the size of the feature map everywhere will be smaller than the input feature map. To keep the size of the input and output feature maps constant, adding "0" to the edge of the feature map during the convolution process is necessary as a padding compensation operation. In addition, the convolution operation is a convolution kernel sliding step that is not a pixel-by-pixel wagon, sometimes skipping a few pixels, where the sliding width is called stride. Suppose here the input sample data of a three-dimensional tensor $I_{\text{in}} \in \text{RC}_{\text{in}} \times w_{\text{in}} \times h_{\text{in}}$, the convolution operation padding is p, and stride size is s, using a filter of the convolution layer, which is also a 3D tensor $F_{\text{in}} \in \text{RK} \times K \times C_{\text{in}} \times C_{\text{out}}$, assuming that the length and width of the convolution kernel are K, and the width and height of the output feature map after the convolution operation for this sample are

$$w_{\text{out}} = \sum \frac{w_{\text{in}} + k - 2p}{\sqrt{s - K}}, \qquad h_{\text{out}} = \sum \frac{\sqrt{h_{\text{in}} - k + 2p}}{s - K}. \tag{1}$$

The output feature data and the input are three-dimensional tensors, which can be expressed as $I_{\text{out}} \in \text{RC}_{\text{out}} \times W_{\text{out}}$. Here, we use a sample feature map to illustrate that the convolutional neural network is trained with input data, which can significantly accelerate the training efficiency and converge more quickly. Usually, the amount of data for one training is batch, and the selection of batch depends on the sample size and memory conditions, so the input data is a four-dimensional tensor, and the first dimension is the value of the batch.

3.1.2. Pooling Layer. The pooling layer, also known as the sink layer, aims to aggregate the data in the local perceptual field of the input feature map into a target value, which is designed to reduce the complexity of the network model and

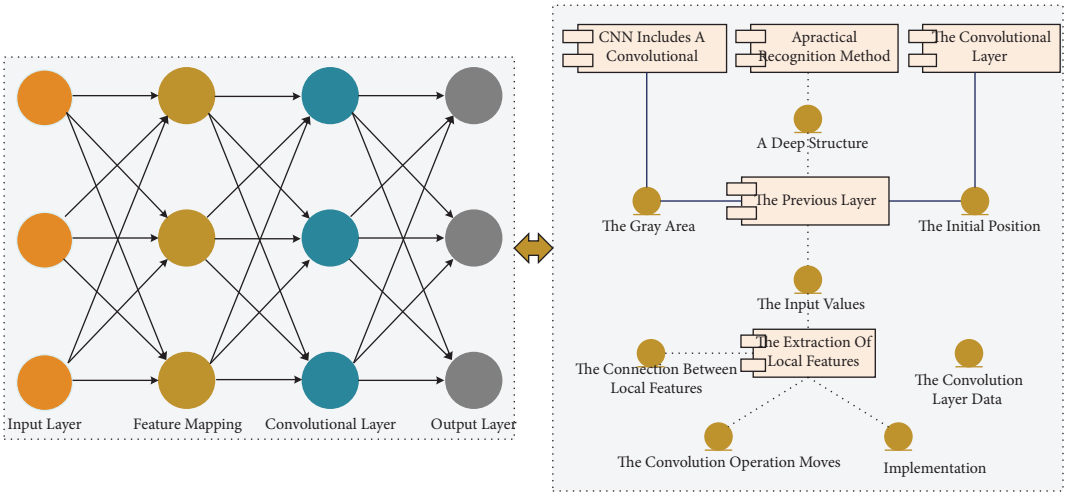


Figure 1: Deep convolutional neural network structure.

avoid overfitting with too many features. The pooling layer is like the convolutional layer in that it uses a convolutional kernel to extract the significant features of a perceptual field region. The only difference with the convolutional kernel of the convolutional layer is that the kernel parameter values are fixed here, and there is no need to use the back-propagation algorithm to update the parameters. Therefore, only the type of pooling layer (average pooling or maximum pooling), the size of the pooling kernel, and the pooling step are required for the training of the neural network, and the average pooling kernel extracts the average and maximum values of the local perceptual field area, respectively. It is defined as

$$f_{(n+1)} = \sum \text{pool}\left(f_j^{n-1} + j\right). \tag{2}$$

3.1.3. Fully Connected Layers. Fully connected layers are generally in the last few layers in modern convolutional neural networks and play the role of feature classification in the whole network. The function of convolutional and pooling layers is to map the input feature representation to the circumscribed feature space. In contrast, fully connected layers map the circumscribed feature representation to the target label space. Earlier, multilayer perceptrons (fully connected networks) consisting of only fully connected layers were frequently used in some simple classification tasks, building a linear mapping of feature transformations. Still, such networks have poor nonlinear capabilities and insufficient processing power under complex tasks, so such networks are now used only in some low-end embedded devices for simple classification tasks.

The data processing process of the fully connected layer is essentially a process of a linear transformation of matrices to vectors, where the parameters that need to be updated using the backpropagation algorithm include the weight parameter matrix and the bias parameter vector. Since the input data of the fully connected layer are vectorized features, assuming that the single sample input feature is $x \in R^m$, the number of output neurons (feature vector length) is n, the weight parameter matrix of the fully connected layer is $a \in Rn^* m$, and the bias parameter vector is $b \in Rn$, and then, the output eigenvector can be expressed as a linear matrix transformation as

$$\text{feature}_{(\text{out}-1)} = \sum \frac{a \times x - b}{\sqrt{ax + b}}. \tag{3}$$

From another perspective, each scalar value of the fully connected layer parameter matrix can be regarded as a convolutional kernel of size 1×1. The fully connected layer can be defined as a convolutional layer when running the network model. For the fully connected layer, the convolutional kernel size can be set to 1×1. For the fully connected layer after the convolutional layer, the convolutional kernel size can be defined as a global convolution of $h \times w$ where h and w are the height and width of the input feature map, respectively. The fully connected layer is usually not very computationally intensive. Still, the number of parameters is very high in the whole network model, and the number of parameters in the fully connected layer can reach more than 95% in the Alex Net and VGG-16 network models. AlexNet is an earlier deep network applied on ImageNet, and its accuracy greatly improved compared to traditional methods. Framework: 5 convolutional layers followed by 3 fully connected layers with ReLU activation function. One improvement of VGG16 over AlexNet is using several consecutive 3×3 convolutional kernels instead of the larger convolutional kernels (11×11, 5×5) in AlexNet. For a given perceptual field (the local size of the input picture for the output), the use of stacked small convolutional kernels is preferable to the use of large convolutional kernels because the multilayer nonlinear layers can increase the network depth to ensure learning more complex patterns, and at a more negligible cost. Still, the computation is less than 10%. With the continuous development of deep learning theory, modern deep neural networks gradually eliminate connected layers or use global average pooling instead of fully connected layers so that the number of layers in the network can be deepened while maintaining a relatively small number of parameters, reducing the training difficulty of the network and the risk of overfitting.

3.1.4. Batch Normalization Layer. Batch normalization can speed up the model's convergence, solve the inherent problem of "gradient dispersion" in deep networks to a certain extent, and avoid the problem of complex parameter updates during backpropagation. In addition, batch normalization applies to deep networks and can play a good role in improving the generalization of traditional shallow simple network models. The batch normalization layer has become essential to almost all neural network models. The specific algorithmic process of batch normalization is mainly divided into four steps; first, for each batch sample data, calculate its overall sample mean and sample variance, assuming that the input batch data are $x = (x^{(1)}, x^{(2)}, ..., x^{(m)})$, which is a feature matrix with sample size m, the statistical characteristics of the batch data.

$$ub = \int_{i=1} \frac{m + x^i}{m - x},$$
$$\sigma b = \int_{i=1}^{m} \frac{\left(x^i - ub\right)^2}{x^i + ub}. \tag{4}$$

After obtaining the batch samples' characteristic mean and characteristic variance, the sample data are batch normalized. The data can be input to the next layer of the network in the form of a standard normal distribution. Suppose the data distribution of each layer of the input network is inconsistent. In that case, the parameters of each layer will have to be adjusted to fit the new data distribution, which undoubtedly increases the difficulty of training the network, a problem known as internal covariate shift. Therefore, we normalize the data so that the input data distribution is the same for each layer.

$$x^i = \sum x^{i-1} + \frac{ub}{\sqrt{\sigma b + e}} + \sqrt{\sigma b - e}. \tag{5}$$

But simply normalizing each layer's features will change the data's original distribution and reduce the feature representation of the samples. The neural network is essentially designed to learn the actual distribution of the features. Finally, the normalized data need to be moderately cheapened to expect an approximate restoration of the distribution of the original features.

$$y^i = \sum r^{x-i} - \beta + y. \tag{6}$$

3.2. Quantitative Assessment Model Design of Swimming Training Effect. Swimming is a kind of sport in the water environment, with the premise of overcoming resistance in the water, using limbs and water interaction, and generating propulsive force to make the body displacement of water sports. Thismobile environment is the most significant and fundamental difference between swimming and other land sports, that is, there is no fixed support in the water. Because the environment is very different, many principles and laws applicable to land human sports cannot be applied to swimming and its unique hydrodynamic principle. That is to reduce the resistance as much as possible while increasing the propulsion resistance, so that as many people as possible can swim in the water. However, due to the difference in physical strength and skills, the degree of pursuing the two is as different as possible. The swimming training process is shown in Figure 2.

Specialized learning of technical movement patterns in the water is required to find relative stability and balance in the water and to exercise the coordination of the neuromuscular system specific to movement in the water, i.e., the same direction, when performed in the water and on land, the coordination between the neuro muscles will change because there is no buoyancy and water resistance on the ground. Therefore, swimmers need to closely integrate physical training with swimming techniques, optimize the power chain as much as possible while maintaining a stable streamlined body posture, and fully exploit and utilize the power of the muscles to generate maximum propulsion. This requires the athlete to have strong enough core area muscles to establish a stable support platform in the water, allowing for coordinated movement of the upper and lower extremities simultaneously to ensure that every part of the body functions efficiently. The aquatic environment in which swimming is located can prevent everyday communication and reduce the ability to regulate auditory and visual senses. Athletes are always relatively independent and need to give full play to their proprioceptive skills. It is difficult for coaches to provide real-time feedback to athletes during training or in the process of learning and improving swimming skills, which brings some hindrance to the formation of motor skills. The neuromuscular or proprioceptive system is very different in the water, so the change of muscle exertion in the water environment makes it much more difficult for athletes to handle the correct body posture in the water and achieve the movements required for competition.

The assessment indicator system plays a normative role as a guideline in training evaluation, making establishing the indicator system critical. In addition to using appropriate assessment methods, establishing a scientifically sound assessment index system is another crucial point in ensuring the accuracy and reliability of assessment results. This is because the index system specifies various assessment points, which are the abstraction and integration of multiple factors for the corresponding training results. One thing is sure; the so-called scientific and reasonable do not exist. To minimize the unreasonableness of the index system, it is necessary to achieve as complete coverage of the training process as possible. By abstracting the commonalities of the specific processes involved in training, through the process of dialectical logic, outline the fundamental features so that according to these features to complete the full coverage of the training process, then, the index system established by these features can be considered to some extent, and it is good coverage of all the assessment points in training. Because the content of the indicator system is highly specialized, in today's swimming training, the establishment of the indicator system often needs to first identify the specific

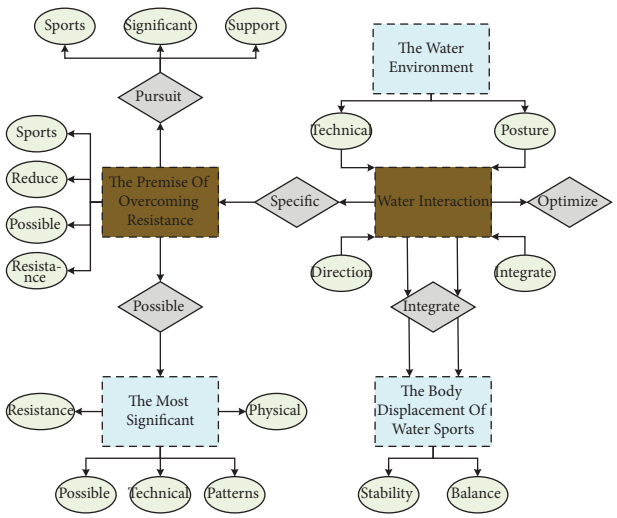


Figure 2: Swimming training process.

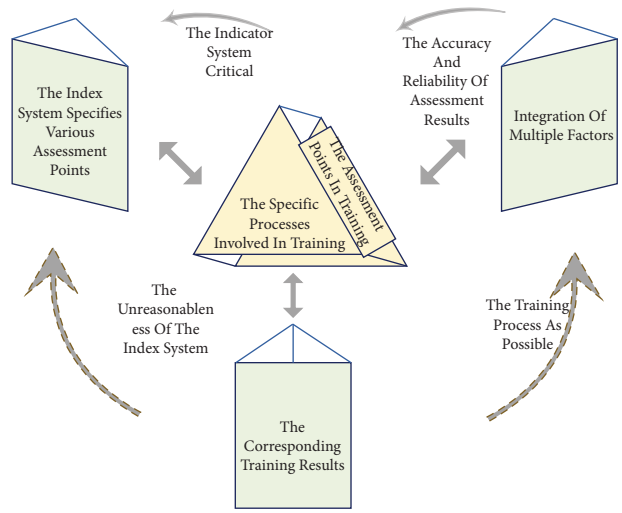


Figure 3: The construction process of the system of quantitative evaluation of swimming training.

type of swimming. After determining the swimming type, cooperate with relevant experts to ensure that the indicator system follows the principles of systematicness, objectivity, development, feasibility and comparability. According to the specific use and protection process of swimming, we analyze and abstract the elements involved in the training. Then, we will continue to discuss, classify, and screen in-depth, and finally determine the phased evaluation index system. The construction process of a quantitative assessment of swimming training is shown in Figure 3.

The assessment index system established through this process generally has comprehensive coverage and is very clear about the professional characteristics of swimming. The indexes continuously evaluated and screened by experts can better reflect the various skill points in using and securing this swimming. However, the attention to numerous details also leads to a more complex indicator system, and the increased complexity directly affects the complexity of solving the final assessment results. In addition, because of the great degree of attention to the details related to a particular swim, and the characteristics of different swims vary significantly from one another, it is necessary to reprofessionalize the analysis of other swims when establishing the index system; that is, it is required to develop other index systems for different swims. Still, it dramatically increases the difficulty of standardization and does not have generality. The training assessment system may be entirely different for other swims, which also increases the cost of training assessment. This project conducted relevant decision research to solve the above problems in the training evaluation system. The most prominent feature of the general index system in the establishment process is the abstraction of the training process according to the characteristics of a specific swim, which is the direct cause of the poor generality of the final system. Therefore, this paper considers changes in this point to make the established index system in line with the conception, in the abstraction of the swimming use and security process, not focus on the details and characteristics of each swim, but from a broad perspective to judge the training process.

The training was performed using supervised machine learning methods to provide a quantitative assessment of swimming functionality using three machine learning models: linear regression, support vector machine (SVM), and support vector regression (SVR), respectively. A more refined FNGS2.0 score was used to make the output more accurate regarding the label data and production of the pretrained models. The input vectors are [SymBrow, SymEye, SymMouth], and the output is the FNGS2.0 score. The pre-trained resulting model is used in the system to fit the quantitative data processed in real time. In this way, the system automatically and efficiently outputs the HBGS scores of swimming training by feeding the captured swimming training data in real time. Since our system is developed on a cell phone, athletes can assess at any time after swim training and obtain a real-time assessment of the swim training function.

4. Analysis of Results

4.1. Analysis of Quantitative Evaluation Model of Swimming Training Effect Based on Deep Neural Network. A deep convolutional neural network contains several convolutional layers. To find the optimal network structure, the change in the number of neurons and the dropout ratio is analyzed by applying network search to the first convolutional layer (first defining the search space and the controller generating the network structure). The test results are recorded as shown in Figure 4. It can be observed that the dropout divergence trend is most apparent when there are 30 neurons, and the maximum AUC value in the test range appears at this time. Later, when the number of neurons was increased, the trend of Dropout0.4 and Dropout0.5 tended to coincide.

The optimal parameters of the first convolutional layer are obtained through the above experiments, and then, the other structural layers of the network perform the tuning operation; the number of neurons per layer is 10–100, the number of convolutional kernels is 2 or 6, the step size is 10,

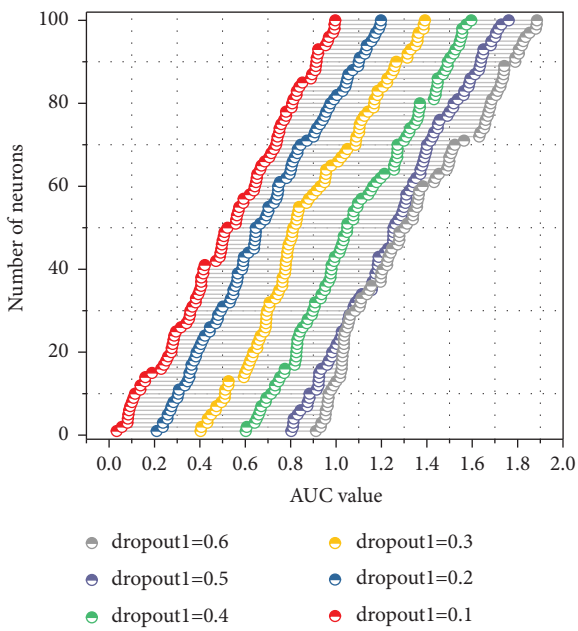


Figure 4: AUC values under the change of convolutional layer parameters.

and the number of relatively optimal parameter units is 50. The number of convolutional kernels is 2. The convolutional neural network's convolutional kernels and pooling layers are changed using the adjusted parameters. Since the pooling layer in CNN has no parameters, dropout's critical role in the model is built. The pooling layer is not included in the model, and the dropout key role is in this model. The adjustment results are shown in Table 1.

To improve the algorithm accuracy, enhance the model generalization ability, and prevent overfitting, another factor affecting the model, regularization, is studied in-depth, and L2 regularization is added to the first layer of CNN in the three major feature extractors of NLP (by reducing the secondary dimensional weights in L2). As the value of L2 regularization increases from zero, the AUC values become unstable, and the model is poorly structured. The overfitting problem of the model is dealt with when dropout tuning is performed on the model, and there is no need to deal with it by L2 regularization.

The mathematical and statistical method focuses on the organization and statistical analysis of the experimental data, which will be counted in an Excel sheet. Programming support: SPSS supports rich data sources and has full data access and management capabilities as well as programming capabilities. Users only need to understand the principles of statistical analysis, and even if they are not familiar with the algorithms of statistical methods, they can quickly obtain the statistical analysis results they need, and it also supports secondary development. Powerful functions: SPSS has complete parts of data input, editing, statistical analysis, report and graph production, etc. It provides a comprehensive data analysis process and covers a full range of statistical analysis methods, such as exploratory analysis of data, partial correlation, analysis of variance, nonparametric test, multiple regression, logistic regression. And the collated

Table 1: Experimental results after adjusting the convolution kernel and pooling layer.

Parameter adjustment	Dropout	AUC
Optimal results	0.1763	0.7788
Convolution kernel = 1	0.1976	0.7695
Convolution kernel = 2	0.1884	0.7742
Convolution kernel = 3	0.1683	0.7681
Convolution kernel = 4	0.1739	0.7735
Convolution kernel = 5	0.197	0.7883
Convolution kernel = 6	0.1749	0.778
Convolution kernel = 7	0.1954	0.7748
Convolution kernel = 8	0.1722	0.7885

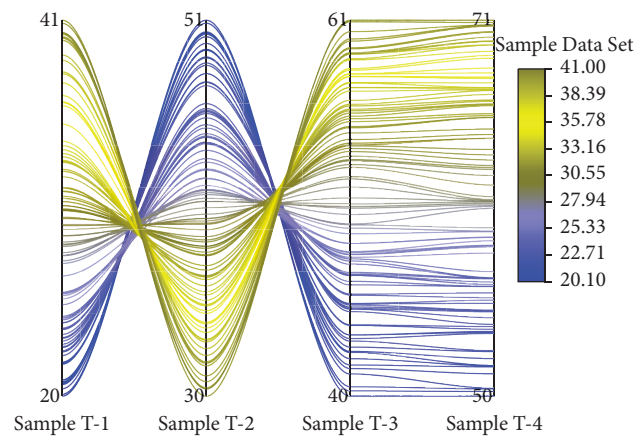


Figure 5: Swimming training effect data test.

data were analyzed using SPSS 20.0 with independent samples t-test. According to the experiment results, the mean ± standard was used to record and characterize the effects. At the same time, the data of some experimental items were organized and analyzed using P-value analysis to determine the correlation of the experimental results. The swimming training effect data test is shown in Figure 5.

Model evaluation is integral to the model development testing and validation process. It effectively discovers the optimal model for expressing the data and how well the model works. In this paper, a confusion matrix is chosen to evaluate the prediction results of the model. The confusion matrix is an intuitive presentation tool for evaluating classification models that the number of correct and incorrect predictions output by the classification model on the data, where each column of the matrix is a sample prediction of the model's forecast, and each row of the matrix is the sample reality.

The real-time data sending and receiving directly affects the immersion of the user's swimming experience; the shorter the delay, the lower the user's vertigo and the stronger the sense of reality. Two methods can test real time; one is to observe the smoothness of the field of view update through the user's experience and directly feel the size of the delay of the area of view update; the smoother the lot of view update, the lower the delay, the better the real time. However, this method lacks a quantitative description of the real-time performance. Another way calculates the frame

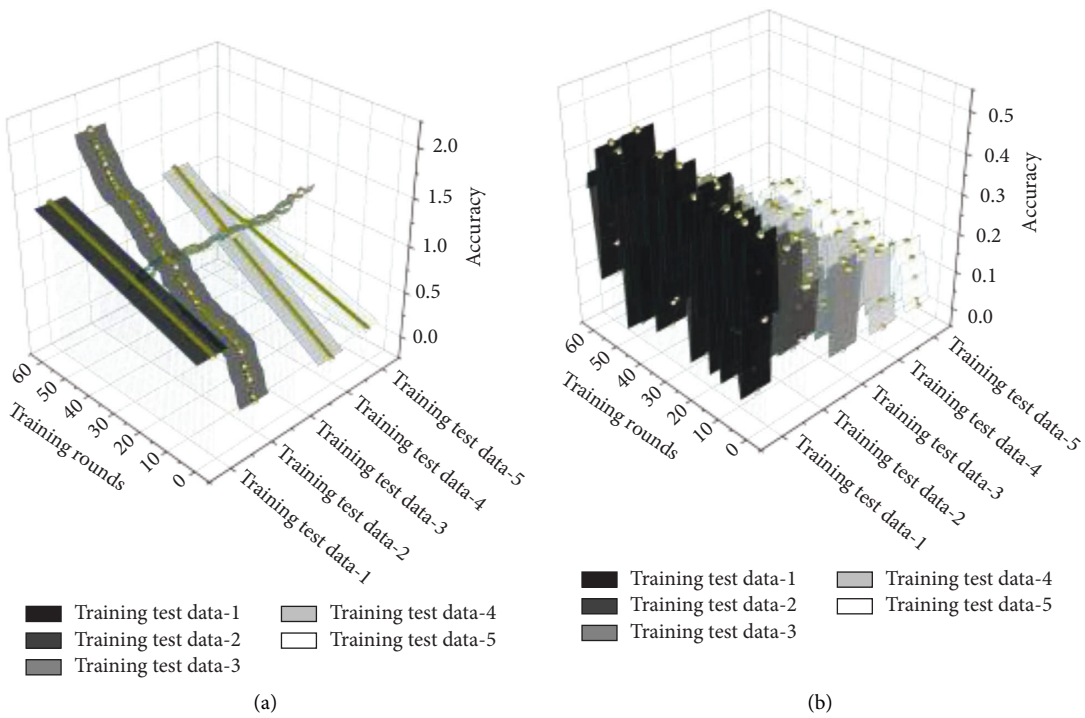


Figure 6: Comparison of model evaluation graphs. (a) Training rounds. (b) Training rounds.

rate by the number of frames of bit-pose data received by the virtual scene program in a fixed time. For Unity3D, the higher the scene refresh frame rate is, the smoother and more natural the motion display is, and a soft motion effect is achieved when the frame rate is above 30 fps. A comparison of the model evaluation graph is shown in Figure 6.

4.2. Quantitative Evaluation Model Implementation of Swimming Training Effect. The effectiveness of the proposed weight-based scaling invariant neural network integration method is verified through experiments on several datasets and various popular deep neural network models. A comparative analysis is performed with recently proposed deep neural network integration methods; the dropout value in the hidden layer is taken as 0.35, the number of neurons and other influencing parameters are taken as the optimal parameters tested in the past, and the input layer in the LR and GBDT models are taken as the features of the third hidden layer, and the dimensionality of the run to the features of the model is 50, which is 301 lower than the original dimensionality, and the specific dimensionality values are combined with linear function and nonlinear function, and the results are obtained after multiple layers of operations, which can be used as a reference standard to reduce the computational load of the model data.

In this CNN model, we choose 1D convolution Conv1D, which is only convolution in the width direction, and do not use a 2D array as the output representation, also known as feature map Conv2D convolution, the width of 1D convolution is 3, not 3×3 2D kernel array, 1D convolution is often used for sequence data, and the input of Conv1d is 3D data. Two-dimensional convolution is commonly used in computer vision and image processing. This model building does not involve feature map convolution, so the one-dimensional convolution Conv1D, which is suitable for this study, is used. Here, in LR and GBDT, input has been established with the upper layer link fully connected layer of features, observation can be seen that the new test dimension (80) only accounts for about 50% of the original size (351), and the features are extracted at the same time the dimensionality reduction is completed. The comparison of AUC before and after the model combination is shown in Figure 7.

The established combined model is debugged, and the obtained data are compared with the single model parameters. The gradient boosting decision tree generates N trees, and after receiving a sample message, each tree goes from the node of the root to the node of the leaf, and when it reaches the leaf node, it is 1 or 0, point or no point. The output value of each tree can be regarded as the output feature, taking the value of 0 or 1, a total of N trees, each tree i has Mi leaves can be seen as having M kinds of combinations, a tree corresponds to a one-hot encoding method, a total of multiple dimensions of new features, as the input vector into the LR model, the output of the result. The above operation is used to verify two things that the experimental results of the established deep learning combinations are generally better than the ordinary single model and to confirm that the deep learning models based on deep neural networks can improve the shallow learning methods of the single model. Collating the experimental results and comparing the metrics, it is observed that the combined models, such as (GBDT + DBPNN), are better than the general models, such as LR/GBDT/. It indicates that the collaborative models established in this study to enhance learning efficiency are

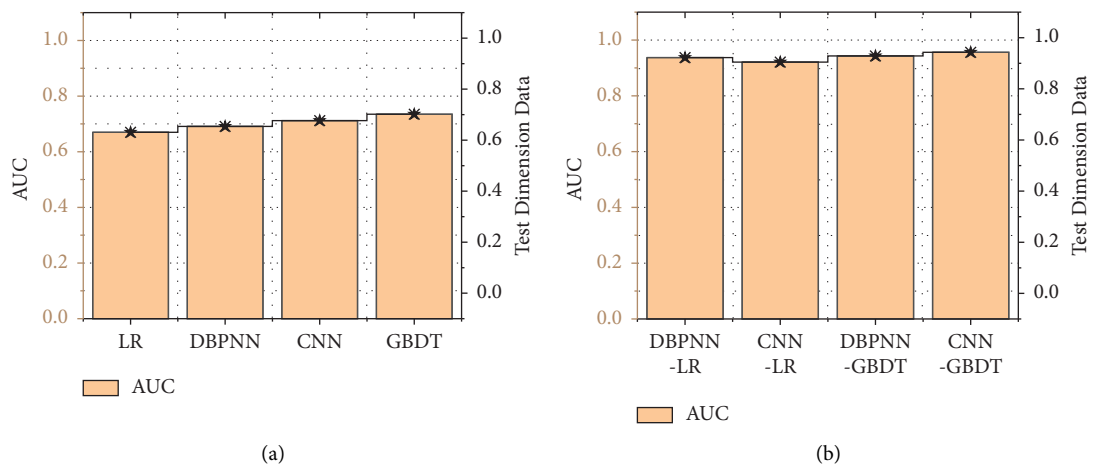


Figure 7: Comparison of AUC before and after model combination. (a) Before model combination. (b) After model combination.

effective and practical. During the experiment, it is relatively challenging to seek the optimal parameters of CNN, but its advantages as feature parameters are also apparent. The evaluation results obtained from several sets of training data by the evaluation method in this paper are compared with the evaluation results in the ideal training state without considering the uncertainties, and the corresponding practical operation results are shown in Figure 8.

The existing index system for training assessment is highly correlated with swimming training, making it necessary to establish a different index system adapted to current swimming for additional training due to the other characteristics and assessment concerns between their swimming training. To solve the abstraction problem to some extent, this paper proposes an assessment method decoupled from specific swims, blurring the guarantee and use characteristics of swim training and obtaining an upper-level index system after the abstraction of assessment points. In the implementation of the assessment, the lower layer can be based on the characteristics of the current swimming training with targeted training project assessment and the assessment of the results of the data obtained from the assessment of categorization and integration, such as the assessment of the use of a training multiple ways of operation together as the assessment of swimming training use operation results, the formation of the input data in line with the structure of the upper layer index system. Such an approach makes it unnecessary to establish separate index systems for assessing each swimming training in the bottom layer. Still, it also ensures the accuracy of the training assessment because the bottom layer is carried out according to the subdivision of different training. The general training assessment only considers the assessment of each technical point in the swimming protection training, ignoring the differences between the simulated training and the operational training in the virtual environment caused by external uncertainties, thus making the results of the mock training assessment may be somewhat different from the results of the actual operation. On the other hand, the use of hierarchical analysis can more clearly describe and quantify the factors in the training assessment, but dynamic changes in the uncertainty of

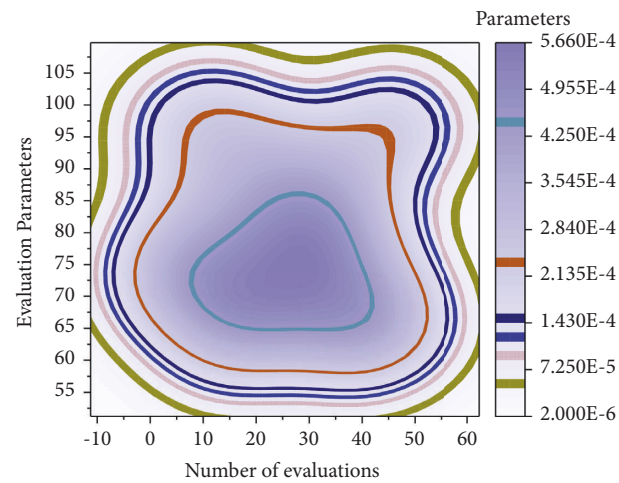


Figure 8: Comparison of training effect evaluation results.

elements cannot be directly quantified. Therefore, this paper proposes using deep neural networks for projection to assist the quantification process of the indeterminate factors, update the index system to match them, and finally get the assessment results closer to the actual operation level through the projection of hierarchical analysis.

5. Conclusion

The development of swimming training is the process of human beings in the continuous understanding of the laws of swimming movement and the physiological and psychological changes of human activity. The emergence of competitive swimming as a social phenomenon has pushed the curtain on swimming training. This paper studies the quantitative evaluation of the swimming training effect based on a deep neural network. DBPNN + GBDT is selected as the algorithm basis for model building, and swimming training test data are introduced for evaluation. The DPBNN used in this paper combines the advantages of the BP neural network and deep learning network algorithms, which enhances the model's massively parallel, distributed

processing, self-organization, and self-learning capabilities and plays a significant role in performing also positive extraction and high-dimensional data processing in terms of automation degree, convolution, and sharing. After completing the deep neural network construction, the fusion gradient boosting decision tree model broadens the data types that can be processed to improve model prediction accuracy further. In this paper, based on the existing evaluation methods, the overall evaluation method and the corresponding index system are improved to establish a higher-level evaluation index system to shield the different characteristics of each equipment training, given the poor generality of the evaluation and the lack of consideration of the fundamental environmental factors in the evaluation results of the simulation training. At the same time, the effectiveness of the proposed evaluation method is verified by introducing a machine learning method that can obtain better data training results in small samples and dynamically obtain the influence weights of variable factors to achieve the coverage and consideration of external factors in the evaluation. The current initialization methods of deep neural networks are overwhelming. Most of them are symmetric random initialization, and the so-called symmetric random initialization means that all the parameters of the neural network ownership value are drawn from some distribution of density distribution about the origin, and the commonly used Gaussian random initialization is the symmetric random initialization. The standard Gaussian random initialization is symmetric random initialization. The activation functions of most of the popular deep network models are Relu functions because their activation regions are not symmetric. The activation for negative values of the input is 0.

Data Availability

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

Conflicts of Interest

The authors declare that they have no conflicts of interest or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work was supported by the Nonfinancial Research Project of Suzhou University (No. 2022xhx143).

References

[1] L. Zhou, H. Yu, and Y. Lan, "Deep convolutional neural network-based robust phase gradient estimation for two-dimensional phase unwrapping using SAR interferograms," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 58, no. 7, pp. 4653–4665, 2020.

[2] J. Huang, J. H. C. Lai, K. H. Tse et al., "Deep neural network based CEST and AREX processing: application in imaging a model of Alzheimer's disease at 3 T," *Magnetic Resonance in Medicine*, vol. 87, no. 3, pp. 1529–1545, 2022.

[3] S. Li, L. Zhang, H. Feng et al., "MutagenPred-GCNNs: a graph convolutional neural network-based classification model for mutagenicity prediction with data-driven molecular fingerprints," *Interdisciplinary Sciences: Computational Life Sciences*, vol. 13, no. 1, pp. 25–33, 2021.

[4] R. Leenhardt, M. Souchaud, G. Houist et al., "A neural network-based algorithm for assessing the cleanliness of small bowel during capsule endoscopy," *Endoscopy*, vol. 53, no. 9, pp. 932–936, 2021.

[5] L. Xiang, G. Guo, J. Yu, and P. Yang, "A convolutional neural network-based linguistic steganalysis for synonym substitution steganography," *Mathematical Biosciences and Engineering*, vol. 17, no. 2, pp. 1041–1058, 2020.

[6] Y. Yao, J. Wang, P. Long, and M. Xie, "Small-batch-size convolutional neural network based fault diagnosis system for nuclear energy production safety with big-data environment," *International Journal of Energy Research*, vol. 44, no. 7, pp. 5841–5855, 2020.

[7] H. Zamyad, N. Naghavi, R. Godaz, and R. Monsefi, "A recurrent neural network-based model for predicting bending behavior of ionic polymer-metal composite actuators," *Journal of Intelligent Material Systems and Structures*, vol. 31, no. 17, pp. 1973–1985, 2020.

[8] J. Lyu and S. Manoochehri, "Online convolutional neural network-based anomaly detection and quality control for fused filament fabrication process," *Virtual and Physical Prototyping*, vol. 16, no. 2, pp. 160–177, 2021.

[9] S. Hou and W. Guo, "Fault identification method for distribution network based on parameter optimized variational mode decomposition and convolutional neural network," *IET Generation, Transmission & Distribution*, vol. 16, no. 4, pp. 737–749, 2022.

[10] A. Mazumdar and P. K. Bora, "Siamese convolutional neural network-based approach towards universal image forensics," *IET Image Processing*, vol. 14, no. 13, pp. 3105–3116, 2020.

[11] Z. Hong, X. Zeng, L. Wei, and X. Liu, "Identifying enhancer-promoter interactions with neural network based on pre-trained DNA vectors and attention mechanism," *Bioinformatics*, vol. 36, no. 4, pp. 1037–1043, 2020.

[12] J. Wang, S. Wang, S. Liu, and Q. Cheng, "Determination of four pesticides in honeysuckle by Echo State Network based on excitation-emission matrix fluorescence spectroscopy technique," *Journal of Food Measurement and Characterization*, vol. 16, no. 1, pp. 431–439, 2022.

[13] Z. Shi, Y. Feng, M. Zhao, and L. He, "A joint deep neural networks-based method for single nighttime rainy image enhancement," *Neural Computing & Applications*, vol. 32, no. 7, pp. 1913–1926, 2020.

[14] K. B. Park and J. Y. Lee, "SwinE-Net: hybrid deep learning approach to novel polyp segmentation using convolutional neural network and Swin Transformer," *Journal of Computational Design and Engineering*, vol. 9, no. 2, pp. 616–632, 2022.

[15] Y. Gao, Y. Chen, H. Feng, Y. Zhang, and Z. Yue, "RicENN: prediction of rice enhancers with neural network based on DNA sequences," *Interdisciplinary Sciences: Computational Life Sciences*, vol. 14, no. 2, pp. 555–565, 2022.

[16] L. Zhang, J. Zhang, J. Ma, and X. Jia, "SC-PNN: saliency cascade convolutional neural network for pansharpening," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 59, no. 11, pp. 9697–9715, 2021.

[17] S. J. Pawan, R. Sankar, A. Jain et al., "Capsule Network-based architectures for the segmentation of sub-retinal serous fluid in optical coherence tomography images of central serous chorioretinopathy," *Medical, & Biological Engineering & Computing*, vol. 59, no. 6, pp. 1245–1259, 2021.

[18] D. K. Dewangan and S. P. Sahu, "RCNet: road classification convolutional neural networks for intelligent vehicle system," *Intelligent Service Robotics*, vol. 14, no. 2, pp. 199–214, 2021.

[19] A. R. Aslam and M. A. B. Altaf, "A 10.13μj/classification 2-channel deep neural network based SoC for negative emotion outburst detection of autistic children," *IEEE Transactions on Biomedical Circuits and Systems*, vol. 15, no. 5, pp. 1039–1052, 2021.

[20] S. Rahman and D. A. Robertson, "Classification of drones and birds using convolutional neural networks applied to radar micro-Doppler spectrogram images," *IET Radar, Sonar & Navigation*, vol. 14, no. 5, pp. 653–661, 2020.

[21] O. Perlman, H. Ito, K. Herz et al., "Quantitative imaging of apoptosis following oncolytic virotherapy by magnetic resonance fingerprinting aided by deep learning," *Nature Biomedical Engineering*, vol. 6, no. 5, pp. 648–657, 2022.

[22] Z. Gao and J. Xiang, "Real-time 3D object detection using improved convolutional neural network based on image-driven point cloud," *Recent Advances in Electrical & Electronic Engineering (Formerly Recent Patents on Electrical & Electronic Engineering)*, vol. 14, no. 8, pp. 826–836, 2021.

[23] F. P. Wang, Z. C. Xu, and Q. S. Shi, "Integrated method for road extraction: deep convolutional neural network based on shape features and images," *Journal of Nanoelectronics and Optoelectronics*, vol. 16, no. 6, pp. 1011–1019, 2021.

[24] W. X. Shen, X. Zeng, F. Zhu et al., "Out-of-the-box deep learning prediction of pharmaceutical properties by broadly learned knowledge-based molecular representations," *Nature Machine Intelligence*, vol. 3, no. 4, pp. 334–343, 2021.

[25] X. Lin, X. Wang, and L. Li, "Intelligent detection of edge inconsistency for mechanical workpiece by machine vision with deep learning and variable geometry model," *Applied Intelligence*, vol. 50, no. 7, pp. 2105–2119, 2020.

[26] A. Di Ieva, C. Russo, S. Liu et al., "Application of deep learning for automatic segmentation of brain tumors on magnetic resonance imaging: a heuristic approach in the clinical scenario," *Neuroradiology*, vol. 63, no. 8, pp. 1253–1262, 2021.

[27] H. Lilienkamp, S. von Specht, G. Weatherill, G. Caire, and F. Cotton, "Ground-motion modeling as an image processing task: introducing a neural network based, fully data-driven, and nonergodic approach," *Bulletin of the Seismological Society of America*, vol. 112, no. 3, pp. 1565–1582, 2022.

Hindawi
Computational Intelligence and Neuroscience
Volume 2022, Article ID 4713311, 9 pages
https://doi.org/10.1155/2022/4713311

Research Article

Visual Information Computing and Processing Model Based on Artificial Neural Network

Junling Wang[1] **and Shuhan Liu**[2]

[1]*School of Journalism & Communication, Lanzhou University, Lanzhou, Gansu 730000, China*
[2]*School of Information Science and Engineering, Lanzhou University, Lanzhou, Gansu 730000, China*

Correspondence should be addressed to Junling Wang; wjl@lzu.edu.cn

Received 9 August 2022; Revised 4 September 2022; Accepted 15 September 2022; Published 30 September 2022

Academic Editor: Ning Cao

This paper analyzes the parallel and serial information processing structure of visual system and proposes a visual information processing model with three layers: visual receptor layer, visual information conduction and relay layer, and information processing layer of visual information computing and processing area. Based on the analysis, abstraction, and simplification of the biological prototype of each layer in the visual system, a framework model of an artificial neural system corresponding to the visual system is proposed. An artificial neural network model is proposed to simulate the mechanism of visual attention. A network model is formed by introducing the saliency mask map as additional information on the benchmark network, and the selective enhancement operation is performed on the extracted features in different regions according to the mask map. The experimental results show that the visual computing processing network model can effectively improve the classification performance of the network when the appropriate saliency mask is used. The visual information computing and processing model network can work effectively for different data sets and different structures of the benchmark network, which is a universal network model. The complexity of visual information computing and processing model network is very small, and the improvement of network performance is not at the cost of increasing model complexity, but in the way of improving network efficiency. The performance of artificial neural network visual information computation and processing model is directly related to the performance of saliency map used as mask map.

1. Introduction

At present, the basic framework of computer vision research is still computational vision theory. The hierarchical idea of this theory holds that a high-level computational problem can be independent of the understanding of the algorithm that performs the computation. Algorithmic problems can be solved independently of the understanding of their physical implementation. This view has already caused controversy. There is a strong objection that the assertion of the independence of the levels confuses two completely different types of problems. The first question is, is it possible to understand algorithms and analytical problems regardless of experimental facts? The answer is clearly no. The second question is, can a given algorithm that performs a task on a given machine be implemented on other machines with different structures? The answer is yes. More than 80% of the external world information acquired by human beings is visual information. Therefore, visual information processing is one of the most deeply studied fields in the whole neuroscience [1]. For how the colorful external world becomes the process of human vision, optic neuroscientists have done a lot of outstanding research work from the retina to the cerebral cortex and made a series of important research results, and these results have laid a solid foundation for the study of visual system models and algorithms.

Under such a development trend, the improvement of the current artificial neural network model has met a bottleneck, and simply increasing the depth of the network could not continue to effectively increase the performance of

the network, but put forward high requirements for the computing performance of the computer. Model complexity of ascension will bring training and ability of the model to express the increase of the difficulty, and now has more training speed is close to and exceed the speed limit of model expression ability, lead to the improvement of traditional network structure, improve the network complexity, and use various means to decrease the difficulty of the training is in the bottleneck. In fact, the current research focus has gradually shifted from improving network performance to reducing model complexity and computational load without reducing network performance [2]. Now, it is urgent to develop a new idea to make further improvement of the artificial neural network algorithm. And the human visual attention mechanism can make the human visual system achieve high performance. If the visual attention mechanism is introduced into the artificial neural network model, the performance of the neural network can be improved by a method that does not increase the complexity of the model.

Based on the information transmission and processing mechanism provided by the visual system prototype, this paper proposes a comprehensive artificial neural network model of the visual nervous system by applying the research results of the artificial neural network, computer vision, and artificial intelligence. In order to further study the visual system, we should make full use of the research results of cells and molecules, and combine the advantages and research strategies of different disciplines. We can refer to the visualization system implementation technology of neural mechanism. Therefore, an artificial neural system model based on the neural mechanism of the visual system is proposed in this paper. The implementation of this model can be integrated by various artificial neural networks according to different basic visual information processing needs.

2. Related Work

With the enhancement of computing performance, artificial neural networks with more parameters and more complex structure can be put into practical application because the implementation of the artificial neural network on the CPU can accelerate the operation speed to ten times or even tens of times [3], which greatly improves the training speed of the artificial neural network and thus accelerates the development speed of artificial neural network technology. The proposed AlexNet used two graphics cards for training and won the championship of several international competitions with an overwhelming advantage [4], which ushered in the second climax of the development of artificial neural network technology. Another famous cat experiment was also carried out: convolutional neural network was used to automatically identify the concept of the cat from video through unsupervised learning [5]. Subsequently, a large number of researchers have invested in the research of artificial neural network technology and constantly refresh records on public data sets. The proposed convolutional artificial neural network achieves 99.79% correct rate on the handwritten database MNIST (Mixed National Institute of Standards and Technology database), and the proposed multicolumn deep neural network achieves 99.77% correct rate on the MINST. Meanwhile, the best result obtained by using the SVM method on MNIST is only 99.4% [6], and the best result obtained by using the K-nearest neighbor method is only 99.3%. For the first time, the classification error rate on ImageNet is lower than the human classification error rate (5.1%) [7]. In some of the larger scale, the number of images, and more variety of database test results, artificial neural network technology is overwhelming. Since then, after several years of research by a large number of researchers, the structure of artificial neural networks has been developed from several layers, dozens of layers in the early years to hundreds of layers. Models such as A1exNet [8], VGG [9], ResNet [10], DenseNet [11], and DPN (digital packet network) [12] are all classical network models appearing in this development process. In addition to these classical structural improvements, there are also the proposed Dropout and Dropconnect [11] regularization methods and the proposed trainable activation function structure maxout [12]. Due to the deepening of network layers in this stage, the focus of research also began to focus on how to solve the problems of gradient explosion, gradient disappearance, overfitting, and so on caused by the increase of network layers, and the artificial neural network technology in this stage gained the name of deep learning. And now, the research of artificial neural network has fallen into a bottleneck.

A classical problem in computer vision is the recognition of objects in images. There are many different forms of object recognition, such as image classification, detection, and image segmentation [13]. Among them, the main goal of the image classification problem is to judge the category of objects contained in a given image, that is, what the target is. The problem of object detection is further based on the problem of image classification, which not only determines what the object is but also locates the spatial position of each object, that is, where the object is. The goal of the image segmentation problem is to find all the pixels corresponding to each object in the image, so it can be regarded as a classification task for each pixel [14, 15]. These three problems are the classical problems in computer vision object recognition task, and also the basic problems to be solved before the construction of complex vision system, which are of great significance in practical applications. In the face of various challenges in the field of computer vision, researchers have proposed a series of algorithms based on statistical features, which are often referred to as traditional computer vision algorithms. Traditional visual algorithms usually rely on statistical methods to extract features in images, such as HOG [16]. After extracting features, these algorithms usually use traditional machine learning algorithms such as logistic regression to get the final results. For different application scenarios, the statistical laws of features are usually different, so the methods of feature extraction in different visual tasks are often different [17]. Traditional computer vision algorithms have achieved good results in some simple visual tasks, but the results are still not ideal in most visual tasks, and it is difficult to meet the complex

requirements of various application scenarios. In the object detection task, CNN-based object detection model-1 can usually be divided into a two-stage detection model [17] and a one-stage side detection model [18].

Generally speaking, the former has a higher detection accuracy, while the latter can achieve a faster detection speed. For example, the proposed MegDet model adopted a two-stage detection framework and achieved 52.5% mmAP in the challenge [19], winning the champion of that year's competition. The proposed Pelee uses a single-stage detection framework and achieves a detection speed of 17fps on the device [19]. The method of designing a special feature extraction network for object detection task is explored, and DetNet is proposed, which can improve the precision of side detection while reducing the computation. In the task of image segmentation, the model based on the full convolutional network [20] has become the mainstream method to solve the problem of semantic segmentation. Among them, a series of models [21] constantly refresh the optimal results of semantic segmentation problems. The proposed GCN model points out the necessity of large convolution kernels and large receptive fields in semantic segmentation tasks [22]. For instance segmentation, the proposed model combines the segmentation model and object detection model in parallel [23] and uses the alignment operation to extract feature I for each candidate target region, which simultaneously improves the accuracy of instance segmentation and object detection. In recent years, the increase in network depth makes the training and optimization of the model increasingly difficult [24–26]. It is difficult to continue to improve the network model by increasing the depth of the network. Therefore, in recent years, the research hotspots in related fields have also begun to change. On the one hand, researchers have begun to study how to simplify the network structure and obtain the maximum performance with the minimum amount of computation while slightly improving or maintaining the network performance. On the other hand, researchers began to study more specific algorithms for data sets combined with more specific application scenarios.

3. Prototype and Model of Visual Information Computation and Processing Based on Artificial Neural Network

3.1. Artificial Neural Network Model for Biological Prototype of Visual Information Computation and Processing. The visual signals are processed by the neural network of the retina and then transmitted from the optic nerve to the visual center. The visual information is computed and processed after the geniculate body of the lateral thalamus. Complex information processing in the computation and processing of visual information leads to visual perception. The neural cells of visual information computation and processing have a columnar fabric pattern. In addition, as in the retina and the external geniculate body, the computation and processing of visual information in visual information processing show both obvious serial characteristics and obvious parallel characteristics. Based on the regulating effect of the external genu body on the visual information flow, the relay visual information processing link controls the visual information flow from retina to visual information computation and processing in a linear or nonlinear adjustable "valve" manner, which is realized by the modeling scheme shown in Figure 1.

(1) Studies on functional architecture of visual information computation and processing show that cells with similar receptive field response characteristics tend to cluster into columnar structures in visual information computation and processing. Therefore, visual information computation and processing are organized as vertical cell populations with similar functional properties.

(2) Serial processing of visual information computation and processing: The proposed hierarchical hypothesis holds that, in visual information computation and processing, more complex receptive fields are formed by the orderly synthesis of simpler receptive fields at lower levels of visual information processing. In fact, the receptive field processes motion and color information by adopting serial hierarchical processing in their respective parallel pathways to gradually extract meaningful information.

(3) Parallel processing of visual information computation and processing: Numerous studies have shown that the visual system is organized into different pathways to transmit and process different aspects of visual information. From the perspective of parallel information processing, a single cell or cell group does not represent a certain characteristic state at the level of perception but only represents some special aspects of the perceived object. In other words, the separate parts do not represent the whole, but the relationship between them constitutes the perception of the whole.

(4) Attention mechanism to realize the interrelation of different parts, the brain must integrate the information processing independently completed in different cortical regions through some mechanism. Research shows that attention must be involved in the process of such synthesis. "Attention" will emphasize the specific properties of an object, highlighting significant visual targets, while ignoring other properties of the object and other objects.

According to the analysis of biological prototype of visual information computation and processing, on the one hand, different visual information attributes are processed separately in different visual areas of visual information computation and processing. On the other hand, there are projections between different visual areas. Feed-forward,

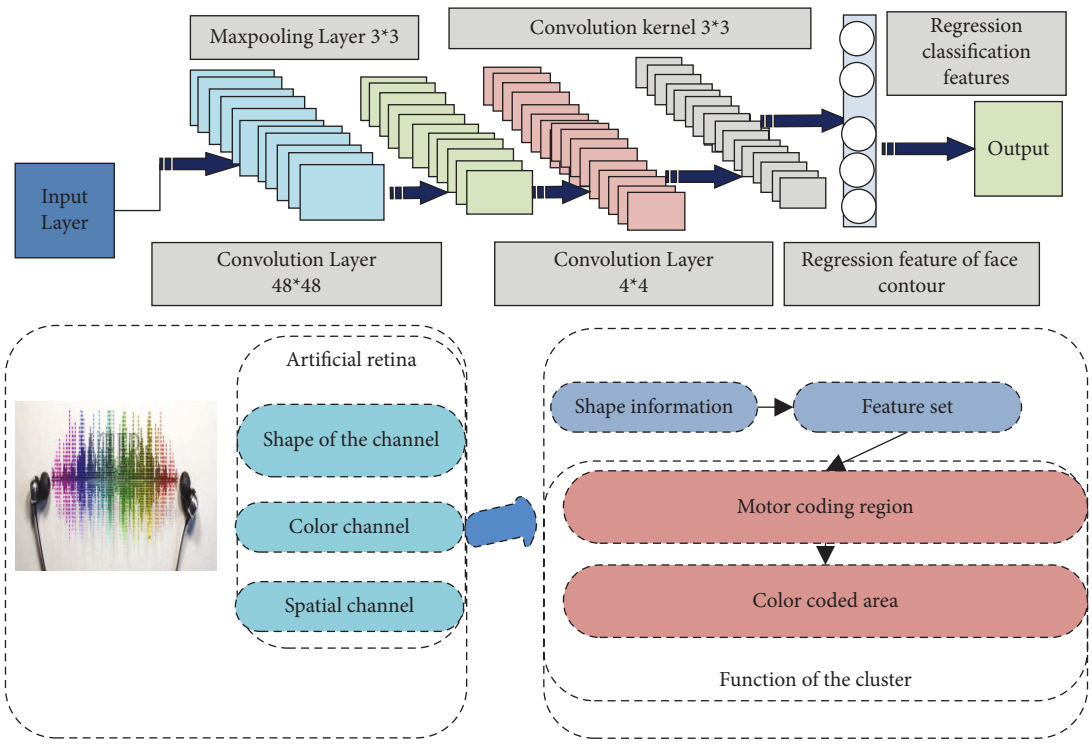


Figure 1: Relay information processing model.

feedback, and local connections in visual interval are the basis for visual information computation and processing to integrate various visual information and produce vision.

Based on assumption that neurons function hierarchically, there should be complex cells that can bring together huge visual information. Therefore, several functional clusters realized by neural networks are designed in the artificial visual information computing and processing model, and the integration of various kinds of visual information is completed by various functional clusters. For example, feature clusters that recognize characters, feature clusters that recognize fingerprints, feature clusters that recognize faces, and so on. Neurobiological studies have shown that the cortex is plastic. Therefore, visual information computing and processing must gradually form specific brain regions by adapting their functions to important stimuli. This kind of self-learning, self-organization, and self-adaptability can be realized by neural network training algorithm.

3.2. Visual Attention Computational Model Based on Target Characteristics. This paper presents a visual attention computation model based on the characteristics of the target itself, which only uses the target information but does not consider the background information of the target. The model includes training stage and attention stage. In the training stage, the color feature is decomposed into three parts: red, green, and blue. The brightness feature includes darker and brighter, the orientation feature includes four directions, and the texture feature also includes two types. Therefore, there are eleven features in total. Background information extraction depends on the target itself rather than on these features. At the same time, the bottom-up saliency map is obtained by the contrast of the graph to be noticed, and the product of these two saliency maps gives the global saliency map. Finally, the size of each saliency region is obtained by maximizing the entropy method. This model has two contributions. One is that the extracted features only depend on the characteristics of the target itself, but not on the background information. The other one is the global saliency plot which is the point-to-point product of the top-down saliency plot and the bottom-up saliency plot. The experimental results show that the model is superior to the pure bottom-up model. When the target to be noticed does not always appear in the background of the training target, or the combination of the target to be noticed and its background is very different from the combination of the training target and its background, the model in this chapter is superior to the VOCUS top-down model and Navalpakkam model.

In the task of finding a red cup in a given input scene, the main flow of the visual attention model based on the features of the target itself is shown in Figure 2. First, the features of the target are extracted from the training image, and each feature is expressed as the mean and standard deviation. Second, the feature information is used to compare the

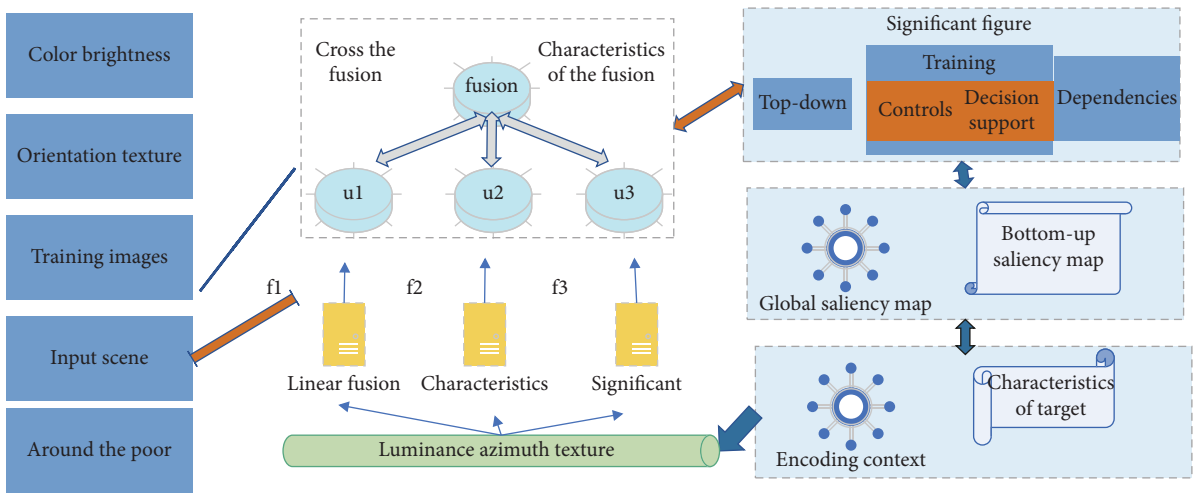


Figure 2: Visual attention model based on target characteristics.

similarity with the features in the input scene to form a top-down saliency map. Third, according to the characteristics of the input scene itself, the bottom-up saliency map is obtained. Finally, the top-down and bottom-up saliency maps are fused to generate a global saliency map to guide visual attention to the possible location of the target.

Error estimation in visual computing is developed on the basis of image processing. It studies the cognitive process of visual information from the level of information processing and studies the computational theory, expression, and computational methods of visual information processing. Error analysis has been a very difficult and rarely studied problem in all stages of visual information processing. However, there is an obvious drawback that it does not reflect the error characteristics of the structure and distribution, and meaningful subtle differences may lead to wrong decisions. The reliability and stability of the visual algorithm are poor, and the calculation results cannot be estimated sometimes and can only be verified by experiments, which fundamentally limits the generality of the visual algorithm. Therefore, how to establish the corresponding error estimation method in each stage of visual information processing is a field worthy of further study in computer vision. The correct estimation of visual computation error will in turn provide a basis for the improvement of the algorithm. The study of a more refined morphological matching error analysis method can effectively improve the reliability and stability of the 3D morphological estimation algorithm because the morphological matching error analysis is based on the structural characteristics of the error, which is more in line with human perceptual characteristics. As a practical technique, its computation has the characteristics of high efficiency and high parallelism.

The work of the visual attention mechanism can be divided into two different working modes. One is a bottom-up, feature-driven visual saliency mechanism, and the other is a top-down, task-driven selective attention mechanism. For an input visual scene, areas with more distinctive features and stronger contrast usually have higher visual saliency, and these areas with higher visual saliency will attract more attention. So when we look at this image, our attention is drawn to the little girl in the image. Therefore, inspired by the visual attention mechanism, in order to further improve the performance of the computer vision model, we propose a structure that can simulate the feature extraction function of the visual attention mechanism in the convolutional neural network. First of all, in order to simplify this problem, we introduce some prior knowledge: we directly use the existing visual saliency algorithms, calculate and process the input image data according to these existing visual saliency algorithms, and obtain the saliency mask map corresponding to each image. These saliency masks are introduced into the training of convolutional neural networks as prior knowledge. In this way, we solve the problem of how the human visual system obtains the global saliency score of the input visual scene by using the existing saliency algorithm.

The input original feature map usually comes from a convolution connection. Here, we express the formula of convolution connection in a general form:

$$\begin{aligned} t_{jk} &= \sum u_{jk} + x_{jk}, \\ x_{jk} &= f\left(z_{jk}\right). \end{aligned} \tag{1}$$

Then, according to the design idea explained above and the connection structure shown in the figure above, we can obtain the forward propagation formula of the connection as follows:

$$z_{jk} = y_{ij} * s_{jk}. \tag{2}$$

After obtaining the forward and backward transfer formulas, we can make a preliminary and simple analysis of its working principle. It can be seen from the formula that, during forward propagation, the value of each pixel on the enhanced feature map is obtained by enlarging the value of each pixel on the original feature map in equal proportion to its significant value for the location pixel. So we can determine, in the forward

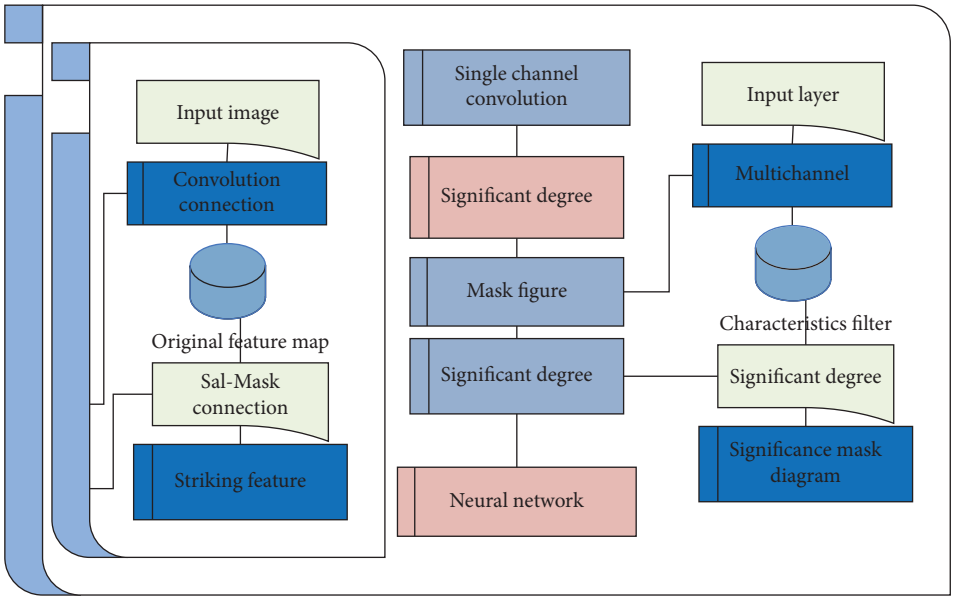


Figure 3: Training multiple input saliency maps.

propagation, after dealing with the connection, a significant degree of mask value larger area on the graph, which is considered more important areas, will be multiples of corresponding amplified to a larger area, by contrast, a significant degree of mask graph value smaller areas, which is considered less important area, only corresponding amplification area into smaller multiples. Therefore, in the process of forward propagation to obtain the final classification result, the features from more important regional parts will obviously have a greater impact on the classification result.

This structure can be regarded as a subartificial neural network consisting of only one layer of convolutional connections, through which the saliency mask map is trained. In fact, as long as the final output is a saliency mask map of the same size as the feature map, we can design a convolutional neural network with any structure, take any number of saliency mask maps as input, and use the error information feedback from the connection to train the network. Using such a scheme to obtain the enhanced saliency mask can theoretically make very fine adjustment and optimization of the input saliency mask. Even further, we can no longer need to use the trained saliency mask map as additional input prior knowledge but can design a network for obtaining the saliency mask map directly from the input image. This is shown in Figure 3.

4. Example Verification

The top-down visual attention model processes tasks that point to the target of a particular task. There are two types of experiments: one is a known task type, and the synthesized scene belongs to this type. The other is the unknown task type. In this case, the task type to be processed is obtained according to the method in the target representation, and most natural scenes belong to this type. Figure 4 shows the experimental results of the single-target composite image scene, and all the experimental results of the single-target scene are shown in Figure 4.

Figure 5 shows a more detailed comparison of MENet with these network models. These comparison results prove that MENet has stronger feature representation ability in different application scenarios.

In visual information processing, the real-time performance is critical, so the size of the gateway time delay is an important indicator of success for the gateway, and the gateway created in order to get the time delay, under laboratory conditions, was measured by the gateway in the CAN. Data as shown in Table 1.

The spatial diagram of visual information computation and processing operator of the two-dimensional artificial neural network is shown in Figure 6. With different values, $g(x, y)$ can constitute different bandpass filters. By selecting different receptive field fovea width parameter D, the purpose of extracting different spatial frequency information can be achieved. The visual information calculation and processing operator of artificial neural network has shown good performance in many aspects. The concentric circle antagonistic receptive field of biological optic nerve cells can be described by the visual information computing and processing operator of artificial neural networks.

Although the visual information computation and processing operator of artificial neural network has good performance in many aspects, it has great shortcomings in the extraction of orientation information because the visual information computation and processing operator of artificial neural network is not directional. Especially, in recent years, many biologists have found that biological optic nerve cells have obvious orientation selectivity, and morphological studies have shown that the elongated oval dendrites of optic cells may be the basis of orientation selectivity.

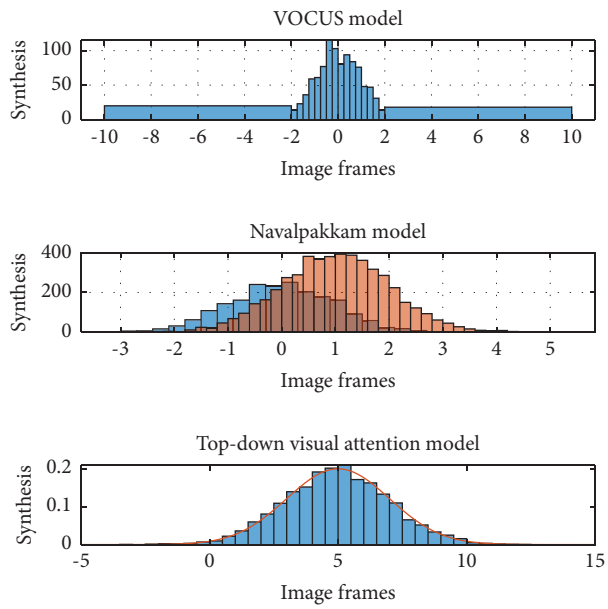


Figure 4: Experimental results of all single-object scenarios.

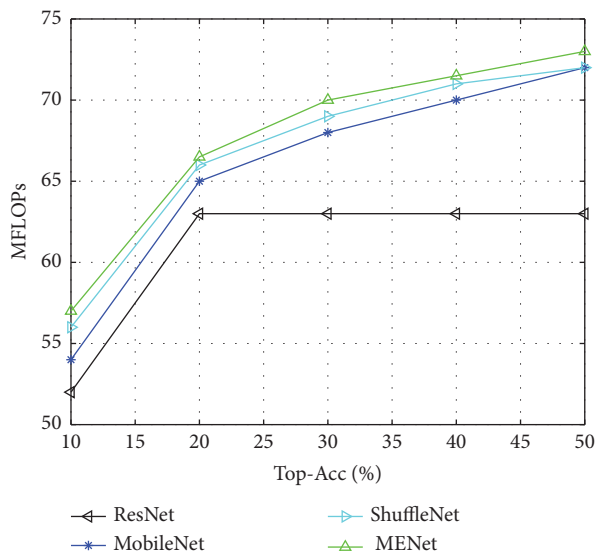


Figure 5: The MENet model outperforms all the compared network models at all four computational magnitudes.

Table 1: Time delay of CAN standard frame.

Baud rate	1.3 ms	4.7 ms	9.5 ms	19.5 ms
220	65	18	10	6
510	65	16	9	5
755	67	18	8	6

The spatial diagram of the visual information computation and processing operator of artificial neural network is shown in Figure 7. Below, the operation mechanism and characteristics of the visual information computation and processing operator of artificial neural network are analyzed.

In summary, the visual information computation and processing operator of artificial neural networks has two important characteristics: (1) we highlight the characteristics of the central region and suppress the surrounding region and (2) the characteristics of direction selectivity. It is these two characteristics of the operator that enable it to simulate the attention mechanism of the nonclassical receptive field of biological vision and to fully take into account the two important elements of gray level and orientation in the extraction of salient area information in the process of visual bottom-up.

5. Conclusion

The visual system has three layers of information processing structure: the visual receptor layer, the visual information

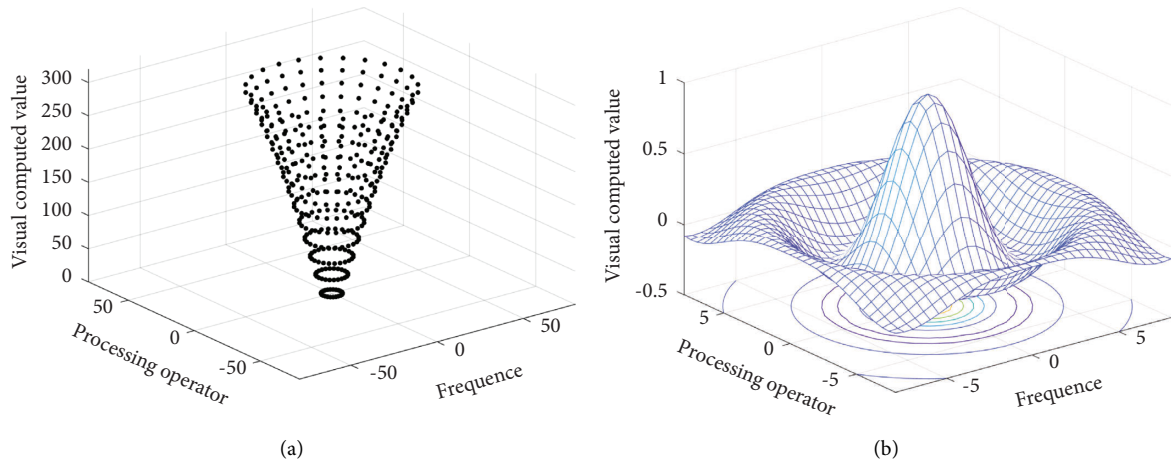


Figure 6: Visual information computation and processing operator spatial diagram of artificial neural network. (a) Visual information computation and processing operator of ON artificial neural network. (b) Visual information computation and processing operator of OFF artificial neural network.

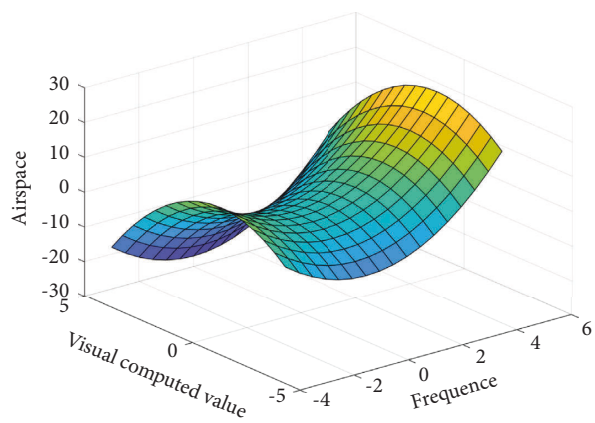


Figure 7: Spatial diagram of visual information computation and processing operator of artificial neural network.

conduction and relay layer, and the information processing layer in the visual cortex area. The layers are not independent of each other, and there are numerous couplings and interactions. Our research shows that modeling the brain's visual system in a reasonable manner, in addition to from the angle of computational neuroscience, provides theoretical support for the further study of visual electrophysiology and experiment guidance, will also be in biological vision and set up a bridge between computer vision, and is expected to be as many images as the carrier of the engineering application provides efficient front-end preprocessing solution. Most visual information processing processes are multiinput, nonlinear, and sensitive to errors caused by noise or discretization, so model-based computer vision systems cannot be applied. In the next step, artificial neural network will become an important means to solve such problems with its superior nonlinear pattern classification performance and strong self-organization and self-learning ability.

Data Availability

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

Conflicts of Interest

The authors declare that they have no conflicts of interest or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work was supported by the Humanities and Social Sciences Planning Project of Ministry of Education: "The Media Image, Discourse Evolution and Communicative Intervention of Anti-intellectualism in Contemporary China" (21XJA860003).

References

[1] H. Badrzadeh, R. Sarukkalige, and A. W. Jayawardena, "Improving ann-based short-term and long-term seasonal river flow forecasting with signal processing techniques," *River Research and Applications*, vol. 32, no. 3, pp. 245–256, 2016.

[2] W. Wu, Z. Zhou, and S. Wu, "Automatic liver segmentation on volumetric CT images using supervoxel-based graph cuts," *Computational and Mathematical Methods in Medicine*, vol. 2016, Article ID 9093721, 2016.

[3] H. Chen, "Research on innovation and entrepreneurship based on artificial intelligence system and neural network algorithm," *Journal of Intelligent and Fuzzy Systems*, vol. 40, no. 2, pp. 2517–2528, 2021.

[4] J. Sinha, "A comparison of network types in artificial neural network-based rainfall-runoff modelling," *International Journal of Applied Research on Information Technology and Computing*, vol. 8, no. 1, pp. 41–50, 2017.

[5] B. Hu, S. Yue, and Z. Zhang, "A rotational motion perception neural network based on asymmetric spatiotemporal visual

information processing," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 28, pp. 1–19, 2016.

[6] J. Liu, L. Lin, and Z. Cai, "Deep web data extraction based on visual information processing," *Journal of Ambient Intelligence and Humanized Computing*, pp. 1–11, 2017.

[7] C. Cui, Z. Jia, and K. Gu, "Research on the information entropy using processing square matrix method based on similarities evaluation model," *Journal of Physics: Conference Series*, vol. 1952, no. 4, pp. 042059–042068, 2021.

[8] S. Hu, "Research on data acquisition algorithms based on image processing and artificial intelligence," *International Journal of Pattern Recognition and Artificial Intelligence*, vol. 34, no. 6, pp. 21–33, 2020.

[9] Z. Wei and G. Lei, "The research on dynamic properties of the small world neural network based on the synaptic plasticity," *Journal of Biomedical Engineering*, vol. 35, no. 4, pp. 509–517, 2017.

[10] L. Guo, H. Lv, and F. Huang, "Research on neural information coding of spiking neural network based on synaptic plasticity under AC electric field stimulation," *International Journal of Pattern Recognition and Artificial Intelligence*, vol. 33, no. 7, Article ID 1959021, 2019.

[11] J. Tan, D. Xia, and S. Dong, "Research on pre-training method and generalization ability of big data recognition model of the internet of things," *ACM Transactions on Asian and Low-Resource Language Information Processing*, vol. 20, no. 5, pp. 21–35, 2021.

[12] B. Li, Y. F. Zeng, and B. B. Zhang, "A risk evaluation model for karst groundwater pollution based on geographic information system and artificial neural network applications," *Environmental Geology*, vol. 77, no. 9, p. 344, 2018.

[13] S. I. Dimitriadis, N. A. Laskaris, and P. G. Simos, "Greater repertoire and temporal variability of cross-frequency coupling (CFC) modes in resting-state neuromagnetic recordings among children with reading difficulties," *Frontiers in Human Neuroscience*, vol. 10, no. 20, pp. 34–46, 2016.

[14] C. Ye, F. Gu, and W. Shi, "Research on the communication signal processing and parallel decoding algorithm based on neural network," *Revista de la Facultad de Ingenieria*, vol. 32, no. 3, pp. 255–263, 2017.

[15] J. Zhang, T. Tian, S. Wang, X. Liu, X. Shu, and Y Wang, "Research on an olfactory neural system model and its applications based on deep learning," *Neural Computing & Applications*, vol. 32, no. 10, pp. 5713–5724, 2020.

[16] G. Zhang, "Research on safety simulation model and algorithm of dynamic system based on artificial neural network," *Soft Computing*, vol. 26, no. 15, pp. 7377–7386, 2022.

[17] Y. Wang, J. Wang, and P. Deng, "Research on the fusion model of sports and medicine based on artificial neural network health analysis and forecast-take atherosclerosis as an example," *Journal of Physics: Conference Series*, vol. 1848, no. 1, pp. 12100–12106, 2021.

[18] Y. Xia, "Research on statistical machine translation model based on deep neural network," *Computing*, vol. 102, no. 3, pp. 643–661, 2020.

[19] J. Zhou, M. Yang, F. Wang, and C Wang, "Research on financial aid model of colleges and universities based on artificial neural network," *Journal of Physics: Conference Series*, vol. 1575, no. 1, pp. 12100–12104, 2020.

[20] L. Liu and W. Ran, "Research on supply chain partner selection method based on BP neural network," *Neural Computing & Applications*, vol. 32, no. 6, pp. 1543–1553, 2020.

[21] R. K. Singh, D. Ray, and B. C. Sarkar, "Mineral deposit grade assessment using a hybrid model of kriging and generalized regression neural network," *Neural Computing & Applications*, vol. 34, no. 13, pp. 10611–10627, 2022.

[22] B. Hong and Y. Zhang, "Research on the influence of attention and emotion of tea drinkers based on artificial neural network," *Mathematical Biosciences and Engineering*, vol. 18, no. 4, pp. 3423–3434, 2021.

[23] Y. Li, "Research on application of BP neural network in library electronic resource quality evaluation," *Revista de la Facultad de Ingenieria*, vol. 32, no. 5, pp. 605–613, 2017.

[24] X. Shen, G. Shi, Y. Zhang, and S. Weng, "Wireless volatile organic compound detection for restricted internet of things environments based on cataluminescence sensors," *Chemosensors*, vol. 10, no. 5, p. 179, 2022.

[25] X. Shen, G. Shi, H. Ren, and W. Zhang, "Biomimetic vision for zoom object detection based on improved vertical grid number YOLO algorithm," *Frontiers in Bioengineering and Biotechnology*, vol. 10, no. 5, Article ID 905583, 2022.

[26] G. Shi, Y. He, and C. Zhang, "Feature extraction and classification of c images based on sparse coding convolutional neural networks," *IEEE Transactions on Instrumentation and Measurement*, vol. 70, Article ID 9501811, 2021.

Hindawi
Computational Intelligence and Neuroscience
Volume 2022, Article ID 1721157, 11 pages
https://doi.org/10.1155/2022/1721157

Research Article

The Analysis of Environmental Cost Control of Manufacturing Enterprises Using Deep Learning Optimization Algorithm and Internet of Things

Jin Qiu[1] and Wenzhuo Chen[2]

[1]*Guangdong University of Science and Technology, Dongguan 523000, China*
[2]*Department of Electronics and Information Engineering, North China Institute of Science and Technology, Langfang 065201, China*

Correspondence should be addressed to Wenzhuo Chen; wenzhuochen@ncist.edu.cn

Received 20 July 2022; Revised 20 August 2022; Accepted 25 August 2022; Published 30 September 2022

Academic Editor: Ning Cao

Under the background of the Internet of things (IoT), the problems between the actual production and the environment are also prominent. The environmental cost control in the production process of manufacturing enterprises are discussed to reduce the environmental cost and promote the improvement of production efficiency. First, the environmental cost under the background of IoT is analyzed. Also, the environmental cost control methods in the production process of traditional manufacturing enterprises are investigated. Second, based on the principle of traditional genetic algorithm, the fast-nondominated sorting genetic algorithm (NSGA-II) of multiobjective genetic algorithm is introduced to complete the optimization of BP neural network (BPNN) algorithm in deep learning (DL), and the multiobjective GA optimization BPNN model is established. Finally, the multiobjective GA algorithm is used to empirically analyze the environmental cost control capability of a paper-making enterprise. It is compared with enterprises with excellent and poor environmental cost control capabilities in the same industry to find out secondary indexes. The results show that environmental costs have long-term and economic characteristics. The global search ability of BPNN optimized by multiobjective GA is improved, and the local optimal dilemma is avoided. Through empirical analysis, it is found that the comprehensive capability of the environmental cost control of the enterprise is better, scored 79 or more, and the indexes of insufficient development and advantages are obtained. As IoT rapidly develops, it is necessary to further improve the ability of enterprises in environmental cost management, which is very important to promote the development of enterprises and enhance their core competitiveness. It is hoped that this investigation can provide certain reference significance for improving the environmental cost management capability of enterprises, increasing production efficiency, and reducing environmental costs.

1. Introduction

The soaring economic growth in China has seen the explosion of manufacturing enterprises, which has brought greater pressure on the ecological environment. Currently, the environmental cost is an important index for the development and production of the green economy and plays a significant role in promoting economic development in China. The environmental cost is first proposed as early as the 20th century based on the analysis of the relationship between economic development and environmental protection [1]. Scholars worldwide have explored the classification of environmental costs, which is discussed from the scope of the enterprise's production activities, the main body of responsibility, as well as the utilization and consumption of resources. Then, numerous classification methods are proposed. With social progression, the concept and classification of environmental costs have been perfected [2]. With further analysis of environmental costs, the control methods of environmental costs are constantly increasing.

Some scholars have analyzed the main drivers of environmental cost management. The main components of environmental cost drivers are the work process and the organizational structure of the enterprise. Therefore, a management and analysis system should be established for enterprise environmental cost management [3]. Some scholars believe that the management of environmental costs of the enterprise should start from the perspective of advanced control. Thus, preventive management should be carried out to reduce the environmental cost, and enterprises are managed and controlled from the source of environmental pollution [4]. Other scholars have determined the environmental cost through the cost accounting concept in the comprehensive environmental and economic accounting management system. Then, it is combined with the ecological footprint model to account for the environmental cost caused by the occupation of ecological resources [5]. At present, scholars worldwide have conducted an in-depth investigation of the management and control methods of the environmental cost. Besides, they have explored the management and control of the environmental cost through various methods.

Afterward, many scholars have evaluated the environmental cost control capability. Some scholars have analyzed the links and factors of enterprise environmental cost control using the fuzzy analytic hierarchy process so that the enterprise can observe the advantages and disadvantages of the application of environmental cost control. To achieve sound and rapid development, an enterprise should recognize its current level of development in the same industry. The same is true for improving environmental cost control capability. In recent years, with the development of DL (Deep Learning) technology, more scholars have applied DL technology to evaluate enterprise development while achieving good results [6]. Here, DL technology and the GA (Genetic Algorithm) are combined to evaluate the environmental cost control capability of the enterprise.

Based on the existing results, the method of environmental cost control of manufacturing enterprises is explored. Meanwhile, related concepts, such as environmental costs of the enterprises, are analyzed. Then, the BP (Backpropagation) NN (Neural Network) algorithm is optimized through the multiobjective GA and DL technology. At the same time, the analytic hierarchy process is combined to evaluate the environmental cost control indexes of the enterprises. The multiobjective GA optimization is used for the index evaluation of insufficient environmental cost control capabilities of the enterprises.

2. Methods

2.1. Environmental Cost Control of Enterprises. Environmental costs have long-term development, and economic environmental costs include two parts: control cost and damage cost. It is the cost incurred by the enterprise's environmental protection measures to comply with environmentally friendly principles. Also, it includes other costs incurred to protect the environment [7]. Enterprises should take the initiative to protect the environment. In the process of production and operation activities, to achieve the purpose of protecting the environment, the activities costs, such as preventive expenditure, maintenance expenditure, and governance expenditure, all belong to the category of the enterprise environment. There are various types of environmental costs of enterprises. According to the period of the production and operation activities that the enterprise is engaged in, it mainly includes three costs before, during, and after the activity [8]. As the name implies, the ex ante cost refers to the cost incurred by enterprises to protect the environment before carrying out production activities. It includes the cost of improving the production process, introducing environmental protection equipment, and training employees on environmental protection knowledge. It is beneficial to reduce the cost of enterprises in production activities. The cost in the process of activity refers to the cost incurred by the enterprises to protect the environment during the production activities. It includes maintenance costs of environmental protection equipment and costs incurred in environmental testing. The ex post cost mainly refers to the cost incurred by the enterprise in the treatment and repair of the environmental pollution generated by itself. The main purpose is to make the environment self-belief and enhance the self-recovery ability.

Environmental Cost Control. Environmental cost control refers to the process by which enterprises control and manage environmental costs through scientific means, thereby predicting, calculating, evaluating, and analyzing environmental costs in an organized manner. According to the requirements of relevant national laws and regulations, enterprises can achieve the goal of green and sustainable operation [9]. Environmental costs are different from the general costs of enterprises. The environmental costs of enterprises cannot be blindly reduced. When managing the environmental cost of an enterprise, the social benefits it brings must be considered. When reducing the environmental cost of an enterprise, it is necessary to consider how to increase the individual benefits of the enterprise as well as ensure the safety and quality of the ecological environment. When conducting environmental cost management of an enterprise, one should not only pursue its operating profit one-sidedly. It is necessary to take the concept of sustainable development as its green business philosophy, thereby effectively managing the environmental costs of enterprises in real time and reducing the pollution degree of the environment caused by enterprises when they are engaged in business activities as much as possible. Enterprises need to pay attention to environmental protection while pursuing operating profits. It makes the social and economic benefits of the enterprise reach a balanced and unified state, thereby laying a solid foundation for the long-term green and healthy development of the enterprise [10].

Characteristics of Environmental Cost Control. It refers to the need for a relatively long development stage when the enterprise controls and manages environmental costs, thereby completing the management and control of the environmental costs step by step. In this process, enterprises need to

pay attention to the requirements for environmental cost control at the current stage. Moreover, it is necessary to predict the environmental costs that the enterprises may incur in the future and take measures in advance. In the initial stage, when enterprises control environmental costs, the manpower, material resources, and economic costs required to invest are relatively high. At this stage, the cost pressure of environmental management faced by enterprises is relatively high. However, with the continuous development of the enterprises, the cost invested in the initial stage may save many economic resources for later development. It is also conducive to the growth of the enterprises' core competitiveness. Economy refers to that the control of environmental costs by enterprises is to maximize the economic efficiency, maximize the benefits of resource utilization, and reduce the pollution of the enterprises to reduce the environmental protection expenditures. The input structure of the enterprises' environmental costs is optimized to reduce the cost pressure, keep the environmental costs at a reasonable level, and maximize the economic benefits of the enterprises [11]. The long-term characteristics of environmental costs are reflected in the development of enterprises. Therefore, starting from the economy of the environmental costs of enterprises, the control of environmental costs is analyzed and explored.

2.2. Multiobjective GA Optimization. There are two main methods for solving multiobjective optimization problems: traditional optimization algorithms and intelligent optimization algorithms. Most of the traditional optimization algorithms convert multiple objectives that need to be optimized into one objective and optimize the solution through a single-objective optimization method. Commonly used methods include the weighting method and the objective planning method. Currently, commonly used intelligent optimization algorithms include GA and PSO (Particle Swarm Optimization) algorithms [12]. Among them, the GA is a global optimization algorithm that searches randomly. By simulating the evolutionary laws of genetic selection and survival of the fittest, the solution of the problem is regarded as a population. The fitness of the solution is used as the basis for judging whether the solution is excellent. The excellent solution is inherited to the next generation, approaching the optimal solution through continuous evolution from generation to generation [13].

When calculating with the GA, it is necessary to pay attention to the calculation of fitness and the determination of the coding method. Among them, fitness is the basis for judging the pros and cons of individuals in a population. Excellent individuals are selected for inheritance through fitness. Then, in the calculation process, the solution of the real problem cannot be directly calculated by the GA. It needs to be encoded and transformed into chromosomes into a form that can be calculated by the GA. This algorithm is simple to operate, with strong versatility and global search capabilities [14].

The process of the GA is shown in Figure 1.

When calculating with the GA, it must first be coded. The process of composing the solution of the problem into a chromosome in a certain order is coding. Conversely, the process of translating the best individual chromosomes obtained by the GA into the solution of the problem is decoding. Currently, the commonly used chromosome coding methods include binary coding, parameterized coding, and direct coding [15].

Initializing the Population. The GA is an algorithm that evolves the population operation. In the initial stage of calculation using the algorithm, an initial population needs to be generated, which may affect the final result.

Fitness Calculation. When the GA is performing genetic operations, the selection of the genetic population is judged based on the fitness value of individuals in the population. The higher the fitness value of an individual, the better the individual. The fitness value of an individual is generally calculated by the objective function of the optimization problem. In the process of this investigation, the objective function to be calculated belongs to the problem of minimization. When transforming it into a fitness function, there are generally two methods of transforming it into an opposite number and a derivative. The transformation method is as follows [16].

$$fitnessfunc(x) = -f(x), \tag{1}$$

$$fitnessfunc(x) = \frac{1}{f(x) + e}. \tag{2}$$

Here, $f(x)$ represents the minimum objective function that is always nonnegative. e represents a greatly small number to prevent the denominator from being zero.

Genetic Operation. The selection operation is used to determine which good individuals can be used as parents to reproduce the next generation. Generally, individuals with high individual fitness values in the population will be retained. Currently, commonly used selection methods include the roulette selection method, random traversal sampling method, tournament selection method, and index sort selection method [17]. The crossover operation will recombine part of the chromosomes of each individual in the parent to produce offspring that combines the information of the parent. The crossover operation may make the generated offspring have higher fitness. Currently, the commonly used crossover operation methods include single-point crossover, two-point crossover, multipoint crossover, and mixed uniform crossover [18]. The mutation operation is to change the genes on the chromosomes of the offspring with a small probability, so that the individuals of the population have diversity, to prevent the population from premature convergence. At the same time, it allows the GA to have the ability of local random search to avoid the occurrence of local optimal phenomena. The method of mutation operation is related to the method of coding. Currently, the commonly used methods include binary

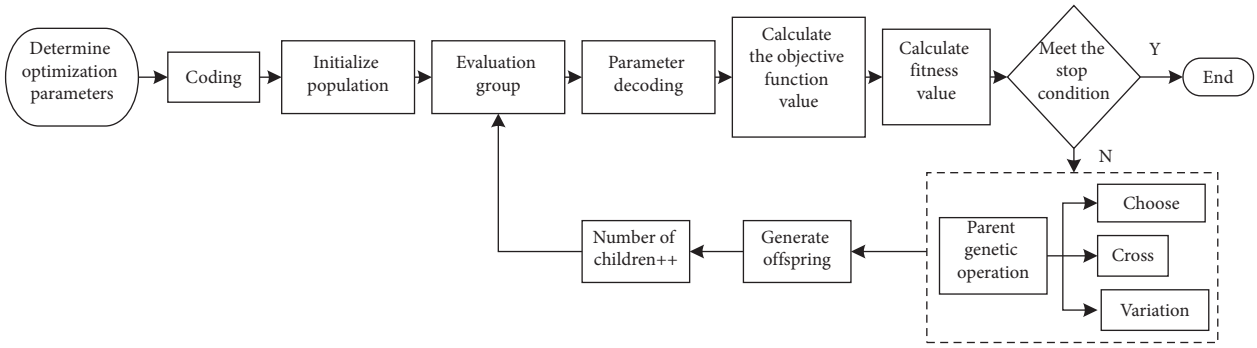


Figure 1: The calculation process of the GA.

mutation, real number mutation, and Gaussian mutation [19].

Parameter Adjustment. The GA mainly involves parameters such as population size, crossover probability, mutation probability, evolutionary algebra, and population initialization method. These parameters are used to determine whether the population will be precocious [20]. The population size will affect the calculation efficiency of the algorithm. The population that is too small will cause diseased genes in the offspring, and too large one will cause the calculation efficiency of the algorithm to decrease. Thus, the population size is usually selected in the range of 100 to 200. The value range of the crossover probability is generally between 0.4 and 0.99. If the mutation probability is too large, the good genes in the population will be destroyed. If it is too small, the diversity of the population will decrease, and the algorithm will easily fall into the trouble of local optimization. Generally, the range of mutation probability is set at 0.00001 to 0.2. Evolutionary algebra will affect the convergence times of the algorithm. Too large will cause the subsequent calculations of the algorithm to be meaningless. Too small will cause the algorithm to converge prematurely, which may cause the algorithm to fail to obtain the true optimal solution. The initialization of the population will determine the number of iterations of the algorithm to a large extent. Therefore, it is necessary to estimate the calculation result during initialization and initialize the algorithm near the global optimal solution.

GA has a strong global search: GA can improve its solving efficiency through parallel computing when facing large-scale or multiobjective problems. The mutation and crossover operations are used to prevent the algorithm from falling into a local optimal state. *Strong Versatility and Controllability*. GA can adjust its coding and genetic operations according to the actual situation and change its mutation and crossover operations so that it has strong controllability and versatility. *Good Scalability and Upgradeability*. The operating steps of the GA are simple, and it can be combined with other algorithms to make up for the deficiencies of the algorithm itself, which can be better applied to practical problems. The GA is self-organizing, adaptive, and self-learning [21].

2.3. Multiobjective GA Optimization. In environmental cost control of enterprises, due to the need to consider multiple optimization objectives, traditional GA cannot solve multiobjective problems well. The multiobjective optimization problem is mainly to ensure that satisfactory solutions can be obtained for multiple optimization objectives [22]. Therefore, multiobjective GA optimization is selected to control the environmental cost of enterprises. In the process of selection, the population is stratified through the dominance relationship among the population. The set of individuals that are not dominated is called the Pareto solution set. The probability of this solution set is passed to the next generation is greater. In the multiobjective GA, it is necessary to find the solution of Pareto and find the optimal solution in the solution set [23]. The fast NSGA-II (Nondominated Sorting GA) with elite strategy in the multiobjective GA is selected to explore the environmental cost control of enterprises. The multiobjective GA runs fast, and the solution set has good convergence. At the same time, the complexity of the NSGA-II is lower [24]. The calculation process of this algorithm is shown in Figure 2.

DL technology has been widely used in the field of computer vision research. Meanwhile, it has an excellent performance in image and time-series data processing. Currently, there are feedforward NN, RNN (Recurrent Neural Network), and self-organizing NN under DL technology [25].

BPNN is a typical multilayer feedforward NN, which mainly relies on the BP of errors during training. The weights and thresholds of the network are adjusted through BP, thus solving the nonlinear problem. The learning process of the BPNN can be divided into two stages of forwarding propagation and backward propagation [26]. The typical BPNN structure includes an input layer, hidden layer, and output layer. The signal is forwarded layer by layer in the BPNN until the output layer. Therefore, the NN is a parallel multilayer feedforward network [27].

Construction of Multiobjective GA Optimization. BPNN is optimized through the multiobjective GA, and the GA-BP (multiobjective GA optimization BPNN) model is constructed. The constructed GA-BP model includes three parts: the establishment of the BPNN model, the optimization of

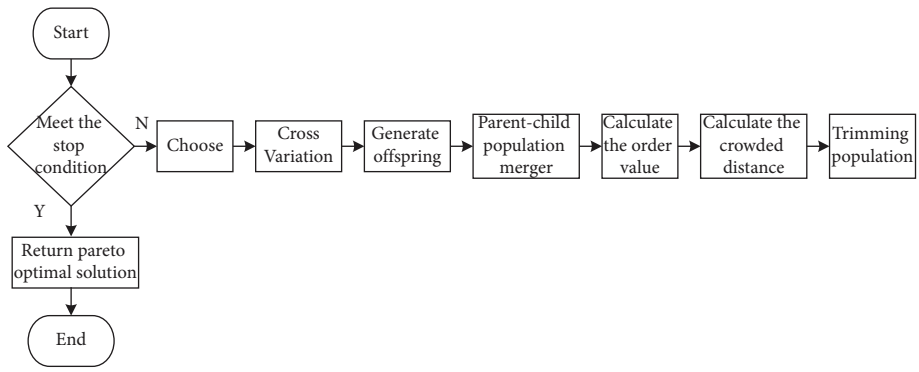


Figure 2: The calculation process of the NSGA-II algorithm.

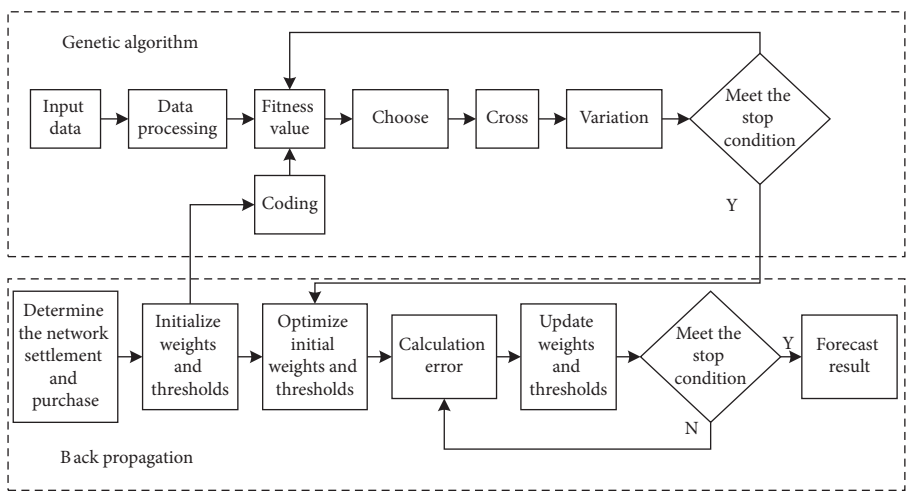


Figure 3: The process of the GA-BP model.

BPNN with GA, and the prediction of optimized BPNN [28]. The process of the GA-BP model is shown in Figure 3.

2.4. Evaluation of Environmental Cost Control of Enterprises Based on the GA-BP. The NN relies on plenty of sample data for training and performs pattern recognition based on the features of the sample. The input feature vector, network weight, and threshold value in the NN jointly determine the output. Thus, the choice of the feature vector is greatly important [29]. The input data for NN training should be distributed with spatial continuity. Sample data with correlated features should be selected for input to reduce the similarity of the samples and keep the data dimension consistent.

The GA-BP model is established on the MATLAB platform, and coding operations are performed. The input dimension of the network is determined by the number of feature vectors. The input network has the highest probability in each mode, which is the recognized mode [30]. The feature vector is compared with the feature vector of poor or excellent samples in the same industry. The features with larger gaps are taken as guidance data. The NN selects data through random sampling and optimizes it by orthogonal experiment [31]. The networks under different parameters are screened. The GA is mainly to initialize the weights and thresholds of the NN after screening.

In manufacturing enterprises, different types of manufacturing systems have great differences in basic information. The GA-BP model can evaluate and optimize the environmental cost control capability of manufacturing enterprises. A paper-making enterprise is taken as the research object to analyze its environmental cost control capability. First, the indexes and weights of environmental cost control are determined using the analytic hierarchy process [32]. Through the analysis of the actual operation of the enterprise, based on the principle of relevance and comprehensiveness, the environmental control indexes of the enterprise are divided into 4 primary indexes and 16 secondary indexes. There is a corresponding relationship between the primary indexes and secondary indexes. The determined evaluation indexes and corresponding numbers are shown in Table 1.

The analytic hierarchy process combined with the fuzzy comprehensive evaluation method is used for index evaluation, thereby getting the final weight of the index. The comprehensive membership degree, comprehensive score, and index score rate are calculated.

After each piece of data is calculated, the data are transformed. Through data discretization and normalization

Table 1: Evaluation indexes of environmental cost control of enterprises based on the GA-BP model.

	Primary index (A)	Secondary index (B)
Evaluation indexes of environmental cost control capability of enterprises	Internal input index of environmental protection $A1$	Environmental training education fee of employees $B1$
		The research and development fee of green products $B2$
		The update of environmental protection technology and equipment $B3$
		Environmental protection system construction, department establishment $B4$
	Cleaner production index $A2$	Expenditure on purchasing environmentally friendly raw materials $B5$
		Energy-saving $B6$
		Pollutant emission reduction $B7$
	Pollution control effect index $A3$	Hazardous waste storage expenditure $B8$
		The comprehensive utilization rate of waste $B9$
		The compliance rate of pollutant discharge $B10$
		Staff health status $B11$
	External influence index $A4$	Number of accidental pollution incidents $B12$
		Environmental protection tax and pollution fine $B13$
		Environmental information disclosure of enterprises $B14$
		Environmental complaints $B15$
		The review situation of environmental protection department $B16$

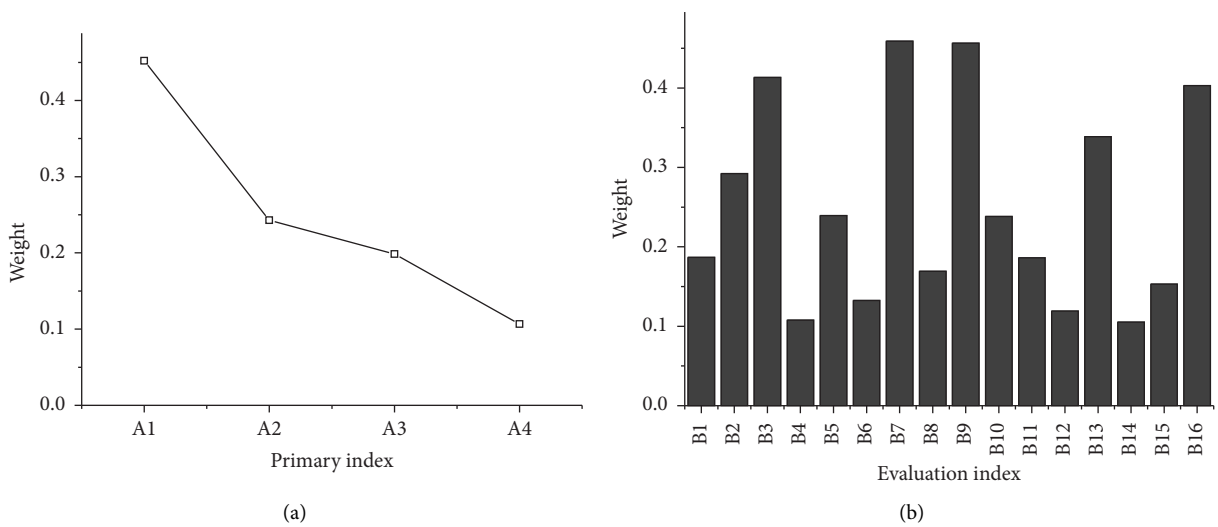


Figure 4: The calculation results of the weights of each index. (a) The calculation results of the weights of the primary indexes. (b) The calculation results of the weights of the secondary indexes

processing, the feature vector of each data index is obtained [33]. The feature vector data are input into the trained GA-BP model [34–37]. The gap of the environmental cost control capability between the paper-making enterprise and the same type of enterprise is obtained. The enterprise's scoring results in the entire industry are obtained. Then, the corresponding poor samples under its primary indexes are selected and analyzed [38–40]. Meanwhile, the control scheme of the IoT platform is adopted to ensure the ease of operation and multiterminal compatibility of enterprises in environmental control [41–43]. Specifically, the Socket-Websocket bridging module is employed to ensure the real-time of IoT so that the Web application in the ordinary browser can control the environmental monitoring equipment in real time, like industrial control software [44–48].

3. Results and Analysis

3.1. Index Weight Analysis. Through the analytic hierarchy process combined with the fuzzy comprehensive evaluation method, the weight of each index is obtained. The results are shown in Figure 4.

Figure 4 illustrates that among the primary indexes, the internal input index of environmental protection $A1$ has the highest weight, indicating that internal input has a relatively large impact on the environmental cost control capability of the enterprise. Among the secondary indexes, the indexes, such as the update of environmental protection technology and equipment $B3$, pollutant emission reduction $B7$, comprehensive utilization rate of waste $B9$, and the review situation of environmental protection department $B16$, have relatively high weight values. The results indicate that these indexes have a relatively large impact on the primary indexes.

3.2. Analysis and Calculation Results of Comprehensive Membership Degree. The calculation of the comprehensive membership degree is obtained through a QS (Questionnaire Survey). A total of 150 environmental QSs are issued. Then, 120 valid QSs are recovered, with a valid rate of 80%. The subjects of the QS include, but are not limited to, internal personnel of the enterprise related to environmental protection. According to the scores, the survey results are divided into five criteria: excellent, good, medium, poor, and very poor. The sum of the survey results of excellent, good, and medium is used as the evaluation criterion for the degree of approval of the index table. The sum of the evaluation results of poor and very poor is used as the basis for calculating the degree of disapproval. The final results of calculating the membership degree of each index are shown in Figure 5.

Figure 5 shows that the paper-making enterprises do a better job in controlling the environmental cost in production. Moreover, the evaluation recognition of each index is high. Among them, the membership degree of expenditure on purchasing environmental protection raw materials and updating environmental protection technology and equipment is relatively poor. About 20% of the employees are not in favor of purchasing environmental protection raw materials and updating environmental protection technology and equipment, and the degree of nonrecognition of green product R&D (Research and Development) expenses is more than 15%. The results show that enterprises should increase investment in R&D of green products to improve the level of R&D.

3.3. An Empirical Test on the Cost of Enterprise Environmental Loss. Figure 6 below shows the environmental cost details of X coal enterprise from 2015 to 2020.

Figure 7 shows the loss cost priority analysis based on Figure 6.

Correlation coefficient $R = 0.983$, judgment coefficient $R^2 = 0.966$, adjusted $R^2 = 0.951$, and standard error of regression estimation $s = 4.046$, indicating that the regression effect of this group of data is general. The test results of constant term and coefficient are shown in Figure 8.

Figure 8 displays that when the environmental control cost and the environmental loss cost function are equal, the specific value of the target rate can be obtained, but this value is somewhat idealized. This value is the minimum value of

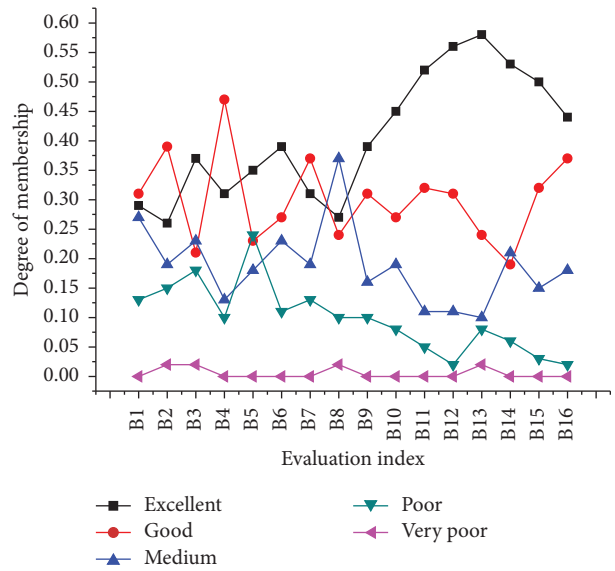


Figure 5: Analysis results of the membership degree of each index.

the total environmental cost function, but it is not necessarily the lowest point of the enterprise's environmental cost, which must be determined according to the actual situation of X coal enterprise.

3.4. Analysis of Environmental Cost Control Capability of Enterprises Based on the GA Optimization. A general evaluation is made on the environmental cost control capability of the enterprise. The primary index of the environmental control capability of the enterprise is evaluated. Figure 9 shows the statistical results of the evaluation.

As can be seen from the above figure, the total score of the environmental cost control capability of the enterprise is above 79. In general, the environmental cost control capability of the enterprise is relatively good. From the primary evaluation index, the score of the cleaner production index is the lowest, only reaching 75%. But relatively speaking, the score of this index is relatively high. In the index scores of this level, the scores of the pollution control effect index and the external control effect index are both above 80%. It shows that these two measures are implemented relatively well in the environmental cost control of the enterprise. However, in the actual environmental protection process, further maintenance and improvement are needed to strive to improve the enterprise's ability to control environmental costs. After calculation, the p-value of these two variables is 0.01, which is less than 0.05, so the research on pollution control effect index and external control effect index has statistical significance. At the same time, following the analysis of these two variables, the control effect of enterprises on environmental cost is further understood to formulate relevant strategies to improve the enterprise's environmental cost control capability.

3.5. Comparison with Other Samples in the Same Industry. The multiobjective GA optimization is used to compare the environmental cost control index of this paper-making

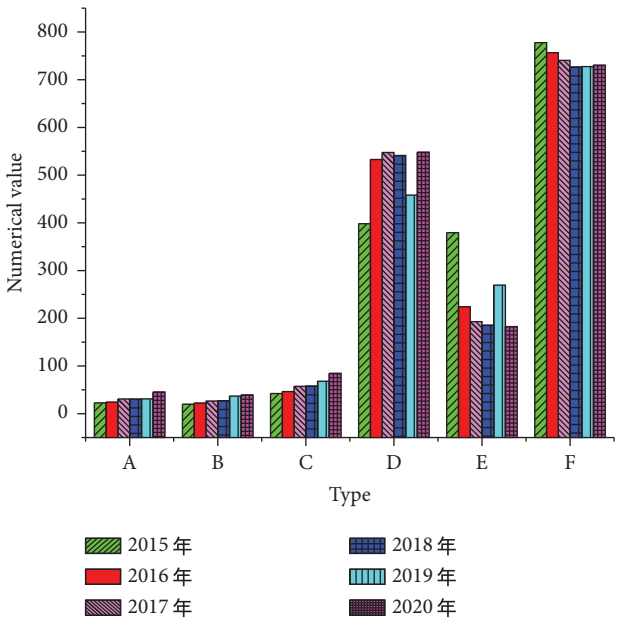


Figure 6: Environmental cost of X coal enterprise in 2015–2020 (A: environmental prevention cost, B: environmental detection cost, C: environmental control cost, D: environmental internal loss cost, E: environmental external loss cost, and F: environmental loss cost).

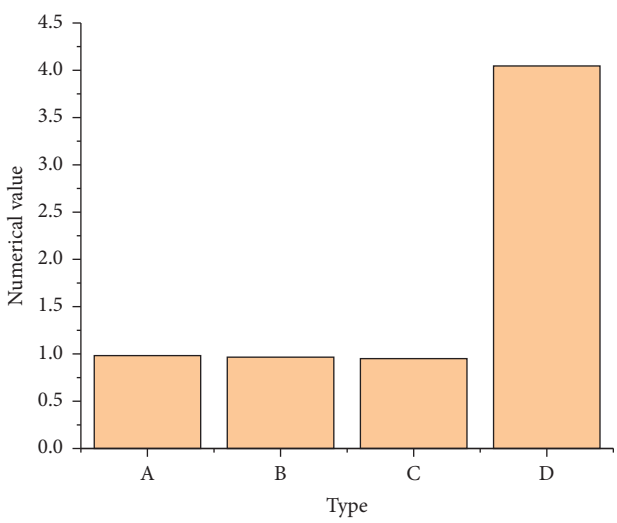


Figure 7: Loss cost priority analysis (A: R; B: R^2; C: adjusted R^2; and D: standard skewness error).

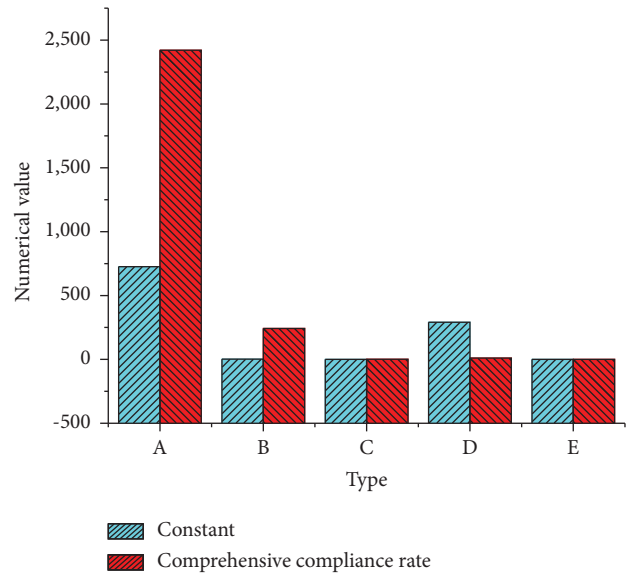


Figure 8: Test results of constant term and coefficient (A: B-value; B: standard error; C: beta value; D: T-value; and E: significance).

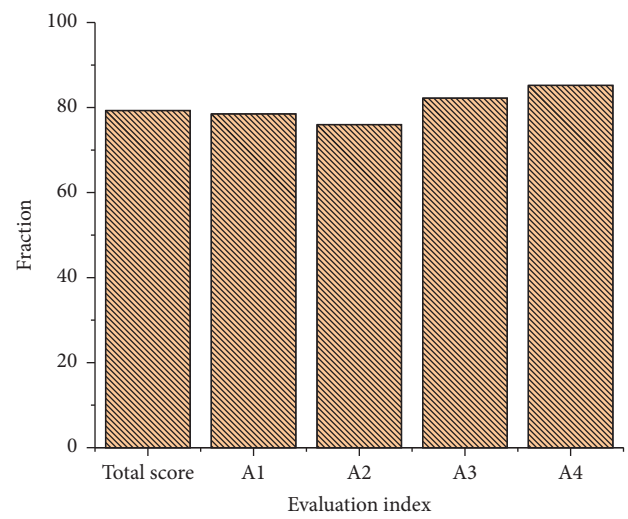


Figure 9: Analysis results of environmental cost control capability of the enterprise based on the genetic optimization algorithm.

enterprise with the feature vector of the poor and excellent enterprises in the same industry. Through the search of the algorithm, Table 2 shows the statistics of the analysis results of the secondary indexes with the largest sample gap characteristics under each primary index.

As can be seen from the above table, when compared with excellent enterprises, the secondary indexes with the largest gap between this enterprise and excellent enterprises are the update of environmental protection technology and equipment, pollutant emission reduction, the compliance rate of pollutant discharge, and environmental protection tax and pollution fine. It shows that the enterprise needs to focus on improving these sample features in its subsequent development. When compared with the features of poor enterprises, the enterprise performs well in several aspects, such as research and development fees of green products, energy-saving, the number of accidental pollution incidents, and environmental complaints. The enterprise needs to continue to maintain the advantages of these indexes in the subsequent development process. At the same time, it should improve its deficiencies and approach the standard of environmental cost control of excellent enterprises. Moreover, the enterprise needs to improve its environmental cost control capability, while maintaining a steady rise in its core competitiveness.

Table 2: Analysis results of the secondary indexes with the largest sample gap feature under each primary index.

Primary index	Secondary index (excellent sample feature)	Secondary index (poor sample feature)
Internal input index of environmental protection $A1$	The update of environmental protection technology and equipment $B3$	The research and development fee of green products $B2$
Cleaner production index $A2$	Pollutant emission reduction $B7$	Energy-saving $B6$
Pollution control effect index $A3$	The compliance rate of pollutant discharge $B10$	Number of accidental pollution incidents $B12$
External influence index $A4$	Environmental protection tax and pollution fine $B13$	Environmental complaints $B15$

4. Conclusion

The environmental cost control of manufacturing enterprises in the development process is mainly explored. Through the analysis of environmental costs and other related concepts, it is found that environmental costs have long-term and economic characteristics. Therefore, it is necessary to pay attention to the management and control of environmental costs in the development process to lay a solid foundation for the development of enterprises. The principle of traditional GA is analyzed. The multiobjective GA is used to improve the BPNN, which can improve the global search ability of the BPNN, avoid the dilemma of local optimization, and improve the algorithm's computing performance. Finally, the multiobjective GA optimization is used to empirically analyze the environmental cost control ability of a paper-making enterprise. It is found that the comprehensive score of the environmental cost control capability of the enterprise is above 79. It is proved that the environmental cost control capability of the enterprise is better while obtaining indexes of insufficient development and advantages. It is necessary to further improve the environmental cost management capability of the enterprise to promote the core competitiveness of its development. At the same time, the simulation analysis indicates that the total score of enterprise environmental cost control capability is more than 79, and the overall enterprise environmental cost control capability is better. In terms of the primary evaluation index, the score of the cleaner production index is the lowest, only 75%. But the score of this index is relatively high. Compared with the same level of indexes, the scores of the pollution control index and external control index are more than 80%.

Still, there are some shortcomings; the focus is on the analysis of the principles of the research methods. Due to space limitations, the BPNN and modeling methods are briefly introduced. The investigation is more inclined to theoretical analysis and lacks certain practice. It is hoped that, in the follow-up, this model method can be applied to plenty of practical investigations to further verify the performance of the proposed model.

Data Availability

The raw data supporting the conclusions of this study can be obtained from the corresponding author upon request.

Ethical Approval

This article does not contain any studies with human participants or animals performed by any of the authors.

Consent

Informed consent was obtained from all individual participants included in the study.

Conflicts of Interest

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

Acknowledgments

This work was supported by S & T Program of Hebei (Grant No.: 22375411D).

References

[1] J. Xu, Q. Huang, C. Lv, Q. Feng, and F. Wang, "Carbon emissions reductions oriented dynamic equilibrium strategy using biomass-coal co-firing," *Energy Policy*, vol. 123, pp. 184–197, 2018.

[2] A. Coskun and N. Karaca, "KOBİ`lerde çevresel maliyetlerin sınıflandırılmasına yönelik bir öneri: metal i?şleme sektöründen bir uygulama," *Ekoloji*, vol. 18, no. 69, pp. 59–65, 2008.

[3] J. Wang, R. Wang, Y. Zhu, and J. Li, "Life cycle assessment and environmental cost accounting of coal-fired power generation in China," *Energy Policy*, vol. 115, pp. 374–384, 2018.

[4] M. M. Borrego-Marín, C. Gutiérrez-Martín, and J. Berbel, "Estimation of cost recovery ratio for water services based on the system of environmental-economic accounting for water," *Water Resources Management*, vol. 30, no. 2, pp. 767–783, 2016.

[5] D. Ayes, "Environmental and material flow cost accounting: principles and procedures," *Journal of Cleaner Production*, vol. 18, no. 13, pp. 1347-1348, 2010.

[6] J. Wang, Y. Ma, L. Zhang, R. X. Gao, and D. Wu, "Deep learning for smart manufacturing: methods and applications," *Journal of Manufacturing Systems*, vol. 48, pp. 144–156, 2018.

[7] S. Tesfamariam and R. Sadiq, "Risk-based environmental decision-making using fuzzy analytic hierarchy process (F-AHP)," *Stochastic Environmental Research and Risk Assessment*, vol. 21, no. 1, pp. 35–50, 2006.

[8] M. Lin, Z. Li, J. Liu et al., "Maintaining economic value of ecosystem services whilst reducing environmental cost: a way

to achieve freshwater restoration in China," *PLoS One*, vol. 10, no. 3, Article ID e0120298, 2015.

[9] X. Zhang, Y. Zhou, and Q. Han, "Game theory-based environmental LCC control behavior analysis," *Journal of Cleaner Production*, vol. 211, pp. 1527–1533, 2019.

[10] Y. S. Kim, T. M. Do, M. J. Kim, B. J. Kim, and H. K. Kim, "Utilization of by-product in controlled low-strength material for geothermal systems: engineering performances, environmental impact, and cost analysis," *Journal of Cleaner Production*, vol. 172, pp. 909–920, 2018.

[11] X. Zhao, Y. Zhao, S. Zeng, and S. Zhang, "Corporate behavior and competitiveness: impact of environmental regulation on Chinese firms," *Journal of Cleaner Production*, vol. 86, pp. 311–322, 2015.

[12] Y. B. Pan, X. Xiao, J. J. Li, Y. Huang, and Z. F. Zhou, "Carbon flow cost control of coal-fired power plant based on "energy flow–value flow"analysis," *Applied Ecology and Environmental Research*, vol. 17, no. 2, pp. 4869–4882, 2019.

[13] H. Li and Q. Zhang, "Multiobjective optimization problems with complicated Pareto sets, MOEA/D and NSGA-II," *IEEE Transactions on Evolutionary Computation*, vol. 13, no. 2, pp. 284–302, 2009.

[14] A. Volkanovski, B. Mavko, T. Boševski, A. Čauševski, and M. Čepin, "Genetic algorithm optimisation of the maintenance scheduling of generating units in a power system," *Reliability Engineering & System Safety*, vol. 93, no. 6, pp. 779–789, 2008.

[15] H. Zhang, B. Lennox, H. Zhang, P. R. Goulding, and A. Y. T. Leung, "A float-encoded genetic algorithm technique for integrated optimization of piezoelectric actuator and sensor placement and feedback gains," *Smart Materials and Structures*, vol. 9, no. 4, pp. 552–557, 2000.

[16] H. B. Wang, X. G. Li, P. F. Li, E. I. Veremey, and M. V. Sotnikova, "Application of real-coded genetic algorithm in ship weather routing," *Journal of Navigation*, vol. 71, no. 4, pp. 989–1010, 2018.

[17] Y. Delice, E. Kızılkaya Aydoğan, and U. Özcan, "Stochastic two-sided U-type assembly line balancing: a genetic algorithm approach," *International Journal of Production Research*, vol. 54, no. 11, pp. 3429–3451, 2016.

[18] T. Nakata, M. Sanada, S. Morimoto, and Y. Inoue, "Automatic design of IPMSMs using a genetic algorithm combined with the coarse-mesh FEM for enlarging the high-efficiency operation area," *IEEE Transactions on Industrial Electronics*, vol. 64, no. 12, pp. 9721–9728, 2017.

[19] E. C. Li and Y. Q. Ma, "The application of improved adaptive genetic algorithm in the optimization of discrete variables," *Journal of Discrete Mathematical Sciences and Cryptography*, vol. 21, no. 2, pp. 417–421, 2018.

[20] X. Wang, X. Chen, Q. Cui, and Z. Yang, "An improved two-step parameter adjustment method for the optimization of a reservoir operation function model based on repeated principal component analysis and a genetic algorithm," *Journal of Hydroinformatics*, vol. 21, no. 1, pp. 1–12, 2019.

[21] D. Gong, J. Sun, and Z. Miao, "A set-based genetic algorithm for interval many-objective optimization problems," *IEEE Transactions on Evolutionary Computation*, vol. 22, no. 1, pp. 47–60, 2018.

[22] E. Bradford, A. M. Schweidtmann, and A. Lapkin, "Efficient multiobjective optimization employing gaussian processes, spectral sampling and a genetic algorithm," *Journal of Global Optimization*, vol. 71, no. 2, pp. 407–438, 2018.

[23] A. Kumar, D. Barman, R. Sarkar, and N. Chowdhury, "Overlapping community detection using multiobjective genetic algorithm," *IEEE Transactions on Computational Social Systems*, vol. 7, no. 3, pp. 802–817, 2020.

[24] A. Balas, F. Valentina, and G. R. Emilia, Eds., *Handbook of Deep Learning Applications*, Springer, vol. 136, New York, 2019.

[25] Z. Dong, H. Jia, and M. Liu, "An adaptive multiobjective genetic algorithm with fuzzy-means for automatic data clustering," *Mathematical Problems in Engineering*, vol. 2018, Article ID 6123874, 2018.

[26] A. Roy, K. Sanjiban, and T. Y. Sekhar, "A deep learning-based artificial neural network approach for intrusion detection," *International Conference on Mathematics and Computing*, Springer, Singapore, 2017.

[27] S. Dong, Y. Zhang, Z. He, N. Deng, X. Yu, and S. Yao, "Investigation of support vector machine and back propagation artificial neural network for performance prediction of the organic Rankine cycle system," *Energy*, vol. 144, pp. 851–864, 2018.

[28] X. Z. Wang, T. Zhang, and L. He, "Application of fuzzy adaptive back-propagation neural network in thermal conductivity gas analyzer," *Neurocomputing*, vol. 73, no. 4-6, pp. 679–683, 2010.

[29] Q. Zhang, L. Zhuo, J. Li, J. Zhang, H. Zhang, and X. Li, "Vehicle color recognition using Multiple-Layer Feature Representations of lightweight convolutional neural network," *Signal Processing*, vol. 147, pp. 146–153, 2018.

[30] S. Jeschke, "Industrial internet of things and cyber manufacturing systems," *Industrial Internet of Things*, pp. 3–19, Springer, Cham, 2017.

[31] X. Yao, J. Zhou, and J. Zhang, "From intelligent manufacturing to smart manufacturing for industry 4.0 driven by next-generation artificial intelligence and further on," in *Proceedings of the 2017 5th international conference on enterprise systems (ES)*, IEEE, Beijing, China, September 2017.

[32] D. Zhang, J. Shen, P. Liu, Q. Zhang, and F. Sun, "Use of fuzzy analytic hierarchy process and environmental gini coefficient for allocation of regional flood drainage rights," *International Journal of Environmental Research and Public Health*, vol. 17, no. 6, p. 2063, 2020.

[33] P. Sefeedpari, T. Vellinga, S. Rafiee, M. Sharifi, P. Shine, and S. H. Pishgar-Komleh, "Technical, environmental and cost-benefit assessment of manure management chain: a case study of large scale dairy farming," *Journal of Cleaner Production*, vol. 233, pp. 857–868, 2019.

[34] Z. Lv, J. Guo, and H. Lv, "Safety poka yoke in zero-defect manufacturing based on digital twins," *IEEE transactions on industrial informatics, advance online publication*, 2022.

[35] Y. Zhang, Z. Pan, J. Yang et al., "Study on the suppression mechanism of (NH4) 2CO3 and SiC for polyethylene deflagration based on flame propagation and experimental analysis," *Powder Technology*, vol. 399, Article ID 117193, 2022.

[36] B. Cao, Y. Zhang, J. Zhao, X. Liu, Ł. Skonieczny, and Z. Lv, "Recommendation based on large-scale many-objective optimization for the intelligent internet of things system," *IEEE Internet of Things Journal, Advance Online Publication*, 2021.

[37] L. Yan, S. Yin-He, Y. Qian, S. Zhi-Yu, W. Chun-Zi, and L. Zi-Yun, "Method of reaching consensus on probability of food safety based on the integration of finite credible data on block chain," *IEEE Access*, vol. 9, pp. 123764–123776, 2021.

[38] Y. Feng, B. Zhang, Y. Liu et al., "A 200–225-GHz manifold-coupled multiplexer utilizing metal waveguides," *IEEE Transactions on Microwave Theory and Techniques*, vol. 69, no. 12, pp. 5327–5333, 2021.

[39] W. Zheng, X. Liu, and L. Yin, "Research on image classification method based on improved multi-scale relational network," *PeerJ Computer Science*, vol. 7, p. e613.

[40] Z. Cao, Y. Wang, W. Zheng et al., "The algorithm of stereo vision and shape from shading based on endoscope imaging," *Biomedical Signal Processing and Control*, vol. 76, Article ID 103658, 2022.

[41] K. Ma, Z. Li, P. Liu et al., "Reliability-constrained throughput optimization of industrial wireless sensor networks with energy harvesting relay," *IEEE Internet of Things Journal*, vol. 8, no. 17, pp. 13343–13354, 2021.

[42] C. Cheng and L. Wang, "How companies configure digital innovation attributes for business model innovation? A configurational view," *Technovation*, vol. 112, Article ID 102398, 2021.

[43] B. Cao, J. Zhao, Z. Lv, and P. Yang, "Diversified personalized recommendation optimization based on mobile data," *IEEE Transactions on Intelligent Transportation Systems*, vol. 22, no. 4, pp. 2133–2139, 2021.

[44] C. Lu, Q. Liu, B. Zhang, and L. Yin, "A Pareto-based hybrid iterated greedy algorithm for energy-efficient scheduling of distributed hybrid flowshop," *Expert Systems with Applications*, vol. 204, Article ID 117555, 2022.

[45] J. Li, K. Xu, S. Chaudhuri, E. Yumer, H. Zhang, and L. Guibas, "Grass: generative recursive autoencoders for shape structures," *ACM Transactions on Graphics*, vol. 36, no. 4, pp. 1–14, 2017.

[46] J. K. Gupta and S. K. Gupta, "Iot based statistical approach for human crowd density estimation-design and analysis," *Acta Informatica Malaysia*, vol. 4, no. 1, pp. 22–25, 2020.

[47] F. Azam, R. Munir, M. Ahmed, M. Ayub, A. Sajid, and Z. Abbasi, "Internet of things (iot), security issues and its solutions," *Science Heritage Journal*, vol. 3, no. 2, pp. 18–21, 2019.

[48] M. Ahmad and A. Sri, "Conceptual design impacts in new normal era: the use of artificial intelligence (AI) and internet of things (IOT) (case studies: class room And restaurant)," *Acta Informatica Malaysia*, vol. 6, no. 2, pp. 34–37, 2022.

Hindawi
Computational Intelligence and Neuroscience
Volume 2022, Article ID 3351196, 9 pages
https://doi.org/10.1155/2022/3351196

Research Article

Application of Visual Recognition Based on BP Neural Network in Architectural Design Optimization

Rui Liang,[1,2] Po-Hsun Wang,[2,3] and Linhui Hu[4]

[1]*School of Architecture and Urban Planning, Guangdong University of Technology, Guangzhou, Guangdong 510090, China*
[2]*Faculty of Innovation and Design, City University of Macau, Taipa, Macau 999078, China*
[3]*State Key Laboratory of Subtropical Building Science, South China University of Technology, Guangzhou, Guangdong 510090, China*
[4]*College of Art and Design, Guangdong University of Technology, Guangzhou, Guangdong 510090, China*

Correspondence should be addressed to Linhui Hu; hlh@gdut.edu.cn

Received 20 July 2022; Revised 30 August 2022; Accepted 8 September 2022; Published 30 September 2022

Academic Editor: Ning Cao

In order to establish the mapping relationship between architectural design parameters and building performance and optimize architectural design parameters, an architectural design optimization method based on BP neural network is proposed. The selected main design parameters of building ventilation include spacing coefficient, air outlet area, and height from the bottom of the window sill to the ground. Take the comprehensive performance of building ventilation design as the main optimization objective to optimize the building design. First, nine groups of samples of building optimization design are obtained through uniform experimental design. Then, based on the architectural design sample data obtained by BP neural network training, the mapping relationship between architectural design parameters and building performance is established, and based on this mapping, the optimal design parameters of the building are calculated. The research results have a certain reference value for architectural design optimization.

1. Introduction

In recent years, with the rapid development of China's basic industry, building structure design also needs to make continuous adjustments and changes with the development of the times [1–4]. At present, architectural design is the basic condition to ensure the safety and stability of the main body of the building structure [5]. Good structural design optimization can not only improve the appearance quality and functional value of the building but also help to improve the quality of people's daily lives [6].

However, the development of high quality has put forward higher requirements for architectural design [7, 8]. At present, most architectural designs pay too much attention to form and light function, and the current design situation of pursuing standardization and a lack of innovation conflicts with it. The architectural design needs optimization and innovation urgently [9, 10]. The main design problems include unreasonable apartment design, incomplete functional space, and insufficient community supporting facilities. Understand the important influence of design factors (such as house type and orientation) on life experience (such as ventilation).

According to statistics, among the many reasons for engineering quality problems, design-related accounts for 40%, ranking first [11–13]. It is precise because of the lack of refined design and design optimization that the defects in the design are not avoided or the landing effect is not good. The presentation effect is inconsistent with the publicity effect or causes quality defects, affecting normal use [14].

With the rapid development of computer technology, the whole society is haunted by machine learning, big data, and artificial intelligence [15, 16]. These technologies are constantly promoting the progress of human civilization and also affect everyone's studies and lives all the time. The goal of artificial intelligence is to understand the thinking mode

of the human brain so that robots can learn and imitate this intelligent and unique way of processing information and can better serve human beings, improve productivity, and promote human progress. Deep learning is in full swing in the field of computers [17–19]. In recent years, it has also been a preliminary attempt to use deep learning to recognize and classify objects in the field of robots. However, because the deep learning model requires a large number of data samples, sometimes tens of thousands or even hundreds of thousands of sample data, the sorting of target recognition parts is limited by the difficulty of identifying object image collections [20–23]. In the complex building design environment, visual recognition based on BP neural network is one of the important means for building designers to obtain building environment information [24]. However, the current vision recognition technology based on BP neural network is far from reaching the intelligent level of the human vision system and cannot recognize and understand any building environment like human beings. Therefore, the application of visual recognition in building design optimization needs further research [25–27].

This study applies the visual recognition method based on the BP neural network to architectural design optimization. Firstly, the main problems in the application of the visual recognition method in architectural design are analyzed, and then, a visual recognition system based on the BP neural network is constructed. Finally, the system is applied to the actual architectural design optimization. The research method of this study has certain reference significance for architectural design optimization.

2. The Concept and Principle of Architectural Design

2.1. Architectural Design Concept. Zhan Kesheng, who first put forward the theory of optimization method in residential building design in China, theory of "optimal residential building design method" in 1984, divides residential building design into three systems, residential use function, building technical economy, and building modeling, and then parameterizes the design factors of each system, establishes a mathematical model, and carries out systematic mathematical evaluation [28, 29]. The optimal room unit and residential unit type are obtained in turn, up to the residential area. The proposed method is based on the social development reality at that time. It is a simple theoretical analysis and research. The optimization method adopts a more complex mathematical model, but it is difficult to match the rapid development of residential buildings in China.

Urban planning is to dynamically solve and coordinate the connection between various buildings and the overall image of the building complex, to continue the history of the city, and to look forward to the future of the city from an ecological and sustainable point of view. Urban planning and design are the premise of architecture and garden construction and provide conditions for the required space [30]. The progress of urban planning and design research also provides an unprecedented broad world for the development of architecture and gardening. The design of buildings reflects the characteristics and requirements of urban planning. Our country is constantly making progress and constantly advancing with the pace of the times, and strengthening architectural design strictly reflects the characteristics of urban planning. As we are constantly innovating, we should always follow the forefront of the times, pursue innovation and development, and make our buildings more powerful.

With the continuous progress of science, people's living standards are constantly improving, and people's pursuits are also constantly updated. China's urban planning will continue to change with the pace. With "the theory of a round sky and a round place" as a whole cultural concept, square has become the ideal model of human living space, which has laid the ideal form of China's square city from the deep-seated cultural awareness. Architectural design solves the coordination between individual building functions and the urban environment under the influence of urban regulatory detailed planning [30, 31]. These will be the characteristics of the city. Urban planning is the analysis and design of urban space, that is, the coordination of the regional space of human activities and the relationship between regional spaces [32]. In terms of breadth, it was previously believed that the three periods of point planning of industrial layout, linear development of transportation layout, and area planning with the goal of creating a better environment, and improving residential and working apartments did not solve the social disease of urban development [33]. We should explore a more holistic and holistic planning, that is, to establish an overall separation and development model of "nature space human system" within the territory.

In modern cities, residential buildings are mostly dense, and local wind fields are complex, making it easy to form local wind fields. The setting of architectural design parameters directly affects the wind pressure on the building surface, thus affecting the ventilation effect of the building room. How to reduce the negative impact of adverse wind fields on residents lives is a major challenge for architectural design optimization.

2.2. Design Principle. People-oriented is at the core of the scientific concept of development. It aims to achieve people's all-round development, pays attention to people's needs, and emphasizes full respect, understanding, and support for people [32]. Architecture is to provide services for people, so the principle of being people-oriented must be emphasized in architectural design, which should not only try to meet people's basic needs, including user needs, health needs, safety needs, etc., but also fully respect people's personalities and design in combination with people's personality characteristics.

Nowadays, with the increasingly serious problem of energy shortages, how to reduce energy consumption and save energy while developing the industry has become an important research topic [34, 35]. Some nonrenewable energy is extremely valuable. If human beings exploit it without restraint, sooner or later, they will face the day of energy exhaustion. Although some renewable energy

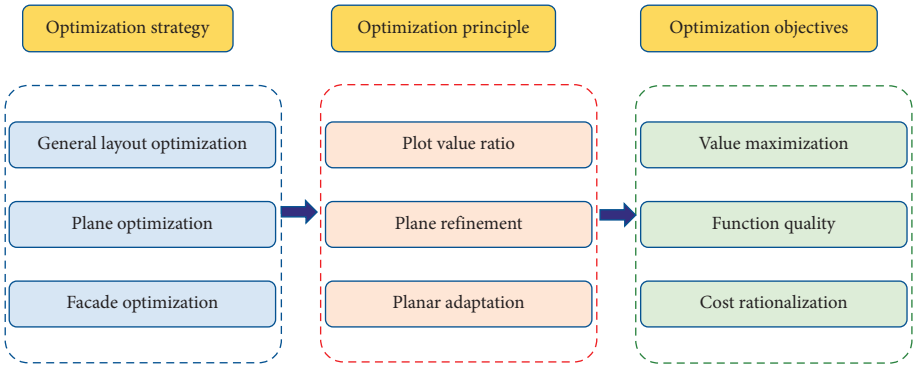


Figure 1: Correspondence diagram of optimized strategies, principles, and objectives.

sources are inexhaustible, they also need to incur a certain cost. Therefore, in the development of the construction industry, we must pursue low consumption and energy conservation, comprehensively use various energy-saving technologies and energy-saving materials in architectural design, maximize energy efficiency, reduce energy consumption, and reduce unnecessary energy waste [36].

The protection of the ecological environment and the construction of ecological civilization are the common responsibilities of all mankind. Only by building eco-friendly buildings, making them live in harmony with the ecological environment, becoming a part of the urban ecological composition, and avoiding damage to the ecological environment, can we conform to the scientific concept of development. Eco-friendly architectural design also lies in conforming to nature, respecting the objective laws of nature, and looking for ways to get close to nature. In other words, it is to combine architectural design with environmental planning [37].

No matter what type of architectural design, people should first meet the basic requirements of living comfort to ensure that the designed building is suitable for people to live. In real life, many factors will affect the comfort of buildings, such as air quality factors, temperature factors, humidity factors, light factors, and environmental factors. The common feature of these factors is that they all belong to the object of human perception. Therefore, in order to effectively create a comfortable and livable building, these factors should be fully considered in the design to ensure that people can obtain a comfortable living experience to the greatest extent.

3. Challenges Faced by Machine Learning in Related Fields

Although machine learning classification algorithms can deal with many complex classification problems, with the data becoming more complex and diverse, machine learning classification algorithm has encountered new challenges in learning objectives and classification efficiency.

3.1. Low Efficiency of Target Learning and Classification. Data in different application fields show high-dimensional characteristics. With the increase of redundant and irrelevant information in the data, the performance of machine learning classification algorithm decreases and the computational complexity increases. Machine learning classification algorithms generally need to use large samples to learn effectively. Big data does not mean that the number of training samples is sufficient. When the sample size is small and the features contain a large number of irrelevant features or noise features, the classification accuracy may be low and overfitting may occur.

Machine learning classification algorithms generally assume that the data set used for training is balanced, that is, the number of samples in each class is roughly equal, but in reality, the data are often unbalanced. Existing studies usually deal with imbalance and high-dimensional problems separately, but in practice, there are often data with the dual characteristics of imbalance and high-dimensional. At present, the strategies, principles, and objectives of machine learning optimization are shown in Figure 1.

In addition to the common binary classification problems, there are a lot of multiclassification problems in practical applications, especially the multiclassification problems of high-dimensional data, which pose challenges to the existing machine learning classification algorithms.

At present, data instances in the application of machine learning classification algorithms are represented by a large number of features. A good classification model depends on a feature set with high correlation. Removing irrelevant and redundant features can not only improve the accuracy of the model but also reduce the running time. Therefore, research on feature selection is becoming more and more important for the development of machine learning classification algorithms.

3.2. The Optimization Concept is Not Mature Enough. Integrating the green building design concept into the architectural design is an important content in the development of modern architecture. The beauty and use efficiency of buildings can be significantly improved through the green building design concept, and the energy conservation and consumption reduction and rational utilization of resources in the green building design can also have a good significance for environmental protection and have high application value. The infiltration of green building design concepts into architectural design needs to follow certain

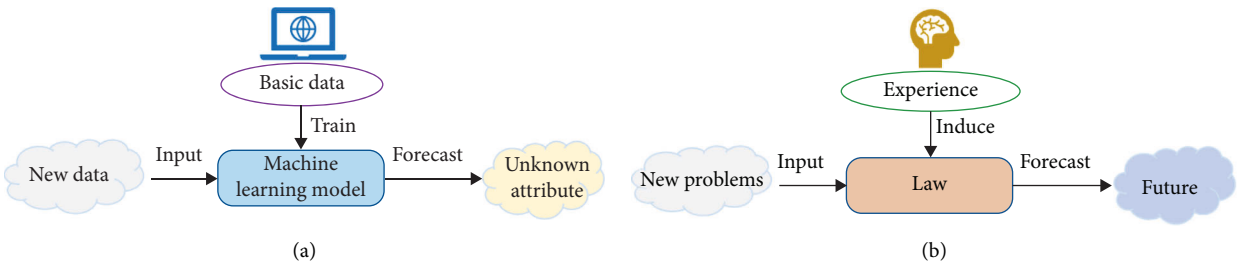


Figure 2: The difference between (a) machine learning and (b) human thinking.

principles and strategies. According to the green building design concept, discussing the building site selection, application of environmental protection materials, and optimization of the housing structure, greening of landscape and other contents in modern architectural design can further optimize the architectural design and create a high-quality building project with ecological environmental protection significance and sustainable development.

The concept of green building design has a high application value in architectural design. When designing green buildings, we often need to consider the principles of environmental protection, efficiency, livability, and economy. The infiltration of green building design concept in current architectural design makes the design of building structure more reasonable and superior. Integrating green building design concept into building site selection design, environmental protection material selection, building structure design and green landscape design can better ensure the construction quality. The application of green building design concept in architectural design is gradually increasing. The integration of green building design concept can not only ensure the artistry and practicality of architectural design itself but also further improve the effect of building energy conservation and environmental protection, thereby promoting the development and progress of the construction industry.

3.3. Machine Learning Is Rarely Used in Architectural Design Optimization. With the development of computer technology and artificial intelligence, machine learning is becoming more and more widely used. Whether in military or civil fields, there are opportunities for machine learning algorithms. Machine learning is to automatically summarize the logic or rules behind things by selecting appropriate algorithms from data and predict the development of things according to the inductive results (models) and new data. Machine learning is to let the machine find the correlation in the data through the learning of a large amount of data [38–40]. In addition to relevance, human intelligence has no causal relationship with today's artificial intelligence. Machine learning does not understand causality. It only knows correlation, so this is the biggest difference between machine learning and human thinking, as shown in Figure 2.

Because the most difficult step of machine learning is to refine the problems in real production and life into machine learning problems, what this requires is researchers' deep insight into the actual problem itself. Moreover, no matter how accurate the prediction of the machine is, its result is worthless if it is not to answer human needs. Therefore, applying BP neural network to architectural design optimization and using machine learning methods to solve the problems existing in architectural design is a new research method, which is of great value for the development of architectural design optimization methods. However, when using the BP neural network for architectural design optimization, the most important thing is to build a machine learning problem. The optimization problem of architectural design is abstracted as a machine learning problem for subsequent neural network training and parameter optimization.

4. Construction of Architectural Design Optimization System Based on Machine Learning

In order to establish the mapping relationship between building design parameters and building ventilation performance and optimize building ventilation design parameters, a method based on uniform design and neural network is proposed. Take indoor ventilation as the optimization goal. In this study, the architectural design parameters that affect the natural ventilation effect of residential buildings are divided into two categories: outdoor design parameters and indoor design parameters. The outdoor design parameters are plane layout and building spacing, and each design parameter is designed with three change levels. The indoor design parameters are two influence parameters: ventilation outlet area and the distance from the outlet to the floor), and each design parameter is designed with three change levels. The sample data of optimal design are obtained through uniform test design method, and the mapping relationship between design parameters and bridge performance is established through BP neural network training sample data. Based on this mapping, the corresponding design parameters under the optimal performance state of the architecture are calculated, and finally, the parameter optimization is completed. The optimization design of building ventilation is shown in Figure 3.

4.1. Outdoor Parameters' Setting. The most common community layouts of residential buildings are staggered, enclosed, and parallel. This paper mainly studies the impact of the three most common building layouts on the indoor natural ventilation effect, numbered A1, A2, and A3, respectively, as shown in Figure 4.

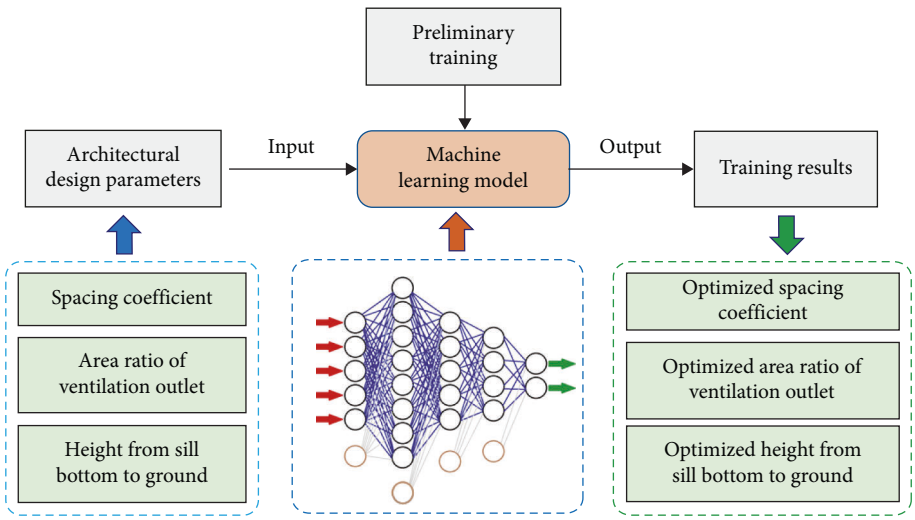


Figure 3: Optimization design of building ventilation.

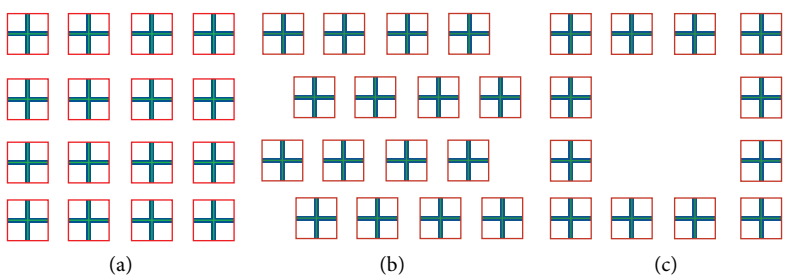


Figure 4: Three models of different building layouts: (a) parallel type; (b) staggered type; (c) enclosed type.

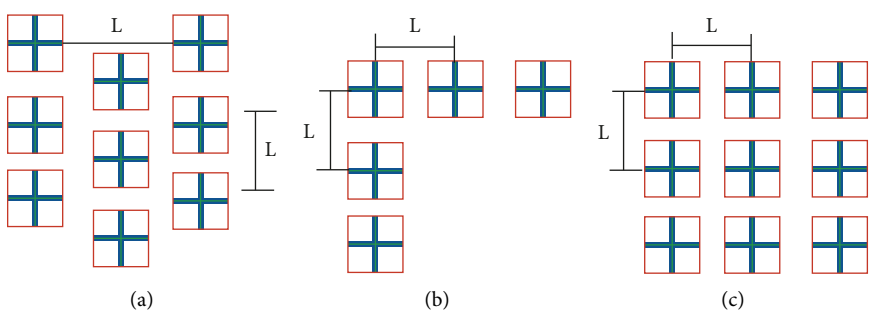


Figure 5: Schematic diagram of horizontal and vertical spacing of buildings.

The design of building spacing is one of the important factors to be considered in the plane design of buildings. The building spacing should meet the relevant provisions of ventilation, fire prevention, sunshine, daylighting, noise prevention, and so on. In this study, the influence of building spacing on natural ventilation is characterized by the spacing coefficient in the building group. The spacing coefficient represents the ratio of building spacing to the height of the main building (the highest building height), as shown in the following formula:

$$\Psi = \frac{L}{H}, \tag{1}$$

where ψ is a dimensionless variable, H is the height of the main building, m, and L is the distance between buildings, m. The horizontal and vertical spacing of buildings with different plane layouts is shown in Figure 5.

Ratio of the openable area of naturally ventilated windows to the area of ventilated rooms σ. To characterize the size of the ventilation outlet area, set three levels of change $\sigma = 8\%$, $\sigma = 10\%$, and $\sigma = 12\%$, numbered F1, F2, and F3 respectively. And set two window openings in the natural ventilation room, and the net height of the window opening is equal, which is 1.5 m. The size is achieved by changing the opening width and the size setting of different ventilation outlet areas.

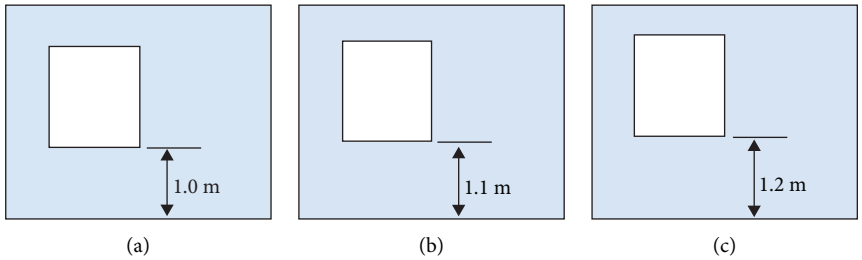


Figure 6: Schematic diagram of distance between air outlet and ground.

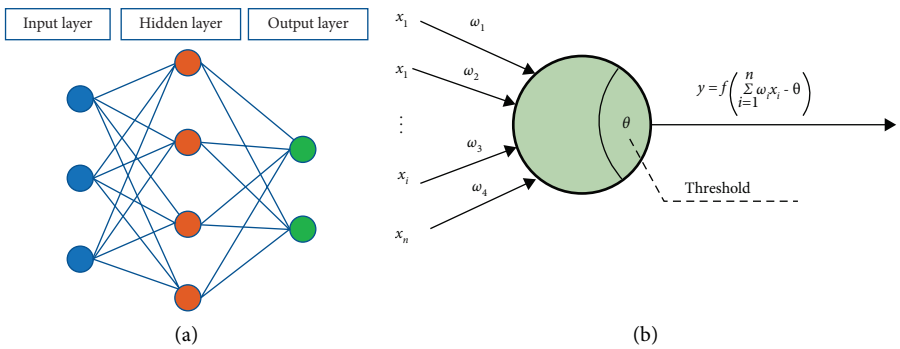


Figure 7: Schematic diagram of typical neural network structure. (a) BP neural network model. (b) Neuron topology.

Studies have shown that the optimal design of the distance from the air outlet to the ground can change the wind field characteristics of natural ventilation rooms. In daily residential buildings, the height from the bottom of the window sill to the ground is generally set in the range of 0.9–1.2 m. In this study, the height from the bottom of the window to the floor is used to characterize the influence of the distance from the air outlet to the ground (floor) on the natural ventilation effect. The height from the bottom of the window that can be naturally ventilated to the floor is set to be 0.9 m, 1.1 m, and 1.2 m, numbered H1, H2, and H3, respectively, as shown in Figure 6.

4.2. Architectural Design Optimization System Based on BP Neural Network

4.2.1. Principle of BP Neural Network. As one of the ways to realize machine learning, neural network is a computer deep learning technology that simulates human brain neural networks. The BP neural network is a typical three-layer neural network. Theoretically, it can approach any continuous function infinitely. In other words, through the training of a certain number of samples, a BP neural network can obtain the linear and nonlinear relationship between some attributes or features. As shown in Figure 7, a typical BP neural network structure mainly includes neurons (input, hidden, and output layers) and connecting lines. The connecting line between neurons represents the weighted transmission of values in a specific direction.

4.2.2. Optimization Objectives. In general, the inevitable consideration in the optimization of building ventilation design is the requirements of ventilation. At the same time, the economy should be fully considered while meeting the ventilation requirements. Therefore, the optimization objective should include the requirements of strength, deflection, and the amount of building materials (representing the economic level). This study takes the strength and deflection of the mid-span section and the concrete dosage of the whole bridge as the optimization objectives. Obviously, this is a multiobjective optimization problem. The calculation of the multiobjective optimization problem is cumbersome, and it is difficult to establish an accurate mathematical model. Therefore, it is considered to use the formula scoring method to convert the multiobjective optimization problem into a single objective optimization problem. Three optimization goals into one optimization goal are as follows:

$$K = \frac{\Psi}{\varphi} + \frac{S}{s} + \frac{H}{h}, \tag{2}$$

where K is the design comprehensive performance for building ventilation, Ψ is spacing coefficient adopted in design, 0.5–1.0, Φ is the lower limit of spacing coefficient, 0.5, S is the area ratio of ventilation outlet, range is 9%–11%, s is the lower limit of area ratio of ventilation outlet, 9%, H is height from sill bottom to ground, m, range is 1–1.2 m, and h is the lower limit of sill bottom height from the ground, m.

Then, according to formula (2), calculate the maximum building spacing coefficient, ventilation outlet area, and the height from the bottom of the windowsill to the ground corresponding to a group of building optimization design parameters under extreme conditions, and then, calculate the comprehensive performance of the building design under the building design parameters.

Table 1: Experimental results.

Number	Ψ	S (%)	H (m)	K
1	0.53	9.1	1.1	1.044
2	0.74	9.8	1.05	1.742
3	0.85	10.2	1.15	1.573
4	0.91	9.6	1.09	1.683
5	0.69	10.8	1.18	1.358
6	0.78	10.1	1.05	1.328
7	0.56	9.4	1.12	1.552
8	0.95	9.9	1.08	1.984
9	0.87	10.9	1.15	1.142

4.2.3. Optimization Calculation of Architectural Design. On the premise of clarifying the optimization objectives and main structural design parameters of building ventilation design, it is necessary to establish a mathematical model between structural design parameters and building performance in order to solve the building design parameters under optimal performance. However, it is usually impossible to establish the explicit functional relationship between building performance and design parameters. Therefore, the BP neural network is considered to establish the mapping relationship between building design parameters and building performance, so as to complete the optimization according to the mapping relationship.

The key of the BP neural network lies in the selection of training samples. The training samples need to be as representative as possible. It is best to evenly distribute the value space of the whole function. Uniform experimental design is an experimental design method based on this demand, so this study considers adopting the uniform experimental method to calculate the initial value of the model. The test factors of this study are three parameters, that is, the test has three factors, and the level number should be three times the number of factors. The comprehensive performance of building ventilation design corresponding to the test design table and each parameter combination is shown in Table 1. The data in Table 1 can be used as the sample data of BP neural network training.

4.2.4. Optimization of Architectural Design Parameters. The BP neural network can imitate the interconnection between neurons in the human brain and the processing method of information and establish the nonlinear mapping relationship between input data and output data by learning the selected training data samples. Usually, the three-layer BP neural network can solve most mapping problems. Therefore, the neural network of this study adopts a three-layer neural network.

Due to the different dimensions of the parameters in this study, there are great differences in values. In order to avoid the small value information being flooded by the large value, all data samples need to be normalized according to

$$Y = 0.8\frac{X - X_{\min}}{X_{\max} - X_{\min}} + 0.1, \tag{3}$$

where Y and X are the values of each group of parameters after normalization and before normalization, respectively,

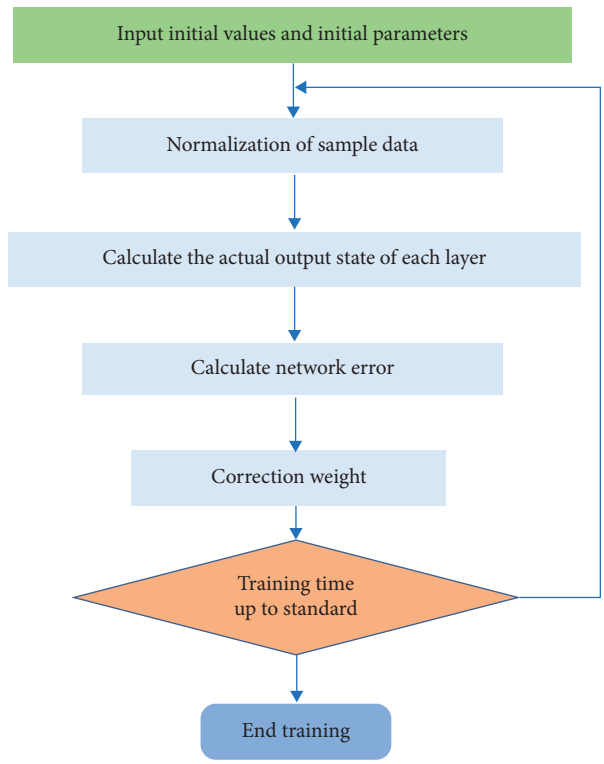


Figure 8: Optimal parameter solving process.

and $X_{\max}$ and $X_{\min}$ are the maximum and minimum values of each group of parameters, respectively. After normalization, the input and output data are all within the range of [0.1, 0.9], which can not only retain the relative information of the original data but also speed up the network learning speed and improve the network convergence ability.

In this study, an empirical formula (3) is used to approximate the number of neurons in the hidden layer:

$$p = \sqrt{m + n + a}, \tag{4}$$

where p is the number of neurons in the hidden layer, m is the number of input layer units, n is the number of output layer units, and a is a positive integer between [1, 10].

In the calculation process, the number of training is 30. The BP neural network established in this study is used to process the sample data in Table 1, normalize the sample data in Table 1 according to formula (3), and then process it with the BP neural network model. The process is shown in Figure 8.

According to the training results of the neural network, the spacing coefficient used in the design optimization of building ventilation design is 0.88, the area ratio of ventilation outlets is 10.02%, the height from the bottom of the window sill to the ground is 1.1 m, and the value of K is 1.521.

5. Discussion

The research for this study is a small part of the related research in the ventilation optimization of architectural design, and the existing machine learning is applied to the

design optimization. The research for this paper involves three kinds of architectural design parameters. Based on the idea of this research, we can further consider the impact of more architectural design parameters on the natural ventilation effect in the future, such as building density, building height, and ventilation outlet shape.

The design spacing coefficient for building ventilation is 0.88, the area ratio of ventilation outlets is 10.02%, and the height from the bottom of the window sill to the ground is 1.1 m.

In addition, the concept of green building can also be taken into account in the model of this study. The concept of green building design has high application value in architectural design. When designing green buildings, we often need to consider the principles of environmental protection, efficiency, livability, and economy. The infiltration of green building design concept into current architectural design makes the design of building structure more reasonable and superior. Integrating green building design concept into building site selection design, environmental protection material selection, building structure design, and green landscape design can better ensure the construction quality. The application of green building design concept in architectural design is gradually increasing. The integration of green building design concept can not only ensure the artistry and practicality of architectural design itself but also further improve the effect of building energy conservation and environmental protection, thereby promoting the development and progress of the construction industry. The research methods and means of the study can provide data support for scholars at home and abroad in theory and methods.

6. Conclusions

Based on the analysis of the main principles and characteristics of current architectural design optimization, this study selects the optimization objectives and parameters of architectural design. Aiming at the optimization of building ventilation design, the building design optimization is carried out based on the method of the BP neural network. The main parameters of architectural design optimization are obtained through calculation. The main conclusions are as follows:

(1) The optimization objective of architectural structural design parameters is determined, and three main structural design parameters that have a significant impact on architectural design optimization are summarized.

(2) The sample data for building design parameter optimization are obtained through the experiment. Then, a BP neural network is used to establish the mapping relationship between building design parameters and building ventilation performance. Finally, based on the existing data and mapping relationship, the optimal structural design parameters of the bridge are calculated.

(3) According to the BP neural network training results, the optimal design parameters used in building ventilation design optimization are as follows: the spacing coefficient is 0.88, the ventilation outlet area ratio is 10.02%, and the height from the bottom of the windowsill to the ground is 1.1 m.

Data Availability

The dataset used in this paper can be obtained from the corresponding author upon request.

Conflicts of Interest

The authors declared that they have no conflicts of interest regarding this work.

Acknowledgments

The 14th Five-Year Plan for the Development of Philosophy and Social Science in Guangzhou (2021GZGJ283). Supported by State Key Lab of Subtropical Building Science, South China University of Technology.

References

[1] K. Shea, R. Aish, and M. Gourtovaia, "Towards integrated performance-driven generative design tools," *Automation in Construction*, vol. 14, no. 2, pp. 253–264, 2005.

[2] L. Moreno-De-Luca and O. J. B. Carrillo, "Multi-objective heuristic computation applied to architectural and structural design: a review," *International Journal of Architectural Computing*, vol. 11, no. 4, pp. 363–392, 2013.

[3] Z. Bian and Z. Q. Zhao, "Generation mechanism of architectural design inspiration based on the theory of brain subconsciousness," *Neuroquantology*, vol. 16, no. 5, pp. 446–453, 2018.

[4] J. R. A. Maier, G. M. Fadel, and D. G. Battisto, "An affordance-based approach to architectural theory, design, and practice," *Design Studies*, vol. 30, no. 4, pp. 393–414, 2009.

[5] N. A. Vasilenko, "Parametric optimization of an architectural object's form as a method to improve its energy efficiency," *IOP Conference Series Materials Science and Engineering*, p. 552, 2019.

[6] B. Steiner, E. Mousavian, F. M. Saradj, M. Wimmer, and P. Musialski, "Integrated structural-architectural design for interactive planning," *Computer Graphics Forum*, vol. 36, no. 8, pp. 80–94, 2017.

[7] Z. Z. Su and W. Yan, "A fast genetic algorithm for solving architectural design optimization problems," *Artificial Intelligence for Engineering Design, Analysis and Manufacturing*, vol. 29, no. 4, pp. 457–469, 2015.

[8] G. Goldman, "Computer Graphics and Architectural design," Computer Graphics-US, vol. 25, no. 3, pp. 174–177, 1991.

[9] A. Yeretzian, H. Partamian, M. Dabaghi, and R. Jabr, "Integrating building shape optimization into the architectural design process," *Architectural Science Review*, vol. 63, no. 1, pp. 63–73, 2019.

[10] G. Mahalingam, "Representing architectural design using virtual computers," *Automation in Construction*, vol. 8, no. 1, pp. 25–36, 1998.

[11] R. F. Passarini, J. M. Farines, J. M. Fernandes, and L. B. Becker, "Cyber-physical systems design: transition from functional to architectural models," *Design Automation for Embedded Systems*, vol. 19, no. 4, pp. 345–366, 2015.

[12] T. Wortmann, A. Costa, G. Nannicini, and T. Schroepfer, "Advantages of surrogate models for architectural design optimization," *Artificial Intelligence for Engineering Design, Analysis and Manufacturing*, vol. 29, no. 4, pp. 471–481, 2015.

[13] T. Mariani, T. E. Elita Colanzi, and S. R. Regina Vergilio, "Preserving architectural styles in the search based design of software product line architectures," *Journal of Systems and Software*, vol. 115, pp. 157–173, 2016.

[14] X. Shi and W. J. Yang, "Performance-driven architectural design and optimization technique from a perspective of architects," *Automation in Construction*, vol. 32, pp. 125–135, 2013.

[15] G. Berseth, B. Haworth, M. Usman et al., "Interactive architectural design with diverse solution exploration," *IEEE Transactions on Visualization and Computer Graphics*, vol. 27, no. 1, pp. 111–124, 2021.

[16] Y. H. Lu, S. W. Wang, Y. Zhao, and C. Yan, "Renewable energy system optimization of low/zero energy buildings using single-objective and multi-objective optimization methods," *Energy and Buildings*, vol. 89, pp. 61–75, 2015.

[17] G. Canestrino, "Considerations on optimization as an architectural design tool," *Nexus Network Journal*, vol. 23, no. 4, pp. 919–931, 2021.

[18] W. H. Huang, L. P. Lei, and G. Fang, "Comparison between four flow stress models characterizing the constitutive behavior of hot deformation of 40Mn steel," *Journal of Materials Engineering and Performance*, vol. 30, no. 12, pp. 9149–9164, 2021.

[19] L. D. Tang, S. Q. Yuan, Y. Tang, and Z. Qiu, "Optimization of impulse water turbine based on GA-BP neural network arithmetic," *Journal of Mechanical Science and Technology*, vol. 33, no. 1, pp. 241–253, 2019.

[20] W. Q. Li, G. H. Xu, Q. H. Xing, and M. Lyu, "Application of improved AHP-BP neural network in CSR performance evaluation model," *Wireless Personal Communications*, vol. 111, no. 4, pp. 2215–2230, 2020.

[21] D. Q. Zhou and Y. H. Guo, "Radar target identification using neural network classifier based on HMM framework," *Journal of Infrared and Millimeter Waves*, vol. 20, no. 2, pp. 107–110, 2001.

[22] L. G. Xu, Y. J. Huang, J. F. Yue et al., "Soot blowing optimization for platen superheater of coal-fired power plant boiler based on heat loss analysis," *Asia-Pacific Journal of Chemical Engineering*, vol. 17, no. 3, 2022.

[23] Z. Y. Lu, Q. Zhang, X. Zhang, and C. J. Tang, "Equivalent grouping-cascaded BP network model and its applications," *Acta Electronica Sinica*, vol. 38, no. 6, pp. 1349–1354, 2010.

[24] K. Lee, J. Park, and S. Lee, *Effectiveness of Different Target Coding Schemes on Networks in Financial engineering*, Springer, Berlin Heidelberg, 2005.

[25] M. L. Feldmann, D. Chrusciel, A. Pohlmann et al., "Architectural and engineering fees from the public institutional perspective," *Journal of Management in Engineering*, vol. 24, no. 1, pp. 2–11, 2008.

[26] D. Zhou, "A new hybrid grey neural network based on grey verhulst model and BP neural network for time series forecasting," *International Journal of Information Technology and Computer Science*, vol. 5, no. 10, pp. 114–120, 2013.

[27] T. W. Liao and L. J. Chen, "Manufacturing process modeling and optimization based on multi-layer perceptron network," *Journal of Manufacturing Science and Engineering*, vol. 120, no. 1, pp. 109–119, 1998.

[28] K. Dou and X. Sun, "Long-term weather prediction based on GA-BP neural network," *IOP Conference Series: Earth and Environmental Science*, vol. 668, no. 1, p. 012015, 2021.

[29] B. Xu and X. Yuan, "A novel method of BP neural network based green building design-the case of hotel buildings in hot summer and cold winter region of China," *Sustainability*, vol. 14, no. 4, p. 2444, 2022.

[30] O. O. Demirba and H. Demirkan, "Focus on architectural design process through learning styles[J]," *Design Studies*, vol. 24, no. 5, pp. 437–456, 2003.

[31] H. Yan, J. Wang, X. Li, and L. Guo, "Architectural design and evaluation of an efficient Web-crawling system," *Journal of Systems and Software*, vol. 60, no. 3, pp. 185–193, 2002.

[32] J. U. Wieding and T. Kretschmar, "Application of a network structure in a knowledge-based system for medical diagnostics," *Computer Methods and Programs in Biomedicine*, vol. 33, no. 3, pp. 159–163, 1990.

[33] N. Gislason, "Architectural design and the learning environment: a framework for school design research," *Learning Environments Research*, vol. 13, no. 2, pp. 127–145, 2010.

[34] T. Hijikata, H. Wakisaka, and T. Yohro, "Architectural design, fiber-type composition, and innervation of the rat rectus abdominis muscle," *The Anatomical Record*, vol. 234, no. 4, pp. 500–512, 1992.

[35] M. Pearce, A. K. Goel, I. L. Kolodner, C. Zimring, L. Sentosa, and R. Billington, "Case-based design support: a case study in architectural design," *IEEE Expert*, vol. 7, no. 5, pp. 14–20, 1992.

[36] C. S. Chan, "Cognitive processes in architectural design problem solving," *Design Studies*, vol. 11, no. 2, pp. 60–80, 1990.

[37] N. Talebi, A. M. Nasrabadi, and I. Mohammad-Rezazadeh, "Estimation of effective connectivity using multi-layer perceptron artificial neural network," *Cognitive Neurodynamics*, vol. 12, no. 1, pp. 21–42, 2018.

[38] K. C. Lee, I. Han, and Y. Kwon, "Hybrid neural network models for bankruptcy predictions," *Decision Support Systems*, vol. 18, no. 1, pp. 63–72, 1996.

[39] H. R. Lin, C. H. Wang, C. J. Chen et al., "Neural bursting and synchronization emulated by neural networks and circuits," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 68, no. 8, pp. 3397–3410, 2021.

[40] Y. Xia and G. Feng, "An improved neural network for convex quadratic optimization with application to real-time beamforming," *Neurocomputing*, vol. 64, pp. 359–374, 2005.

Hindawi
Computational Intelligence and Neuroscience
Volume 2022, Article ID 8906385, 12 pages
https://doi.org/10.1155/2022/8906385

Research Article

Application of Price Competition Model Based on Computational Neural Network in Risk Prediction of Transnational Investment

Xiuxiu Chen [1] **and Xiaoxin Huang** [2]

[1] *School of Digital Economy & Trade, Wenzhou Polytechnic, Wenzhou, Zhejiang 325035, China*
[2] *School of Finance and Trade, Wenzhou Business College, Wenzhou, Zhejiang 325035, China*

Correspondence should be addressed to Xiaoxin Huang; 00204145@wzbc.edu.cn

Received 20 July 2022; Revised 31 August 2022; Accepted 9 September 2022; Published 29 September 2022

Academic Editor: Ning Cao

Aiming at the scenario where edge devices rely on cloud servers for collaborative computing, this paper proposes an efficient edge-cloud collaborative reasoning method. In order to meet the application's specific requirements for delay or accuracy, an optimal division point selection algorithm is proposed. A kind of multichannel supply chain price game model is constructed, and nonlinear dynamics theory is introduced into the research of the multichannel supply chain market. According to the actual competition situation, the different business strategies of retailers are considered in the modeling, which makes the model closer to the actual competition situation. Taking the retailer's profit as an indicator, the influence of the chaos phenomenon on the market performance is analyzed. Compared with the previous studies, this thesis uses nonlinear theory to better reveal the operating laws of the economic system. This paper selects company A in the financial industry to acquire company B in Sweden. It is concluded that company B is currently facing financial difficulties, but its brand and technical advantages are far superior to company A. The indirect financial risk index of company B, that is, the investment environment, is analyzed, and the final investment environment score of the country where company B is located is 90 points, which is an excellent grade by scoring the investment environment of the target enterprise. Combining the investment environment score and the alarm situation prediction score, it is concluded that the postmerger financial risk warning level of company A is in serious alarm.

1. Introduction

Under the current situation of global economic integration, almost all countries in the world play the roles of investors and investees at the same time [1]. A large number of industrial investors are often not limited to looking for investment opportunities in their own countries, and the content of project evaluation such as consideration and income estimation has become a topic of constant research and attention from people in the investment community of various countries, especially the financial benefit evaluation of the project, which has a crucial impact on the implementation and decision-making of the project [2]. The project financial benefit evaluation is to identify and analyze the investment, cost, income, tax, and profit, and other financial expenses and benefits proposed in the project feasibility study report on the basis of the current national fiscal and taxation system and relevant laws, and to estimate the project after the project is completed and put into operation [3]. Project financial benefit evaluation is an important part of project evaluation. The project financial benefit evaluation mainly includes the evaluation of the project's profitability and solvency. Before making a decision on an investment project, risks must be weighed against its expected benefits [4]. Especially for transnational investment projects, the operating environment of the project will have different degrees of risks in terms of economy and politics. Only by fully estimating the possible related risks and correctly predicting the possible returns can investors evaluate investment decisions more carefully [5].

Papers on neural networks are emerging in an endless stream, and their applications in many fields are becoming more and more extensive. Among them, the proposal of a support vector machine has successfully solved the problems of pattern recognition, regression, and density estimation. The proposal of the BP network parallel algorithm has ended the problem that the network learning speed is too slow and it is difficult to adapt to the analysis of large-scale engineering structures. The proposed correction factor can improve the convergence performance of the network very well. The proposal of a conjugate gradient algorithm further improves the training rate of the BP network. These works have strongly promoted the application of neural networks in structural engineering.

The choice of financial strategy for transnational operations determines the orientation and mode of the allocation of financial resources of an enterprise and affects the behavior and efficiency of financial management activities [6]. The financial strategy of a multinational operation is to raise the necessary capital in order to adapt to the overall competitive strategy and to effectively manage and use the capital within the organization. It is an important part of the overall strategy. The financial strategy of multinational operations reflects the comprehensive support for the business strategy and is the implementation and guarantee system of the business strategy [7]. Through dynamic long-term planning, the financial strategy of multinational operations continuously expands the scale and duration of financial resources, improves the rationality of the capital structure, and gives full play to the maximum benefits of financial resources. Correctly formulating and effectively implementing the financial strategy of multinational operations will effectively increase the value of the company [8]. The inflation risk, exchange rate risk, and interest rate risk faced by multinational companies are closely related to national policies. If the operator does not understand the country's political and economic policies and how these policies affect the country's economic development, the operator will not be able to correctly assess the country's risk profile, and it will be impossible to effectively predict inflation, exchange rates, interest rates, and other economic parameters, and it will not be able to provide the necessary macroeconomic environment analysis for the financial decision-making of their own enterprises [9]. When evaluating a country's business environment and economic risks, in addition to measuring economic indicators such as GDP growth rate, inflation rate, and trade deficit, companies should also consider whether the country's regulatory and legal system is fair and whether the government maintains currency values and policies, stability, etc. The better a country's economic situation, the less likely its government will act to harm the interests of multinational corporations operating in that country.

In this paper, an edge-cloud collaborative computing deep neural network method ECCI is proposed based on model compression and intermediate data compression and can select the best dividing point according to application requirements. By using the CPSCA channel pruning algorithm to compress the edge-side model, it can not only further reduce the number of parameters and computation on the edge side, thereby reducing the resource requirements for edge devices, but also effectively accelerate the edge-side model. By adopting the "three-step compression" strategy for the intermediate data to be transmitted, the transmission delay of the intermediate data can be greatly reduced. The optimal division point selection process is modeled as an integer linear programming problem to solve, the optimal division point that can meet the time or accuracy requirements of the application is selected according to application requirements, system factors, and network conditions, and the model is actually deployed. Online direct selling channel retailers will not only consider operating profits but also their market share in their operations. At the same time, traditional offline retailers will not only consider their own profits but also other retailers when formulating price strategies. This paper analyzes the stability of the model equilibrium points and finds three unstable bounded equilibrium points and one Nash equilibrium point with local stability. In this paper, company A in the financial industry is selected to acquire company B in Sweden, and the investment environment of the country where company B is located is rated as excellent. It came to the conclusion that the financial risk level is in serious alarm.

2. Related Work

Relevant scholars have established a perceptron without a hidden layer to describe the structural design problem and the prototype of the structural design neural network model, thus adding a new intelligent method to structural engineering [10]. Scholars further expanded the application of neural networks in structural engineering and discussed the application of BP networks in structural engineering. In the following years, Avdjiev successfully described the pattern of structural materials by using neural networks and experimental data; researchers used neural networks to identify the characteristic information of structural frequency response functions, making it a reality to use neural networks to diagnose structural damage [11]. The nonlinear analysis of the structure is successfully carried out by using the neural network. The related scholars put forward a BP neural network application framework, which has a great influence on the later period, combined with the practical application of BP network in structural engineering [12]. Since then, the application of neural networks in structural engineering has become more and more extensive, and it has gradually become one of the important means used in structural engineering.

Related scholars have analyzed the problem of online direct sales in dual-channel supply chains under random demand and the pricing of traditional retail channels where retailers are responsible [13]. Related scholars have studied the impact of production decisions and optimal pricing in dual-channel supply chains under normal circumstances and when demand changes in emergencies [14]. The researchers analyzed how to adjust the sales price of electronic channels in the supply chain to affect the demand of the whole supply chain so as to achieve the coordination of supply and demand in the whole supply chain [15]. Related

scholars have analyzed the pricing competition and mutual coordination between manufacturers and retailers in the dual-channel supply chain under the conditions of demand uncertainty and joint promotion [16].

Relevant scholars have studied the wholesale price contract and shared profit contract model of the dual-channel supply chain under random demand, analyzed the influence of dual-channel supply chain decision-making and coordination mechanism under demand uncertainty, and proposed a dual-channel supply chain in demand and profit-sharing contracts under certain circumstances [17]. Relevant scholars have studied the pricing problem of the dual-channel supply chain consisting of a single manufacturer and a single retailer when the demand changes, and established a profit and revenue model of the dual-channel supply chain system when the demand has changed [18].

Relevant scholars advocate that the technological innovation of enterprises should be assessed from five aspects, including strategic innovation ability, intellectual resource ability, information resource ability, innovative organizational ability, and innovation basic ability, including both qualitative and quantitative aspects [19]. Through the statistics of these indicators, you can not only draw the situation of enterprises in various subaspects of innovation, and consider them comprehensively, but also see the overall status of technological innovation.

Relevant scholars pointed out that financial risk early warning of cross-border mergers and acquisitions refers to the real-time control and forecasting of financial risks that enterprises may or will face by setting and observing changes in some sensitive early warning indicators based on financial accounting information [20]. Financial early warning includes two aspects: financial risk and early warning. Financial risks can lead to the inability of an enterprise to pay, inability to pay due debts or expenses, and economic phenomena such as insolvency, including operational failure, business failure, insolvency, and insolvency. Early warning refers to issuing warnings in advance to avoid or minimize possible losses.

3. Methods

3.1. Overall Architecture of Computational Neural Network. As shown in Figure 1, it is the overall architecture of our proposed edge-cloud collaborative reasoning algorithm ECCI based on model compression and data compression. Specifically, it is mainly divided into four stages: the side model pruning stage, the intermediate data compression stage, the division point selection stage, and the actual model deployment stage.

The main tasks of each stage are summarized as follows:

(1) Side model pruning stage: In order to reduce the computing delay of edge devices, we use different layers as quasi-dividing points, and adopt the CPSCA channel pruning algorithm. By pruning the division layer and the previous network structure, multiple deployable compression models with different pruning ratios can be obtained for selection, and the accuracy of the obtained model and the size of the intermediate transmission data are recorded.

(2) Intermediate data compression stage: In order to reduce the intermediate data transmission delay, we propose a three-step compression strategy for the intermediate data to achieve the purpose of compressing the intermediate data size.

(3) Best division point selection stage: In order to be able to select the best division point that can meet the user's needs according to the specific requirements of the application for computing delay or model accuracy and changes in system factors, we will look for the edge device and the cloud server. The process of optimal division of points is modeled as an integer linear programming (ILP) problem to be solved.

(4) Model actual deployment stage: The selected best model is actually deployed, part of the network before the division point is deployed to the edge device, and part of the network after the division point is deployed to the cloud server.

3.2. Side Model Pruning Method. In order to further reduce the computing delay of edge devices and the requirements for computing resources of edge devices in the process of edge-cloud collaborative reasoning of deep neural network models, we propose a model compression method, using different layers as quasidivided points. The first half of the network on the edge side is compressed to generate several deployable models with different side-side compression ratios (e.g., 0%, 30%, 60%, 90%) for selection. In order to ensure that the partially compressed model still has high accuracy, and considering the powerful computing power of the cloud server, the second half of the network deployed on the cloud side is not pruned, and the original model structure remains unchanged. In this process, the choice of model compression method is very important. An effective compression method should be able to compress the model to a small enough size while keeping the accuracy of the model within the acceptable range of the application.

We have learned that the attention mechanism is an effective algorithm for measuring the importance of channels. The dual attention mechanism SCA, which combines channels and spaces, has excellent performance. It can measure neural network models from the two dimensions of space and channel.

Therefore, we use the proposed CPSCA pruning algorithm based on the attention mechanism to prune the division layer and the previous network, and the network structure after that remains unchanged.

3.3. Intermediate Data Compression Strategy. Due to the data amplification effect of the intermediate layer, the intermediate data transmission delay accounts for a very high proportion of the total end-to-end computing delay. Therefore, if we want to achieve the purpose of accelerating edge-cloud collaborative computing, how to effectively

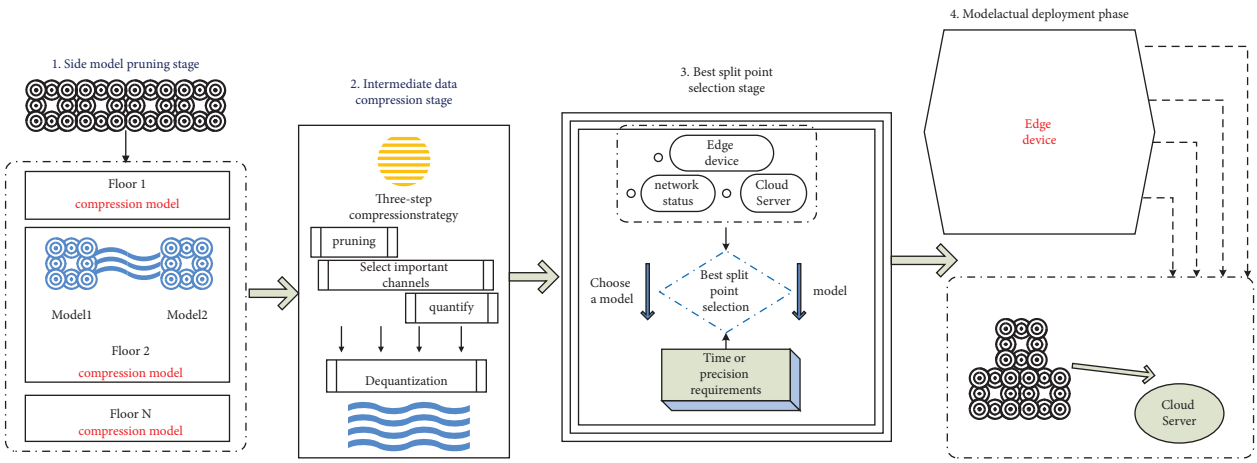


Figure 1: ECCI overall architecture.

reduce the intermediate data transmission delay is a key issue that needs to be solved urgently.

However, the existing work on edge-cloud collaborative reasoning mainly focuses on how to divide the model, and there is less research on how to compress the intermediate feature data.

Therefore, in order to reduce the intermediate data transmission delay, we propose a "three-step compression" strategy for the intermediate data that the edge device needs to transmit to the cloud server.

First, the CPSCA channel pruning technique we adopted in the side model pruning stage can effectively reduce the size of the intermediate data output by the partition layer while reducing the size of the side model. We take the most important layer type in the deep neural network, the convolutional layer, as an example, and compare the effect of channel pruning operation for the side model on the size of the intermediate output data. The size of the output feature map Xoutput is obtained as

$$S_{\text{output}} = \frac{(S-1)\left(H_{\text{input}}-F-2P\right)}{M\left(W_{\text{input}}+F+2P\right)}. \tag{1}$$

Among them, M is the number of output channels in the convolution layer convi, F is the size of the filter, S is the operation step size, and P is the size of the padding. If the pruning ratio of the convolutional layer convi is r, the size of the output feature map obtained after pruning is as follows:

$$S^{P}_{\text{output}} = \frac{(S+1)(r-1)\left(H_{\text{input}}-F-2P\right)}{M\left(W_{\text{input}}+F+2P\right)}. \tag{2}$$

In our actual pruning operation, not only the division layer is pruned but also the division layer and the previous layers are pruned.

For the intermediate output feature map that the partition layer needs to transmit to the cloud server, we use the attention mechanism to select more important regions in the feature map for transmission. The so-called attention mechanism is actually similar to the human visual system. By scanning the global image, it pays more attention to the key areas related to target judgment, while ignoring the unimportant areas. In the process of pruning using the CPSCA algorithm, the channel weights generated by the SCA attention module can accurately measure the importance of each channel of the output feature map. Based on this, we select the key channel that plays an important role in the recognition result in the intermediate output feature map for transmission, which can not only effectively reduce the intermediate data transmission delay but also improve the information processing efficiency of the cloud server after receiving the key information. The "three-step compression" strategy is shown in Figure 2.

Considering that the intermediate data to be transmitted is a 32 bit floating-point type, which not only occupies a high memory but also requires a large amount of calculation, we propose to further use the quantization operation for the important information selected based on the attention mechanism.

We found through experiments that compressing data to 8 bits has better performance than other quantization bits.

In fact, INT8 quantization is a commonly used quantization operation in the industry, which not only compresses data size efficiently but also causes minimal loss of precision. Therefore, we employ INT8 quantization during data compression.

The formula for quantizing data of type FP32 to type INT8 is as follows:

$$M_Q = \sqrt{M_F S} - Z. \tag{3}$$

Conversely, the formula for dequantizing data of type INT8 to type FP32 is as follows:

$$M_F = \frac{S - M_Q}{Z}. \tag{4}$$

MF represents the original intermediate data to be transmitted, which is 32 bit floating-point data; MQ represents the quantized intermediate data, which is 8 bit integer data; Z represents the quantized fixed-point value corresponding to 0 floating point value; and S and Z can be obtained by the following formulas:

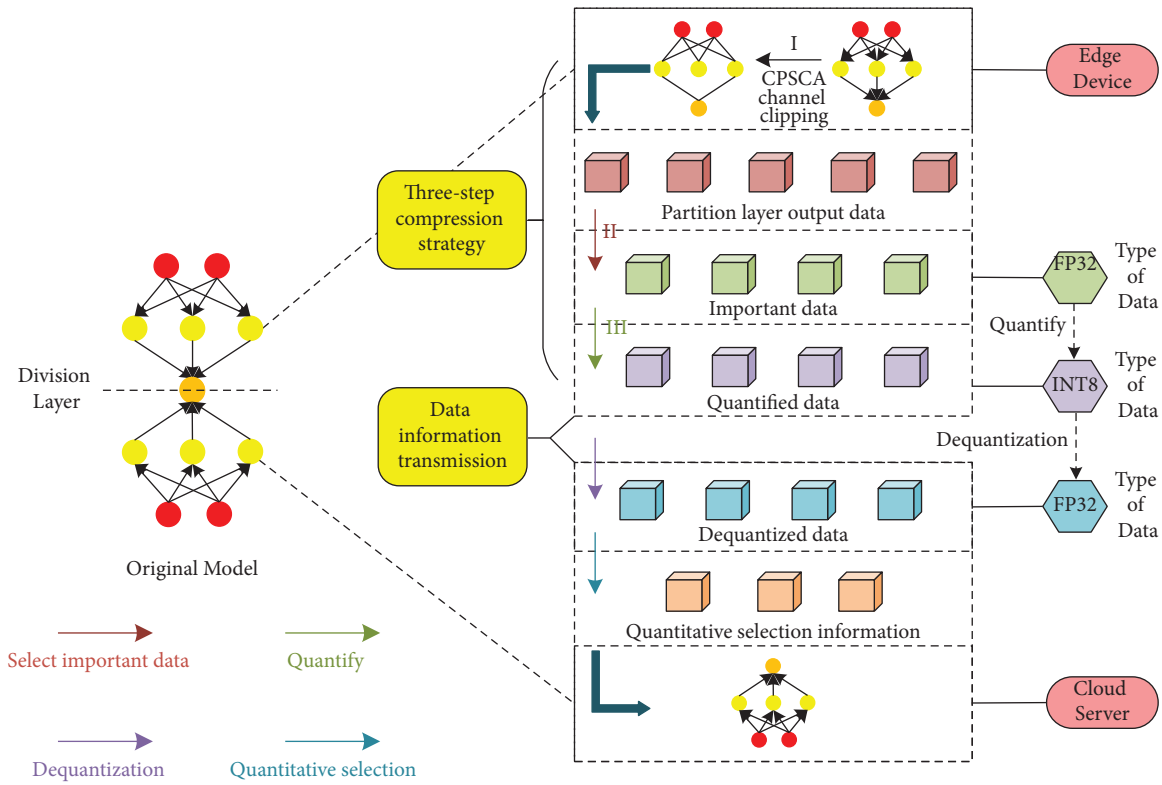


Figure 2: "Three-step compression" strategy.

$$S = \left(M_F^{\max} + M_F^{\min}\right)\left(M_Q^{\max} - M_Q^{\min}\right),$$

$$\% \quad Z = \begin{cases} \left|M_Q^{\max} - M_F^{\max}\right|, M_Q^{\max} \in (-\infty, 0), \\ M_Q^{\max} - M_F^{\min}, M_Q^{\max} \in \left(0, M_F^{\min}\right), \\ M_F^{\max} - M_Q^{\min}, M_Q^{\max} \in \left(M_F^{\min}, M_F^{\max}\right), \\ \left|M_F^{\min} - M_Q^{\min}\right|, M_Q^{\max} \in \left(M_F^{\max}, +\infty\right). \end{cases} \tag{5}$$

3.4. Optimal Division Point Selection Algorithm. In order to determine the optimal division point between edge devices and cloud servers according to the application's different preferences and specific requirements for computing delay or inference accuracy, we propose a dynamic optimal division point selection algorithm that comprehensively considers the DNN model layer. Factors such as type, network bandwidth, device computing power, and the application's specific requirements for latency or accuracy are used to determine the best division point that can meet the application requirements.

First, the total computational delay of the model consists of the following three parts:

$$T_{ij} = \frac{t_{ij}^{\text{edge}} + t_{ij}^{\text{cloud}}}{2t_{ij}^{\text{edge}\longrightarrow\text{cloud}}}. \tag{6}$$

Make predictions by building a regression model. The delay in transmitting intermediate data from the edge device to the cloud server can be calculated by the following formula:

$$t_{ij}^{\text{edge}\longrightarrow\text{cloud}} = D_{ij}R\left(\gamma_1^P + \gamma_2^Q - 2\gamma_3^S\right). \tag{7}$$

The first step is model pruning, the second step is to select important channels, and the third step is the compression ratio of the data quantization operation to the intermediate data, and R is the network bandwidth. Therefore, we can define the objective function for calculating the delay as

$$f = \frac{\prod_{i=1}^{N}\prod_{j=1}^{K}\left(x_{ij}t_{ij}^{\text{edge}} + x_{ij}t_{ij}^{\text{cloud}}\right)}{\prod_{i=1}^{N}\prod_{j=1}^{K} 2x_{ij}t_{ij}^{\text{edge}\longrightarrow\text{cloud}}}. \tag{8}$$

Among them, when $x_{ij} = 1$, it means that the i-th layer is determined as the division layer, and the j-th model of the i-th layer is determined as the actual deployed model. In our algorithm, only one model in one partition layer can be selected for deployment, so constraints need to be set to ensure the uniqueness of the decision variable x_{ij}:

$$\prod_{i=1}^{N}\prod_{j=1}^{K} x_{ij} = 2\sqrt{5} - 1. \tag{9}$$

Indicates that only one x_{ij} has a value of 1, and the rest are all 0.

We can define the accuracy objective function as

$$g = A_{ij}^{O}\frac{\Delta A_{ij}^P - \Delta A_{ij}^Q}{\Delta A_{ij}^P + \Delta A_{ij}^S}. \tag{10}$$

We model the process of choosing the best partition point for the model as an integer linear programming (ILP) problem to solve. First, judge the given time requirement or accuracy requirement. For the given accuracy requirement A, select the compression model j^* in the partition layer i^* to minimize the total delay T.

Thus, the problem can be expressed as

$$\min_{x_{ij}} f = \frac{\prod_{i=1}^{N}\prod_{j=1}^{K}\left(x_{ij}t_{ij}^{\text{edge}} + x_{ij}t_{ij}^{\text{cloud}}\right)}{\prod_{i=1}^{N}\prod_{j=1}^{K} 2x_{ij}t_{ij}^{\text{edge}\longrightarrow\text{cloud}}}, \text{s.t.} \prod_{i=1}^{N}\prod_{j=1}^{K} x_{ij} = 2\sqrt{5} - 1 \quad .(11)$$

By solving the above integer linear programming (ILP) problem, the optimal division point that can meet the application requirements can be selected during the model division process.

3.5. Price Competition Model. This paper studies a two-tier supply chain consisting of a retailer in a traditional offline channel and an online direct-selling retailer. Consumers can get the products they need from both online direct sales channels and traditional offline channels:

Hypothesis 1: The retailer of the online direct sales channel and the retailer of the traditional offline channel sell similar products to consumers, and their competition rules conform to the Bertrand price competition model.

Hypothesis 2: Neither the online direct-selling channel retailers nor the traditional offline channel retailers can fully grasp the market information, they are all bounded rational.

Hypothesis 3: There are no fixed costs for retailers in online direct sales channels and those in traditional offline channels, and the variable cost per unit of product is a fixed value.

Hypothesis 4: When making the price decision for the next period, the retailer of the online direct sales channel will not only consider the profit obtained in this period but also its current market share. This means that the retailer's evaluation of the operation effect of the online direct selling channel is affected by both its own profit status and its market share.

Hypothesis 5: When setting the price of the next period, the retailer of the traditional offline channel will not only consider the profit of this period but also the profit of the competitor, the retailer of the online direct sales channel.

According to the previous assumptions, the price competition among retailers conforms to the Bertrand game model. When setting the price of the next period, retailers in the online direct sales channel will consider their own market share while considering the profit situation.

After the oligopolistic competition pattern is formed, competing for a higher market share is an important strategy to drive competitors out of the competition and increase the degree of product monopoly. Therefore, increasing market share and maximizing their own profits are the important basis for online direct-selling retailers to set the next price. Online direct selling channel retailers will make a trade-off between their own profit maximization and market share when setting the next price.

The changing trend of market share is consistent with the changing trend of sales revenue. In order to simplify the analysis, this paper replaces the impact of market share on price with the impact of sales revenue on price. From the perspective of management, we can understand that the incentive for operation comes from sales profits.

The business preference coefficients not only reflect the retailer's business goals but also affect the retailer's pricing decisions. When the retailer sets prices, it no longer considers the opponent's profit or its own market share, and the system equilibrium solution is also the same as the classic Bertrand game model.

As an online direct selling retailer, in order to gain a higher market share, he will reduce the market pricing of products; at the same time, his competitive strategy will also have an impact on the prices of traditional offline retailers, and online direct selling channel retailers have an impact on the market. This competitive requirement for market share will also force traditional offline retailers to join price competition and lower the prices of traditional offline retailers.

The sensitivity of traditional offline channel retailers to their competitors, the direct sales retailers of online direct sales channels, has a much more complicated impact on the pricing strategies of both parties.

From the above analysis, we know that both online retailers and traditional retailers make the market competition more intense and complex in order to achieve their business goals.

3.6. Corresponding Dynamic Price Adjustment Model. The game between retailers is a long-term and continuous process, and it is not a game to adjust to the optimal position. The price competition between them is a complex dynamic process that is repeated for a long time. Retailers adopt a static expectation method, that is, the price decision of the next period only considers the optimal strategy of the current period, but the use of static expectations requires retailers to fully grasp the market information. In reality, the market information obtained by retailers is often incomplete because obtaining more market information obviously requires more costs. In many cases, due to information distortion and information asymmetry, it is difficult to obtain perfect information without the energy to research the market. Therefore, the use of static expectations overestimates the retailer's ability to predict market prices. At the same time, retailers' pricing processes are not all rational. Therefore, this paper assumes that both traditional offline channel retailers and online direct sales channel retailers adopt bounded rational expectations. This indicates that the retailer's price decision in the next period depends on a local estimate of the current period's marginal utility. If

the retailer's marginal utility is positive in the current period, he will increase the price in the next period. This shows that in each pricing period t, the retailer will determine the pricing of the period $t+1$ according to the pricing of its competitors in order to maximize its own utility. The dynamic adjustment process is as follows:

$$p_i(t+1) = \frac{\partial U_i(p_1, p_2)}{\partial p_i} - \alpha_i p_i^2(t+1), i = 1, 2, \tag{12}$$

where the parameter a_i is the retailer's price adjustment speed. If the retailer can obtain a greater competitive advantage by increasing the price adjustment parameter a_i, so as to achieve its own business purpose, he will continue to speed up the price adjustment. We get the discrete dynamics of the price game for a multichannel supply chain:

$$\begin{cases} p_1(t+1) = p_1(t) - \alpha_1 p_1^2(t+1)\left[\left(\frac{a_1}{b_1} - \frac{c}{p_1}\right) + \frac{\varepsilon p_2}{\beta(c+2p_2)}\right], \\ p_2(t+1) = p_2(t) - \alpha_2 p_2^2(t+1)\left[\left(\frac{a_1 \varepsilon p_2}{b_1 \beta(c+2p_2)}\right) + \frac{b_2}{\nu(c+2p_2)}\right], \\ p_1(t) = p_1(t) - \alpha_1 p_1(t)\left(\frac{a_1}{b_1} - \frac{c}{p_1}\right). \end{cases} \tag{13}$$

From the perspective of actual competition, the economic significance of the stability of the four equilibrium points of the market is given. The prices of the two retailers are both zero, which means that the two retailers have given up operating their own retail channels and withdrawn from the market, or are operating at a loss. Such a market equilibrium is obviously unstable and even impossible to happen. Both the traditional retail market and the emerging online direct selling channel market have huge potential for development, so no matter which channel a retailer withdraws from the competition, its market share will be completely occupied by other retailers.

4. Results and Analysis

4.1. Cross-Border M&A of Company A by Company B in the Financial Industry. This article will select a case of cross-border mergers and acquisitions in the financial industry for empirical analysis. Company A in China's financial industry acquired company B in Sweden in 2020. With the final closing ceremony, company A completed the acquisition of the entire equity of company B for US$1.8 billion, and obtained the ownership of company B.

The current status of China's financial industry is as follows: a lack of core technologies, low brand awareness, difficulties in expanding overseas markets, weak team R&D capabilities, and lack of mature global operation experience and mechanisms. This largely explains the motivation of Chinese financial enterprise A to embark on the road of cross-border mergers and acquisitions. Company A's finance is aimed at customers with low- and middle-income levels in China, and the brand's international influence is not enough.

The products produced by company B are mainly for middle- and high-level users. The brand is famous all over the world for luxury, comfort, and safety. Especially in terms of safety, the technical level of company B has been leading the financial industry technology of the whole world. The operation of company B has been hit hard by the financial tsunami, and its cash flow is facing a serious crisis. Affected by the huge decline in purchasing power, the sales volume is also not as good as before. The continuing operation will face the dilemma of the capital chain being cut off, and it is urgent to find a new buyer to survive.

Analyzing the financial statements of company B, it can be found that company B has been in a state of loss, but the degree of loss varies in different periods. The quarterly financial data of company B from 2017 to 2021 (20 quarters) are shown in Figure 3.

In 2018, it decreased by 20% year-on-year, and in 2019, it decreased by 11% year-on-year. It can be seen that the decline in sales volume has become smaller. This situation provides an opportunity for China's financial industry to seek overseas mergers and acquisitions. Before the acquisition, company A was in good financial condition and had a large amount of cash flow, which could fully meet the requirements of the acquisition amount. The cash flow at the end of the accounting year is increasing year by year. Company A's financial situation is running well, and it is in urgent need of external expansion and pursuing a higher level of technology. Company B meets the requirements of company A in terms of technology and brand. The cash flow of company A before the acquisition is shown in Figure 4.

According to statistics, the investment environment of the country where company B is located has been ranked high in the world, and Sweden basically has an open and friendly investment policy. In order to obtain the international investment environment rating of Sweden more accurately, according to the description of the target enterprise investment environment assessment subsystem constructed, the investment environment rating of the country where the company is located is carried out. The evaluation details are shown in Figure 5.

Company A faces many risks, the most important of which are technology integration risks, labor relations handling risks, cultural integration risks, and brand strategy risks. In order to more accurately understand the risks after mergers and acquisitions, it is necessary to evaluate the investment of the companies after mergers and acquisitions and to predict the risks through the evaluation results.

4.2. Monte Carlo Simulation of Financial Risk of Cross-Border M&A of Company A

4.2.1. The Basic Data Assumptions of the Simulation Model of Company A's Merger and Acquisition of Company B Are Determined. Investment: The initial investment is based on the acquisition amount of US$1.8 billion; in the future, some additional investment will be made for research and development. Assume that 1 year after the merger, an investment of 10,000 million yuan is required to hire new

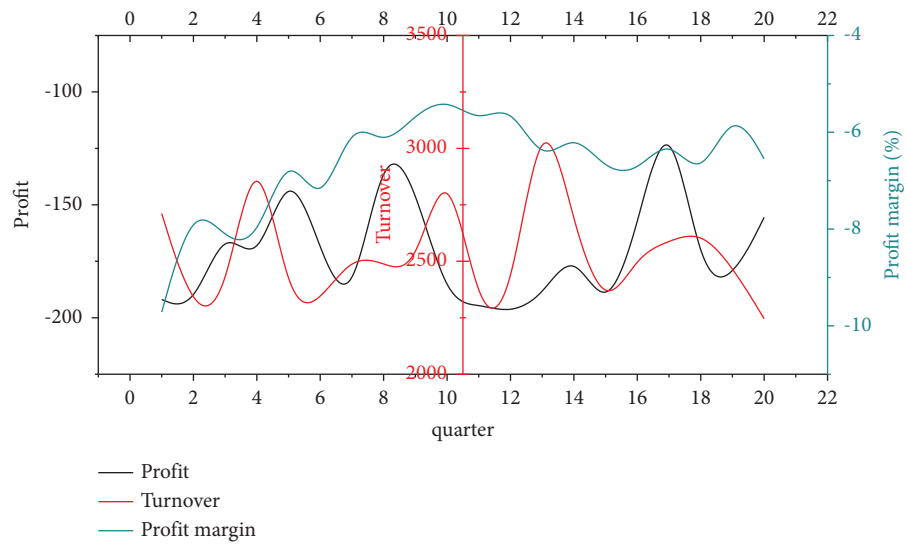


Figure 3: Quarterly financial data of company B from 2017 to 2021 (20 quarters).

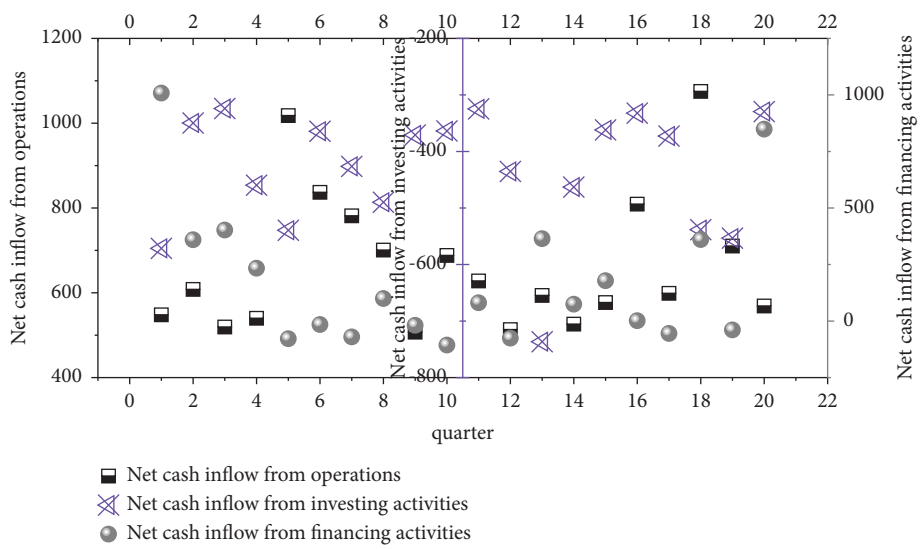


Figure 4: Company A's cash flow before M&A.

employees; 3 years after the merger, another 10,000 million yuan is invested to build a new R&D center in Europe. Assuming continuous operation, this simulation simulates the financial data for 10 years after the merger.

The turnover in 2020 of the base year is a random variable that obeys a triangular distribution. If company A can achieve the maximum production capacity in previous years at the initial stage of the merger, the maximum turnover that can be achieved is 30,900 million yuan. If there are many frictions in the initial stage of the merger, it can only meet the minimum production capacity of company A in previous years, and the minimum turnover is 12,900 million yuan. Since the research time has exceeded 2020, it can also be seen from the financial statements of company A in 2020 that the actual turnover of the year is about 20,200 million yuan, and this value is between the historical maximum and minimum values, so the possible value is set at 20200 million yuan. The company spends a lot of energy on personnel integration, corporate culture unification, standardization, etc., and its turnover changes within an average of plus or minus 10% compared with the previous year. Therefore, it is assumed that the annual growth rate of company A's turnover in the 2nd to 3rd years obeys—a random variable uniformly distributed in the interval of 10%. From the 4th year after the merger, the company's production capacity should return to historical levels, with historical 12-year annual turnover growth. According to the law of rate change to assume the range of annual growth rate of turnover of company A from 4 to 10 years after the merger, through the collected financial data of company A's turnover from 1998 to 2019, we calculated the annual growth rate of each year compared with the previous year. The turnover of company A from 2017 to 2021 (20 quarters) is shown in Figure 6.

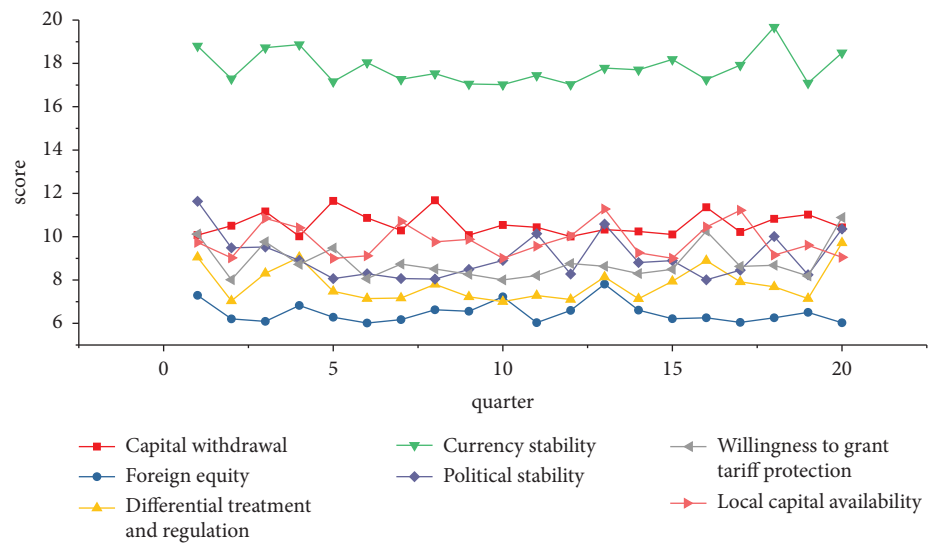


Figure 5: Investment climate rating of the country where company B is located.

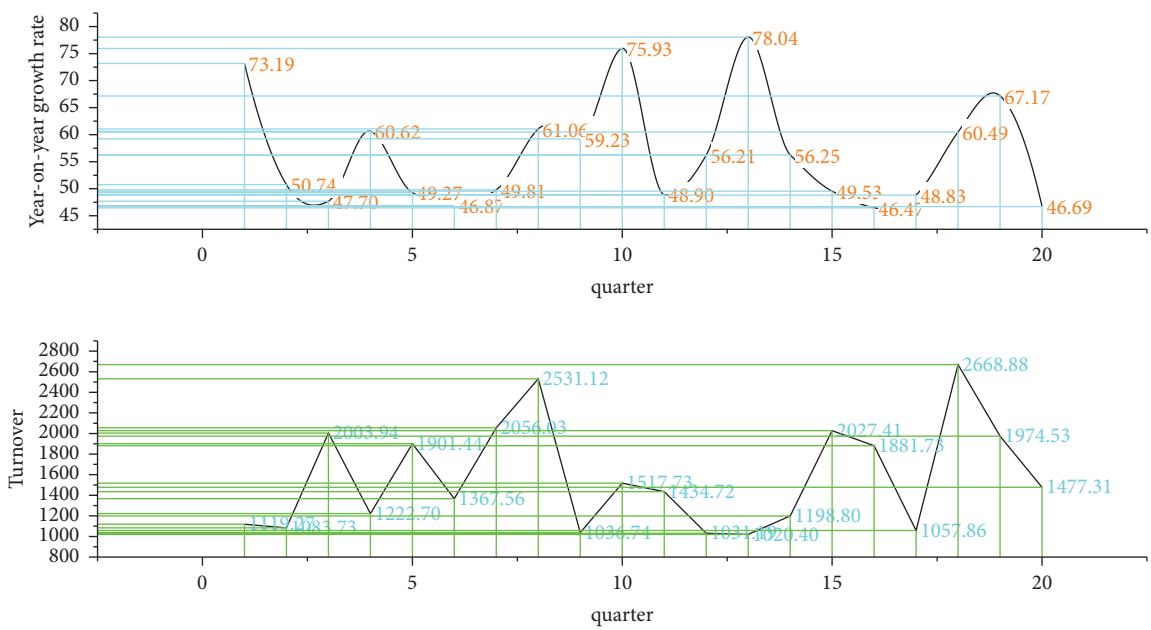


Figure 6: Company A's turnover from 2017 to 2021 (20 quarters).

4.2.2. Calculation Results of a Model Simulation. According to the principle of value selection in the basic data determined by the simulation model, run Crystal Ball software, randomly select random numbers according to the probability distribution determined above, and combine them for calculation. The cross-border investment risk prediction results of the three algorithms are shown in Figure 7.

Through this simulation, the calculation results of an experiment are obtained. The results show that the net present value of the M&A project is 150.616 billion yuan, the internal rate of return has reached 43.54%, and it has a high antirisk ability.

However, one calculation result shows the comprehensive effect of each factor. After random sampling, the repeated calculation is more convincing.

Crystal Ball software facilitates repeated sampling of values and can automatically generate reports on the statistics of predictors.

4.2.3. 1000 Simulated NPV Result Data. Theoretically speaking, the more simulation times, the more correct the result, but the corresponding manpower and material resources are needed to calculate and sort out. Relatively few simulation times will affect the reliability of the simulation

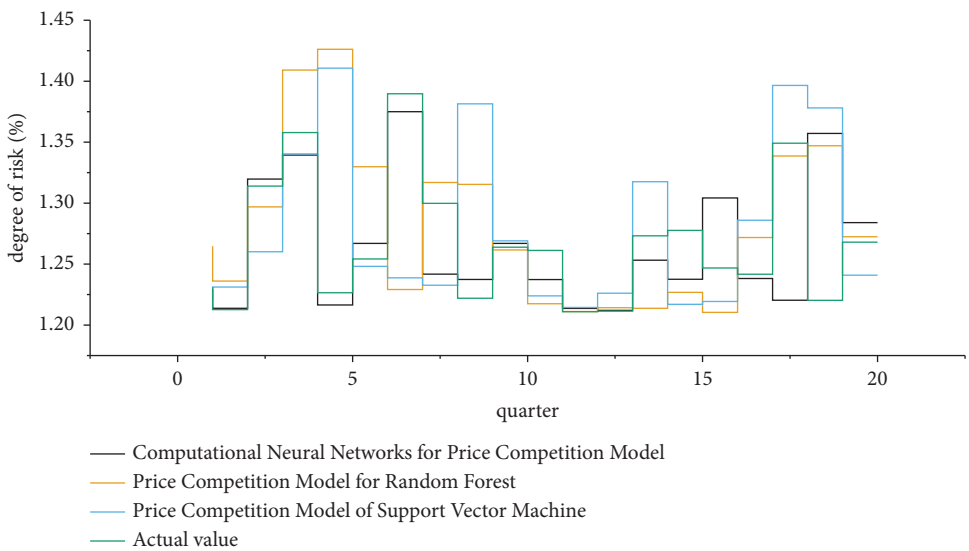


Figure 7: Cross-border investment risk prediction results of three algorithms.

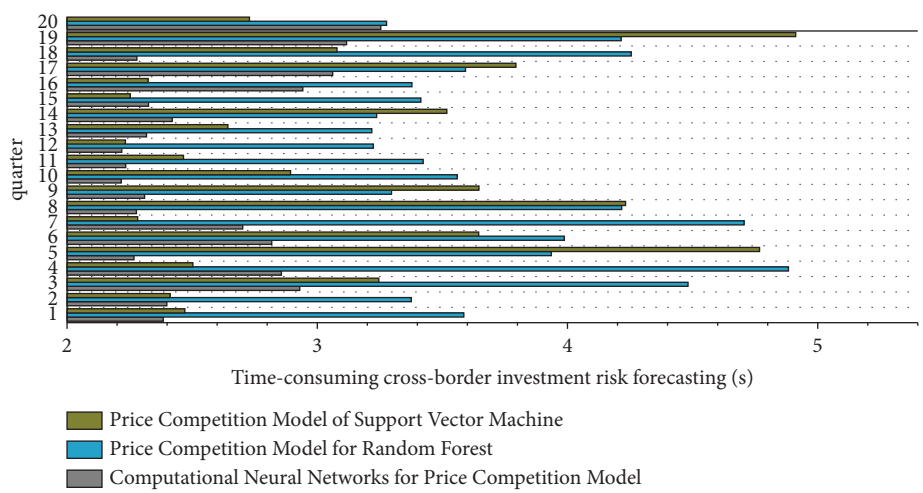


Figure 8: Time consumption of cross-border investment risk prediction by three algorithms.

results, generally between 200 and 500 times. In order to describe the change in NPV more accurately, 1000 times calculations are set this time. The operation process and value are realized with the help of Crystal Ball software. The cumulative probability that the NPV is greater than or equal to zero is $P\ (\mathrm{NPV} \geq 0) = 60\%$. Figure 8 shows the time consumption of the three algorithms for cross-border investment risk prediction.

4.3. Early Warning Analysis of a Company's Cross-Border M&A. After 1000 simulations, the output results of NPV and IRR are obtained. The cumulative probability of NPV greater than or equal to 0 is $P\ (\mathrm{NPV} \geq 0) = 60\%$, and the cumulative probability of IRR between 30% and 56.67% is 32.59%. The investment environment of the country where company B is located has a score of 90 points, and the investment environment can be rated as excellent with few restrictions on foreign investment, and company B is technology intensive and has a brand advantage. From the cumulative probability histogram of IRR, it can be judged that company A has the weak antirisk ability and is in the financial deterioration range. In this way, it can be concluded that company A has certain financial risks in the 10 years after the acquisition of company B, and the police level is set as a serious warning.

After company A successfully acquires company B, it cannot relax its vigilance. In the face of serious warnings, first of all, it must pay great attention to the integration between companies and control production costs. While growing, the cost of sales should be reduced as much as possible; secondly, to retain the advantages of the B brand and maintain its brand influence in the international market, the production technology of company A should learn from company B to narrow the technological gap. At the same time, it also launched the A product into the international market, thereby expanding its market share.

5. Conclusion

In order to solve the problem of too long intermediate transmission time caused by too large intermediate data, a "three-step compression" strategy is proposed to reduce the size of intermediate output data that need to be transmitted. The attention mechanism selects important channels and the third step data quantization operation. In order to meet the application's different preferences and specific requirements for computing delay or inference accuracy, an optimal division point selection algorithm is proposed that comprehensively considers factors such as DNN model layer type, network bandwidth, device computing power, and delay or accuracy requirements. Taking the profit of the system as the evaluation index, the influence of the chaos phenomenon on the operating efficiency of the system is evaluated. We found that the system with the highest efficiency (maximum profit) is when the competitive price of the two retailers is stable at the Nash equilibrium point, and when the system is in a cycle bifurcation or chaotic state, a very small change in price may cause a relatively small change in the market evolution behavior. Therefore, in order to avoid the loss of the system and the profits of each retailer, retailers should be more cautious and scientific when setting prices, especially when the market is chaotic due to free competition, timely analysis of information, and using certain chaos control methods are very necessary. Based on the discounted cash flow model, the postmerger alarm forecasting subsystem uses Monte Carlo simulation for each variable risk factor to examine the comprehensive impact of various risk factors on the net present value. It is implemented with the help of Crystal Ball software. The cumulative probability that the net present value of the later t years is greater than zero; the alertness interval is divided into two dimensions, the investment environment evaluation and the alertness prediction value of the subsystem, to jointly judge the alertness interval after the merger and divide the alertness into five levels: giant police, heavy police, medium police, light police, and no police. This paper selects company A in the financial industry to acquire company B in Sweden and synthesizes the investment environment rating of the country where company B is located and the cash flow of company A after the merger.

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

The work was supported by the Philosophy and Social Science Planning Project of Wenzhou (21wsk039).

References

[1] E. V. Kudryashova, L. A. Zarubina, and I. A. Sivobrova, "Cross-border investment cooperation in the Arctic Region: challenges and opportunities," *Economic and Social Changes: Facts, Trends, Forecast*, vol. 12, no. 1, pp. 39–52, 2019.

[2] B. Niu, K. Chen, L. Chen, C. Ding, and X. Yue, "Strategic waiting for disruption forecasts in cross-border E-commerce operations," *Production and Operations Management*, vol. 30, no. 9, pp. 2840–2857, 2021.

[3] M. Giuffrida, H. Jiang, and R. Mangiaracina, "Investigating the relationships between uncertainty types and risk management strategies in cross-border e-commerce logistics," *International Journal of Logistics Management*, vol. 32, no. 4, pp. 1406–1433, 2021.

[4] D. Soloveva, R. Yamini, and J. Wei, "A merry host makes bolder guests: an analysis of cross-border investment choices of Chinese firms," *Thunderbird International Business Review*, vol. 63, no. 2, pp. 175–190, 2021.

[5] K. H. Lee, "Cross-border mergers and acquisitions amid political uncertainty: a bargaining perspective," *Strategic Management Journal*, vol. 39, no. 11, pp. 2992–3005, 2018.

[6] G. Marcjasz, B. Uniejewski, and R. Weron, "On the importance of the long-term seasonal component in day-ahead electricity price forecasting with NARX neural networks," *International Journal of Forecasting*, vol. 35, no. 4, pp. 1520–1532, 2019.

[7] H. Hewamalage, C. Bergmeir, and K. Bandara, "Recurrent neural networks for time series forecasting: current status and future directions," *International Journal of Forecasting*, vol. 37, no. 1, pp. 388–427, 2021.

[8] N. V. Trusova, T. A. Cherniavska, Y. Y. Kyrylov, V. H. Hranovska, S. V. Skrypnyk, and L. V Borovik, "Ensuring security the movement of foreign direct investment: Ukraine and the EU economic relations," *Periodicals of Engineering and Natural Sciences*, vol. 9, no. 3, pp. 901–920, 2021.

[9] R. Garnier, "Concurrent neural network: a model of competition between times series," *Annals of Operations Research*, vol. 313, no. 2, pp. 945–964, 2022.

[10] H. Chung and K. S. Shin, "Genetic algorithm-optimized multi-channel convolutional neural network for stock market prediction," *Neural Computing & Applications*, vol. 32, no. 12, pp. 7897–7914, 2020.

[11] S. Avdjiev, V. Bruno, C. Koch, and H. S. Shin, "The dollar exchange rate as a global risk factor: evidence from investment," *IMF Economic Review*, vol. 67, no. 1, pp. 151–173, 2019.

[12] A. Baldominos, Y. Saez, and P. Isasi, "Evolutionary convolutional neural networks: an application to handwriting recognition," *Neurocomputing*, vol. 283, pp. 38–52, 2018.

[13] Y. Cheng, D. Wang, P. Zhou, and T. Zhang, "Model compression and acceleration for deep neural networks: the principles, progress, and challenges," *IEEE Signal Processing Magazine*, vol. 35, no. 1, pp. 126–136, 2018.

[14] L. Nalin and G. T. Yajima, "Commodities fluctuations, cross border flows and financial innovation: a stock-flow analysis," *Metroeconomica*, vol. 72, no. 3, pp. 539–579, 2021.

[15] K. M. Cherry and L. Qian, "Scaling up molecular pattern recognition with DNA-based winner-take-all neural networks," *Nature*, vol. 559, no. 7714, pp. 370–376, 2018.

[16] B. Lim, S. Zohren, and S. Roberts, "Enhancing time-series momentum strategies using deep neural networks," *The Journal of Financial Data Science*, vol. 1, no. 4, pp. 19–38, 2019.

[17] J. Li, G. Gao, L. Ma, T. Zhao, H. Qu, and F. Chen, "Analysis of profit models for cross-border power interconnection

projects," *Global Energy Interconnection*, vol. 2, no. 5, pp. 457–464, 2019.

[18] P. Ong and Z. Zainuddin, "Optimizing wavelet neural networks using modified cuckoo search for multi-step ahead chaotic time series prediction," *Applied Soft Computing*, vol. 80, pp. 374–386, 2019.

[19] M. Labouré, "Pensions: the impact of migrations and cross-border workers in a small open economy," *Journal of Pension Economics and Finance*, vol. 18, no. 2, pp. 247–270, 2019.

[20] Z. Deng, B. Wang, Y. Xu, T. Xu, C. Liu, and Z Zhu, "Multi-scale convolutional neural network with time-cognition for multi-step short-term load forecasting," *IEEE Access*, vol. 7, Article ID 88058, 2019.

Hindawi
Computational Intelligence and Neuroscience
Volume 2022, Article ID 3531011, 9 pages
https://doi.org/10.1155/2022/3531011

Research Article

A Study on the Application of BP Neural Network Based on Visual Recognition in Regional Economic Forecasting

LingYan Meng

Universiti Pendidikan Sultan Idris (UPSI), Faculty of Management and Economics, Tanjong Malim 35900, Perak Darul Ridzuan, Malaysia

Correspondence should be addressed to LingYan Meng; menglingyan@yangdong-edu.com

Received 20 July 2022; Revised 26 August 2022; Accepted 1 September 2022; Published 26 September 2022

Academic Editor: Ning Cao

The economic growth in the new normal is no longer limited to the total amount and scale of economic growth in the traditional and neoclassical periods, but has changed to "quality" and "development" under the dual requirements of historical changes and tasks of the times. The quality of regional economic growth is an important part of the quality of China's economic development and an important part of the quality of China's economic development in the new era. Therefore, this paper proposes a BP neural network based on visual recognition in a regional economic prediction model and conducts application experiments. This regional economic forecasting model is relying on data technology for economic panel data mining, then graphical processing of panel data, followed by the selection of visual recognition technology for economic panel map analysis, to derive its various component coefficients, and finally then using the BP neural network to fit the prediction, at the same time, through long-term and short-term prediction, to predict the future development quality of each region's change trends and fluctuations, to predict the institution's role, so as to avoid major transitions and deteriorating alarms, and to provide support for the macroregulation of regional economic development quality.

1. Introduction

For China's economy, 2018 is the beginning of an epoch and a new century and also the "first year of quality," the quality of economic growth will be anew historical coordinate of China's economic development [1, 2] and also an important milestone of China's economic development, and also the Chinese economy has entered a new period of "quality development" [3]. Looking back, China's economic development has been rapid since the reform and opening up, with the nominal GDP growing at 14.53% per year from 1978 to 2016, with an annual incremental growth rate of 1/3, creating a "Chinese miracle" in the past three decades and laying the foundation for China's Chinese dream. In the twenty-first century, with the restructuring of the global division of labor system, China's strategic position has become increasingly prominent, and China's economic development path is at a new stage of development [4–8].

Looking at the evolution of China's economic development, since the reform and opening up, China has entered a phase of rapid development. During the phase of quantitative growth, China is used to taking the route of "path dependence," focusing on the results of economic development, under the concept of "cost-benefit," and putting. The problem of "economic growth" is simply reduced to the problem of "conditions" and "results." Considering economic growth from the perspective of dynamic evolution, the traditional model of investment accumulation only hopes to obtain continuous income growth, but ignores the multidimensional endogenous scale-reward superposition effect, which is a nonsustainable and uneconomic behavior [9]. Moreover, the early blanket development model, which applied the same growth factor accumulation mechanism to different regions, ignored the fact that the initial resource allocation would have an increasing degree of influence on economic growth as it was gradually driven, leading to large

growth effects under the same model. Fortunately, at the end of the twentieth century, the world was thinking about the eternal topic of economic development, and instead of focusing only on the conditions and consequences of economic development, economists tried to give more connotations to economic growth from a larger perspective and proposed an extended concept of economic growth quality to explain the function of endogenous growth paths of economic growth and the accompanying social progress, coordinated development [10], environmental costs, and other deep developmental impacts are comprehensively assessed. We can see that the quality of economic growth is a new development goal, which is a new and contemporary practice of economic growth.

Under the new economic development model, the goal of development has changed from "quantity" to "quality," which is the need of the times. The report of the 19th National Congress mentions "quality" several times, summarizing China's achievements in recent years while clearly indicating that the main contradiction in China is imbalance and insufficiency. The fundamental contradiction is reflected in the direction of economic development [11]. The goal of economic development has shifted from a quantitative stage of solving the basic food and clothing problem of the population to a higher, green, coordinated, higher and higher level of development, which is an important issue for China's economic development.

Therefore, based on an empirical study on the quality of China's regional economic development, a new development model is proposed, and then the problems of the development path of the quality of China's economic growth at this stage are evaluated, and a feasible way for its future development quality is provided for the future development direction [12].

2. Introduction to Related Theories

2.1. BP Neural Network Theory. BP neural network is one of the most commonly used models, and its use is very wide, but its practical application is mostly only applicable to specific environments compared with our proposed time series model. The BP neural network is based on the intelligence of the human brain, and although it is not fully mature [13], its potential is huge, and it is foreseeable that the human brain is not just a special case in processing information. Therefore, starting from a generalized neural network model, which aims to be applied to the full range of data sequences, it is reasonable to assume that the method can perform good simulations and predictions in many aspects. The BP neural network first consists of three levels of input-output and hidden, while the hidden layer is likely to include multiple levels as well, and multiple nodes can exist on all three levels, and the propagation direction of messages is unidirectional. They are connected to each other by weights. The schematic diagram is shown in Figure 1.

The training of the BP neural network is mainly to process the information of the input layer over the excitation function and get the required information and transform it into the required information and pass it from one layer to another by the excitation function, if there is only one layer, then by the double processing of function and connection weights, that is to get our required result. Then, we decide whether to repeat the intermediate process by comparing our computed output with the real output. Obviously, we can change the excitation function and the number of implicit layers to achieve control over the above process [14].

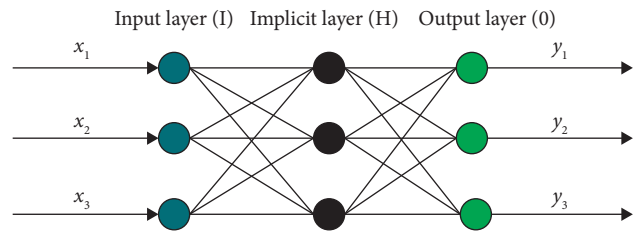


Figure 1: Principle diagram of BP neural network.

2.2. Visual Identity Theory. Visual identity (VI), also called VIS, is the abbreviation of the English Visual Identity System. Its meaning is to unify and standardize all the visualized corporate images and to communicate the company's image to the public through VI. VI includes corporate name, brand logo, standard fonts, printing fonts, standard graphics, standard colors, slogans, business reports, product descriptions, and software systems. It contains at least ten aspects, i.e., product packaging, production environment and equipment, exhibition space and equipment, means of transportation, office equipment and supplies, work clothes and accessories, advertising facilities and audio-visual products, public relations and gifts, factory flags and signs, instructional signs and signage [15]. Intuitive, communicative, and infectious are the important features of visual identification. Visual identity is to take the basic elements of corporate identity and show them effectively through strong policies and management systems, so as to form the inner visual imagery of the company and convey the spirit and business concept of the company through the unification of visual symbols, thus effectively enhancing the brand image of the company and products.

2.3. Data Mining Theory. It is the basic theory of data mining, mining algorithm, data warehouse, visualization technology, knowledge representation, knowledge maintenance and reuse, heterogeneous data mining, web data mining, etc.

The main role of data mining is to discover the types of patterns from data, to predict future trends and behaviors, and thus to provide knowledge-based forward-looking decisions. At present, the main functions of data mining are as follows:

2.3.1. Conceptual Description. It describes the connotation of a specific object and summarizes its relevant characteristics. Specifically, it includes characteristic description and differentiation description; the former refers to the

commonality of objects, and the latter refers to the difference between different categories of objects.

2.3.2. Contact Information. Data correlation is a kind of database knowledge with important significance. If two or more variables have a certain pattern, they are called correlations. Association can be classified as simple association, temporal association, and causal association. Correlation analysis aims to discover the network of relationships hidden in the database [16]. The rules generated with association analysis have a certain degree of reliability because the relevant functions of the data in the database are not known. Association analysis finds an association rule that shows the frequent simultaneous occurrence of attribute values in a particular set of data.

2.3.3. Classification and Forecasting

(1) Category. It includes categorization and determination of the category based on the properties of the object under test. It is a kind of knowledge that has the characteristics of the same nature and a kind of knowledge that distinguishes the characteristics from other things. Based on this, a classification algorithm based on decision trees is proposed. The algorithm is a decision tree structure based on a collection of examples, which is instructive. There are also statistics, rough sets, neural networks, support vector machines, etc.

(2) Prediction. Time series type information is used as the basis for predicting the attribute values or values of a particular sample, based on historical information and new information as input. Traditional statistical methods, neural networks, machine learning, etc., are time series prediction methods.

2.3.4. Cluster Analysis. Cluster analysis is the objective grouping of objects with similar properties according to their attributes. Compared with classification methods, clustering is grouping according to the properties of the data itself, and grouping is carried out without artificial grouping. By clustering or grouping them according to the principles of similarity within majorizing classes and similarity between minorizing classes, the resulting clustering results in greater similarity in the same class and lower similarity between classes.

2.4. Theories Related to Regional Economic Growth. Regional economic growth, in a narrow sense, is expressed in monetary terms, that is, the growth of GDP. In a broad sense, regional economic growth also includes controlling population and expanding demand for goods. At present, there are two different development models of regional economic growth in China.

2.4.1. Balanced Development Theory. In the book "Industrialization in Eastern and Southern European Countries," the British economist Rosensteordan put forward the "Big Push" theory. This theory is a strategic concept for developing countries to start industrialization [17]. It has four main elements. First, at a minimum critical investment scale, synergistic investment in several industrial sectors can achieve "externalities." Second, investing in complementary industrial sectors to increase industrial output can overcome the problem of narrow markets and shortage of demand in developing countries; simultaneous investment in complementary industries can reduce production costs, improve supply, and promote economic development. Third, vigorously promote domestic and foreign investment from financial sources. Fourth, capital should be invested in capital construction and light industry.

Naxos' strange circle of poverty doctrine. The shortage of capital is an important constraint to the economic growth of developing countries, which is caused by the contradiction between the supply and demand of capital. In the problem of insufficient supply of capital, low income causes low saving capacity, resulting in shortage of capital, which increases productivity and reduces income. From the perspective of capital demand, low income leads to a decrease in purchasing power, which leads to a shortage of investment, the accumulation of capital becomes difficult, and a decrease in productivity, which eventually leads to low income. To break the circle of poverty, it is necessary to increase savings, increase investment, and form reciprocal demand among industries to transform this vicious circle into a virtuous one.

2.4.2. Theory of Unbalanced Economic Development. In his book "A Brief Discussion of the Concept of Growth Poles," François Perroux puts forward a concept of growth poles based on the imbalance of economic development. The core content of the growth pole theory is that in a certain region, there is a certain innovation capacity, or an important industry, which, to a certain extent, can concentrate capital and technology, so as to achieve a certain scale effect and thus drive the economic development of the surrounding areas. It can be seen that the economic development in China is not balanced and has a polarizing effect.

The growth poles play the role of absorption and diffusion in regional economic development, which is reflected in four aspects, namely, innovation and demonstration, scale effect, external economy, and aggregation [18].

Therefore, for a region to achieve economic development, it is necessary to promote the economic development of its surrounding areas by developing a growth pole and using its polarization and radiation diffusion effects.

The American economist Hirschman, in his Economic Development Strategies, regards the imbalance strategy as the best way to achieve economic development. He argues that economic development is not achieved in all areas, and whenever it occurs in a certain region, it concentrates economic growth in the original direction of development. The formation of such growth poles inevitably leads to unevenness between regions, and in areas where regional power is concentrated, priority tends to be given to development.

This theory provides the theoretical basis for our economic development, development mechanism, and

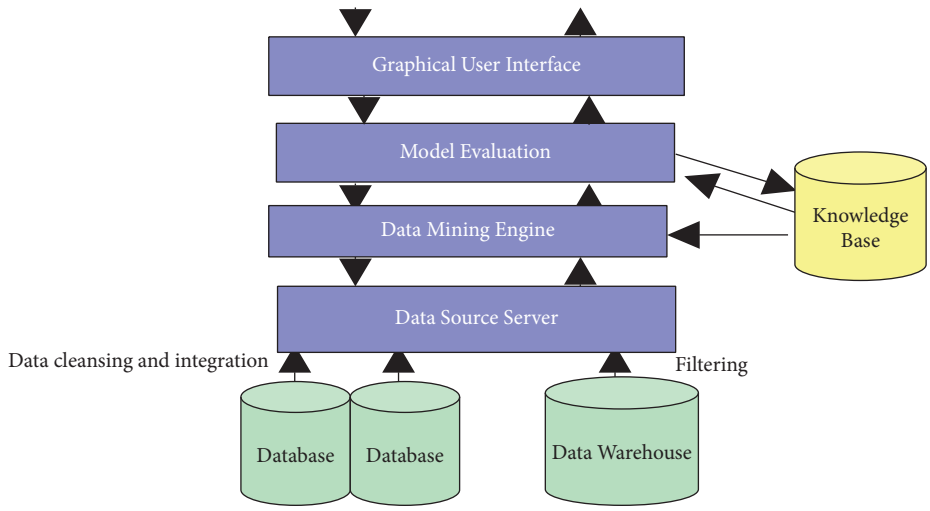


Figure 2: Data mining model.

government intervention. Based on the realistic economic growth and development mechanism, it provides new ideas for the economic development of less-developed regions. However, it ignores the lack of markets and social, economic, and infrastructure development in less developed regions, which leads to the ineffective operation of various incentive mechanisms.

3. Application Method Design

3.1. Regional Economic Data Mining Model. The research on economic forecasting from the regional perspective of China's macroeconomic forecasting mainly focuses on the construction of the regional economic forecasting system. However, before regional economic forecasting, data need to be obtained and processed first. The model diagram of this data mining system is shown in Figure 2.

3.1.1. Data Source. A data source provides data needed for data mining, including a database or set of databases, data warehouses, spreadsheets, and other types of information repositories.

3.1.2. Data Source Server. It filters, cleans, and integrates data sources to meet users' needs for data mining.

3.1.3. Data Mining Engine. A data mining engine is the core part of the system, and it includes several functional modules such as characterization, association, classification, cluster analysis, evolution, and deviation analysis.

3.1.4. Pattern Evaluation. Eliminating unnecessary or redundant patterns and filtering them based on thresholds of interest to obtain interesting patterns. GUI: an interactive interface between the system and the user that presents interesting patterns (knowledge) in various forms in an intuitive way, such as charts and tables.

3.1.5. Knowledge Base. The knowledge base is stored, and guides the process of data mining and the interpretation and evaluation of models.

3.2. Regional Economic Forecasting Model Construction. Since regional economic development is nonlinear, unstable, and unbalanced, it is difficult to ensure its effectiveness with linear methods.

From a long-term perspective, the system established should not only be in line with the comparative level of the eight balanced regions but also with the overall development trend during the new normal.

In the short term, the forecast of the quality of regional economic development should aim to meet the needs of time and individuals. First, monthly data should be studied in order to achieve the timeliness of regional economic development quality forecasting. Secondly, the purpose of short-term forecasting is different from the long-term preference trend, where the study from the individual perspective focuses on the fluctuations of individual regions and requires the establishment of a unified and uniform logical system under the main forecasting indicators [19]. Under the requirement of these two objectives, this paper proposes a coarse-set-based preprocessing method by simplifying the coarse set and synthesizing it into a comprehensive forecasting index, which can reduce the complexity of the information system and optimize the subsequent system effectively while ensuring the same forecasting performance.

This regional economic forecasting model is based on data technology for economic panel data mining, then graphical processing of the panel data, followed by the selection of visual recognition technology for the analysis of the economic panel map to derive its component coefficients, and through the synthesis of the four main aspects of the index, a BP neural network is used for forecasting. Its model diagram is shown in Figure 3.

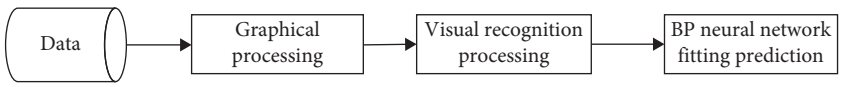


Figure 3: Diagram of the economic forecasting model.

Table 1: Fitted prediction test table.

Number	Network structure	Decidability factor			
		Training set	Validation set	Test set	Total data set
1	2	0.8276	0.8504	0.8425	0.8373
2	5	0.8125	0.7713	0.8508	0.8082
3	5–8	0.8454	0.7614	0,8617	0.8365
4	5–10	0.8321	0.8764	0.8443	0.8245
5	10–20	0.8354	0.9026	0.7714	0.8321
6	40–40	0.8754	0.8675	0.8865	0.8767
7	50–100	0.8435	0.9123	0.8454	0.8367
8	100–200	0.8504	0.8243	0.8941	0.8588
9	200–400	0.8478	0.8759	0.8132	0.8587
10	20-4020	0.9443	0.9156	0.9633	0.9456
11	50-100-50	0.8697	0.8367	0.8388	0.8553
12	20-40-40-20	0.8723	0.8505	0.8061	0.8343
13	10-20-20-10	0.8542	0.9578	0.8937	0.9032
14	10-20-20-20-10	0.8521	0.9352	0.9389	0.9405

4. Application of Experimental Analysis

4.1. Data Preparation. The experimental data in this paper are all from China Economic Network, based on the data from 2000 to 2021; the most important data are the important factors affecting regional GDP growth. Using data mining methods, from monetary policy, investment and consumption, exchange rate, total retail sales of social goods, urbanization process, science and technology innovation, etc., a total of 963,025 pieces of data were mined, and finally, by using data mining and data processing in the model, the data according to regions and forecast indicators were classified [20].

4.2. Regional Economic Fitting and Forecasting Experiment. The regional economic fitting prediction experiment is a fitting prediction model based on the visual recognition technology with eight nodes as the input layer, and the results are shown in Table 1.

It can be seen that for the selection of eight input layer nodes, the best result of 20-40-20 is 20-40-20. This indicates that the internal structure mechanism of the neural network of 20-40-20 can better explain the inner rules of this time series.

Figure 4 shows the prediction fit image of this time series with eight nodes in the input layer and 20-40-20 in the implied layer (the red line is the fitted curve, and the black line is the real curve). As we can see, the prediction results are excellent.

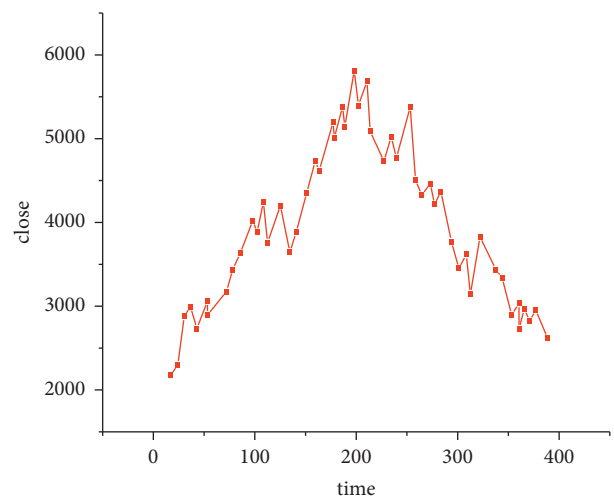


Figure 4: Fitted prediction diagram for the eight nodes of the input layer.

4.3. Short-Term Trend Forecasting of China's Regional Economy. In the short term, forecasts are made using the regional economic forecast model, and by modifying the monthly information from 2011 to 2019, the best eight regions are selected among the eight regions for optimal forecasting [21]. Output the forecast values of the quality of economic growth of major regions from January to December 2020, as shown in Figures 5–12.

Looking at the short-term forecast, there is a clear downward trend in all regions around the Chinese New Year in February this year, which is in line with the basic pattern of the economy. There is a slight increase in volatility throughout the year. According to this forecast, in addition to the traditional first-quarter volatility, the North Coast will also see a brief downturn in the second quarter, followed by an upward phase of volatility [22]. Similar to 2019, the eastern coast has 3–4 phases of large waves with a major trough in October. The southern coast is generally showing a

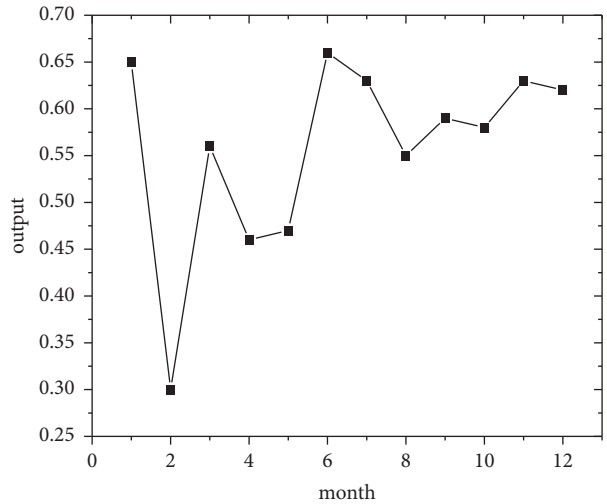


Figure 5: North Coast monthly economic growth quality forecast, 2018.

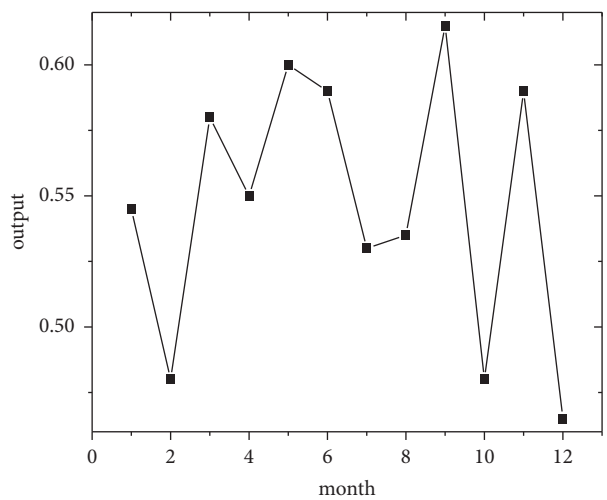


Figure 6: Eastern Coast monthly economic growth quality forecast, 2018.

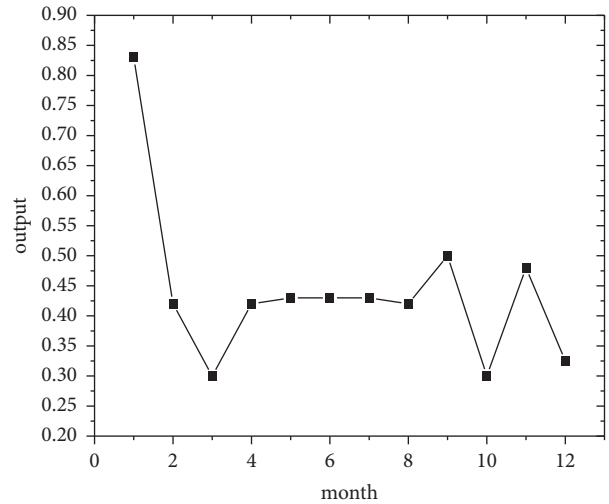


Figure 7: South Coast monthly economic growth quality forecast, 2020.

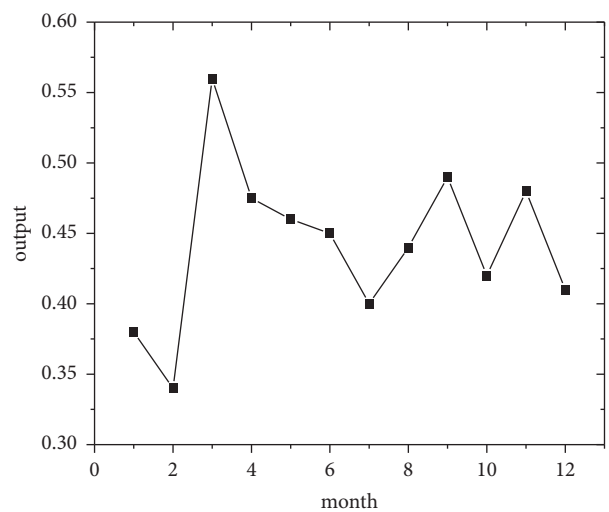


Figure 8: Northeast monitoring monthly economic growth quality forecast, 2020.

high start to 2020, with insufficient momentum for subsequent improvement, and has not yet recovered to its starting level. The Northeast started the year or was affected by last year's volatility, with a low starting point for economic growth and a sawtooth shape in the third and fourth quarters, with an unfavorable final fall. Compared with 2017, the economic development of the Central Yellow River region is noteworthy, with two serious recessions in February and March, and in the subsequent adjustment, the momentum effect is not obvious, and the fall point at the end of the year is low, which may have an impact on the beginning of next year. The Central Yangtze River region has outperformed this year, and after the volatility of the first two quarters, the third and fourth quarters are likely to be much better. The monthly trend in the Southwest region is similar to 2019, with a relatively small range in all months except the first quarter, with steady economic growth, but the overall pull-up mechanism is weak in its role and needs to be consistently improved. This year, June is the peak of the great northwest region, excluding the anomaly, there is a certain pulling effect around the middle and third quarter of each year, presumably the second and third quarter intersection for the region to enhance the quality of economic growth opportunity, and the subsequent trend has far-reaching impact. Future regulation can be advanced in this point in time, in order to make time judgments from the forecast perspective, reversing the long-term trend.

4.4. Long-Term Trend Forecast of China's Regional Economy. In the long-term forecast, we focus on the overall development trend of regional economic development, try to dig out the rate of improvement or slowdown from the global

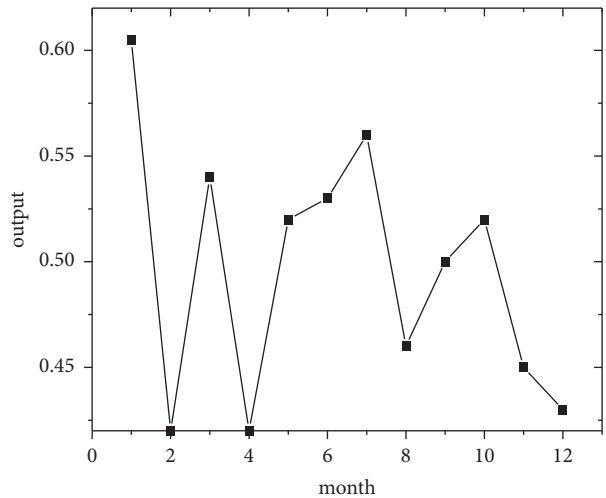


Figure 9: Monthly economic growth quality forecast for the Central Yellow River in 2020.

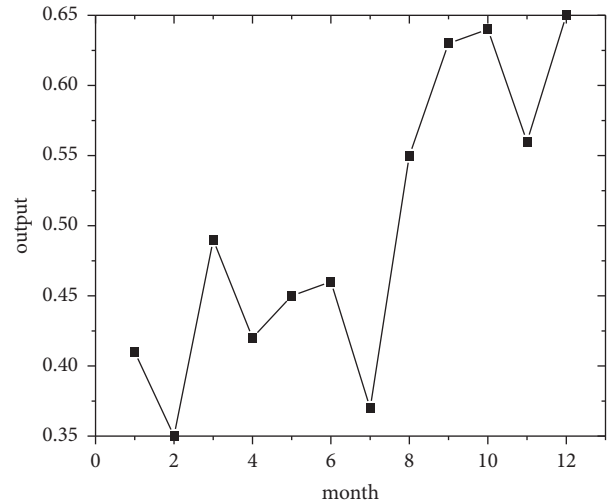


Figure 10: Central Yangtze River monthly economic growth quality forecast, 2020.

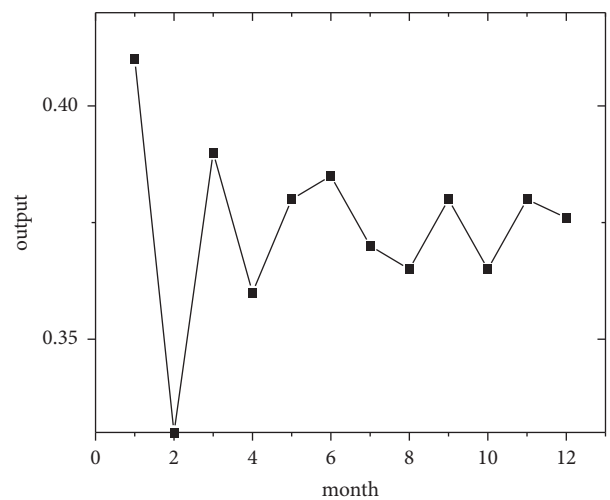


Figure 11: Southwest monthly economic growth quality forecast, 2020.

perspective, and analyze in detail the differentiation status among individual regions and the possibility of achieving regional catch-up[23]. Meanwhile, using the established regional economic forecasting model, the regional economic development quality forecasting indexes from 2000 to 2021 are divided into five stages for cyclic learning and correction within five years, and finally, the implied level is set as a single level with the number of neurons as 5, so as to obtain the development trend of five years.

The black part in Figure 13 shows the real quality in the past, and the gray one is the expectation for the future. From the time dimension, we can see that the quality of economic development in the eight regions will continue to improve in the longer term. According to the annual historical analysis of the three ladder pairs, the North and East coasts remain at the top of the rankings, and since 2017, the quality of economic growth in these two regions will level off, while the East Coasts will continue to take advantage of efficiency gains and have better overall quality than the North Coasts. There will be changes in the configuration of the second tier. In the past 17 years of quality forecasts, the southwest region has continued to maintain its growth potential and will surpass the middle reaches of the Yangtze River in the future, officially ranking among the second tier. The original South Coast and the two midstream regions will see little improvement in the quality of economic growth over the next five years, proving that there is spatial overlap between the three regions mentioned above. In the early stage of economic development, the south coast police situation is dominant and weak in terms of operational processes, results, and benefits, and its priority is the efficiency-driven quality improvement path, and its future quality status is still ahead of the two midstream regions, but it is also necessary to pay attention to the insufficient growth base caused by the long-term revaluation state of the quantity and quality of input factors, and the level of human capital and physical capital supply after the efficiency improvement of industrial upgrading The match with the level of human capital and physical capital supply after the efficiency of industrial upgrading, with a view to continue to maintain consistency with the level of quality of the first tier. In the two mid-tier regions, the tendency is to adopt the path of improving quality in the investment phase, which is effective but does not generate endogenous economic growth momentum in the long run, and when conditions accumulate to a certain extent, a distortion, slow and digestive capacity distortion, will be generated, and it will take more time and effort to improve production efficiency by then. The quality of the third echelon in the Northeast is still slow to improve, and the problem of transformation and upgrading will continue to be a structural resistance; in the short term, the success of the Great Northwest to solve the problem of locking the lower end of the quality level also depends on the improvement measures in many aspects of development

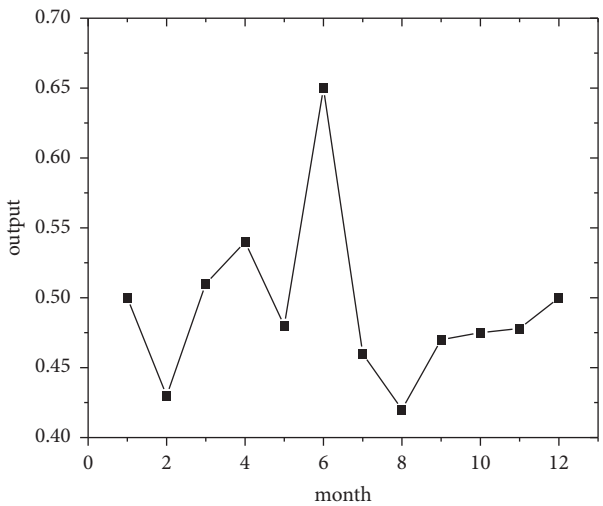


Figure 12: Great Northwest monthly economic growth quality forecast, 2020.

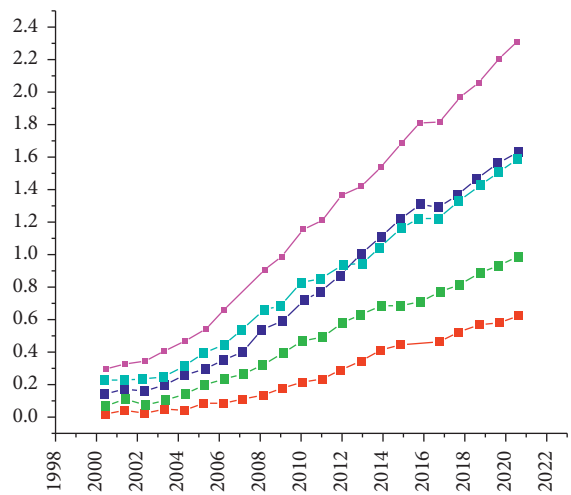


Figure 13: Long-term trend projections for the quality of economic growth in the eight regions.

conditions, processes, results, and benefits, and at the level of strategic adjustment.

Overall, the first tier of regional quality remains stable and continues to lead over the next five years. The second tier, with the addition of Southwest and the stability of the three regions, is expanding significantly, but is encountering many obstacles to break through to the first tier. The third level continues to improve but lacks effective incentives and innovative ways to achieve great growth, and the superior engine for achieving great growth has not yet taken shape, and the path to improvement remains elusive. Based on the forecast and analysis of the history and current situation of economic development, how to break the status quo, stabilize the advantages, tap new effective growth engines, and achieve a dynamic balance of high quality is the key to quality regulation in the future.

4.5. Policy Shift of China's Regional Economic Growth Quality Improvement. The transformation of regional economy from quantitative growth to qualitative growth requires mutual coordination between national macrocontrol and regional macrocontrol. At the level of conditions, we should actively explore the new structural characteristics of factor endowment, comply with the background of the new normal economic development law, and improve the quality of supply; take health-oriented financial capital, and guide the flow of capital to the real and emerging green industries by reducing the financial leverage of each region; and take a flexible approach directing human capital flows to high-tech and innovative industries. At the process level, reforms are used to further develop market dynamics, tap potential growth drivers in the transition period, and effectively innovate the social security system to maintain a smooth, fair, stable, and green operating mechanism and an environment for efficiency. To measure the value judgment of regional economy and development from the perspective of evolutionary results and radiating benefits, to bring into play regional synergy under the logic of development and quality, and to achieve dynamic balance in the per capita sense under the guidance of national macropolicies, complemented by differentiated macrocontrol policies with long and short-term interaction matching the regional quality status.

To achieve all-round quality control in the whole process of economic development, the system of rapid development such as total control, individual regulation, speed regulation, and short-term regulation must be abandoned. On the basis of the new normal and the characteristics of the new era, according to the characteristics of the eight regions, the focus of regional development is on a unified logic, long-term and short-term interactive regulation, the regulation of coordinated development, and effective low-end supply improvement. We should put the regional characteristics of unified and differentiated regulation ideas throughout each regulation and steering measures, accelerate the establishment of regional economic, scientific and technological innovation, and financial and human resources coordinated the development of the regional industrial system, and build a regional economic market mechanism effective. Regional micro subjects have vitality, regional macro regulation, and control of the economic system, and constantly enhance the innovation and competitiveness of China's regional economy, to achieve long-term sustainable regional economic growth. The regional economic innovation and competitiveness of China's regional economy will be continuously enhanced, and long-term sustainable regional economic growth and all-round quality improvement will be achieved.

5. Conclusion

The transformation of regional economy from quantitative growth to qualitative growth requires mutual coordination between national macrocontrol and regional macrocontrol. At the condition level, we should actively explore the new structural characteristics of factor endowment, comply with the background of the new normal economic development law, and improve the quality of supply; health-oriented

financial capital by reducing the financial leverage of each region, guide the flow of capital to the real and emerging green industries; and guide the flow of human capital to high-tech and innovative industries in a flexible manner. Economic forecasting is the application of the theory and method of forecasting to economic activities and the study of future economic phenomena. This paper uses theories and methods from several disciplines, including economics, probability theory, mathematical statistics, modern management science, system theory, and computer science, and applies appropriate mathematical modeling techniques to analyze and forecast the development trend of the research object.

This paper starts from the basic theoretical framework of macroeconomic growth quality, summarizes the current research basis and shortcomings from the empirical point of view, reconstructs the research idea of regional economic growth quality, therefore proposes the BP neural network in the regional economic forecasting model in visual recognition, and uses data mining technology for data preparation and the regional economic data set for model fitting. The model is then fitted to the regional economic dataset for prediction experiments and long-term and short-term economic prediction experiments, and finally, it is concluded that the model is well applied and has good feedback for long-term and short-term regional economic prediction.

Data Availability

The dataset used in this paper are available from the corresponding author upon request.

Conflicts of Interest

The authors declared that they have no conflicts of interest regarding this work.

References

[1] K. Chang and B. Tsai, "The effect of DEM resolution on slope and aspect mapping," *Cartography and Geographic Information Systems*, vol. 18, pp. P69–P77, 1991.

[2] C.-Z. He, *Self-organizing Data Mining and Economic Forecasting*, Science Press, Beijing, China, 2005.

[3] G. Zhao, *Data Mining*, Xi'an Jiaotong University, Xian, China, 2003.

[4] Z. Jing and Xiaogang, *Data Mining Algorithms and Their Engineering Applications*, Mechanical Industry Press, New york, USA, 2006.

[5] X. Huifeng and K. Xiaodong, "Wenyu," *Intelligent Data Mining Techniques*, Northwestern Polytechnic University Press, Chennai, 2005.

[6] Li Qiudan, "Research and platform implementation of data mining related algorithms," *PhD Thesis*, Dalian University of Technology, China, 2004.

[7] J. W. Han and M. Kamber, "Data mining: concepts and techniques. San francisco," *CA. Morgan Kaufman*, pp. 2–4, 2001.

[8] D. Hand, H. Mannila, and P. Smyth, *Principle of Data Mining*, pp. 1-2, MIT Press, Cambridge, CA, 2001.

[9] I. Witten, *H Frank E. Data Mining Machine Learning Tools and Techniques with JAVA*, Implementations, San Francisco, CA. Morgan Kaufman, 2000.

[10] L. Minxia, "Technical methods and applications of data mining and knowledge discovery," *Journal of Yuncheng College*, vol. 2, 2005.

[11] U. Fayyad, F. Piatetsky-Shapiro, and P. Smyth, "From data mining to knowledge discovery in databases," *AI MAGAZINE, fall*, vol. 52, no. 8, pp. 28-29, 1996.

[12] S. C. An, *Data Warehousing and Data Mining*, Tsinghua University Press, Beijing, China, 2005.

[13] J. Zhiquan, *Research on Self-Organizing Data Mining Model Based on GMDH Principle*, Dalian Maritime University, China, 2004.

[14] C.-L. Chu and En-T. Ma, "Research on local government debt risk prediction in China–analysis based on BP neural network," *Journal of Guangxi College of Finance and Economics*, vol. 29, no. 05, pp. 58–67, 2016.

[15] C. Fangli and Z. Xinhong, "Some understanding on macroeconomic monitoring and forecasting system [J]," *Special Economic Zone*, vol. 10, pp. 33-34, 1994.

[16] L. Hao, H. U. Xinyue, and L. I. Li, "Research on the artificial neural network model for the credit risk analysis of commercial banks. Systems engineering-theory practice," vol. 25, no. 9, pp. 55-56, 2001.

[17] Y. Ichikawa and T. Sawa, "Neural network application for direct feedback controllers," *IEEE Transactions on Neural Networks*, vol. 3, no. 3, pp. 224–231, 1992.

[18] S. E. Yang and L. Huang, "Financial crisis warning model based on BP neural network," *Systems Engineering-theory Practice*, vol. 25, no. 12, pp. 12–19, 2005.

[19] T. D. Sanger, "Optimal unsupervised learning in a single-layer linear feedforward neural network," *Massachusetts Institute of Technology USA*, vol. 2, pp. 459–473, 1989.

[20] J. H. Park, Y. S. Kim, I. K. Eom, and K. Y. Lee, "Economic load dispatch for piecewise quadratic cost function using Hopfield neural network," *IEEE Transactions on Power Systems*, vol. 8, no. 3, pp. 1030–1038, 1993.

[21] Y. Zhang, Y. Xu, and G. Ding, "License plate character recognition algorithm based on filled function method training BP neural network," *Computer Engineering Science*, vol. 31, no. 5, pp. 59–61, 2009.

[22] J. Hu and X. Zeng, "An efficient activation function for BP neural network," *International Workshop on Intelligent Systems*, vol. 36, no. 5, pp. 1–4, 2009.

[23] S. R. Gunn, "Support vector machines for classification and regression," *ISIS Technical Report*, vol. 14, pp. 85-86, 1998.

Hindawi
Computational Intelligence and Neuroscience
Volume 2022, Article ID 8045968, 12 pages
https://doi.org/10.1155/2022/8045968

Research Article

Security Analysis of Social Network Topic Mining Using Big Data and Optimized Deep Convolutional Neural Network

Kunzhi Tang,[1] Chengang Zeng,[2] Yuxi Fu,[3] and Gang Zhu[4]

[1]*College of Engineering & Computer Science, The Australian National University, Canberra 2615, Australia*
[2]*School of International Education, Zhejiang Normal University, Jinhua 321004, China*
[3]*Department of Science and Technology, Beijing Normal University-Hong Kong Baptist University United International College, Zhuhai 519087, China*
[4]*Chinese Academy of International Trade and Economic Cooperation, Beijing 100000, China*

Correspondence should be addressed to Chengang Zeng; zengchengang@zjnu.edu.cn

Received 11 July 2022; Accepted 6 September 2022; Published 23 September 2022

Academic Editor: Ning Cao

This research aims to conduct topic mining and data analysis of social network security using social network big data. At present, the main problem is that users' behavior on social networks may reveal their private data. The main contribution lies in the establishment of a network security topic detection model combining Convolutional Neural Network (CNN) and social network big data technology. Deep Convolution Neural Network (DCNN) is utilized to complete the analysis and search of social network security issues. The Long Short-Term Memory (LSTM) algorithm is used for the extraction of Weibo topic information in the memory wisdom. Experimental results show that the recognition accuracy of the constructed model can reach 96.17% after 120 iterations, which is at least 5.4% higher than other models. Additionally, the accuracy, recall, and $F1$ value of the intrusion detection model are 88.57%, 75.22%, and 72.05%, respectively. Compared with other algorithms, the model's accuracy, recall, and $F1$ value are at least 3.1% higher than other models. In addition, the training time and testing time of the improved DCNN network security detection model are stabilized at 65.86 s and 27.90 s, respectively. The prediction time of the improved DCNN network security detection model is significantly shortened compared with that of the models proposed by other scholars. The experimental conclusion is that the improved DCNN has the characteristics of lower delay under deep learning. The model shows good performance for network data security transmission.

1. Introduction

In recent years, the Internet technology has made unprecedented progress with the continuous acceleration of the globalization process [1]. The number of smart grid terminals has increased, big data, cloud computing, and other technologies have been applied in various fields, and the amount of data on the Internet has been exploded [2]. The Internet has gradually entered the era of big data with the new features of high value and high transmission speed. Data provides more support for people to find and extract useful information. The cross-age progress of the Internet has made social networks larger and more diverse in forms. Various social networks, such as QQ, WeChat, and Weibo, have completely changed the life [3].

However, the development of science and technology is a two-edged weapon. The rapid technological progress brings not only a new way of life but also the subsequent network information security issues [4]. In social networks, users are often threatened by network security issues including network attacks, private data leakage, misuse, and theft of confidential information. In the current network information security monitoring system, there are many methods and strategies for managing network information security [5]. However, all these methods have varying degrees of deficiencies in terms of security and operability when facing

a huge amount of data. They cannot detect security topic information on the network in time, especially the effective information contained in the massive data of social networks. Traditional information search technology is difficult to search and extract in time [6]. Therefore, to ensure data and information security, timely and effective social network security topic mining and analysis models are particularly important for the development of information security and network security [7].

To sum up, in today's diverse social networks, it is particularly important to mine and analyze the security topic data on social networks and meanwhile ensure the security of users' social network information and avoid the leakage and data loss of users' private data [8]. At present, the main problem is that the topics discussed on social networks are mixed, and the inadvertent behavior of users on social networks may leak their private data and threaten their privacy. The main contribution of the research lies in the establishment of a network security topic detection model by combining deep convolutional neural network (DCNN) and social network big data technology. The main innovation lies in the use of DCNN to complete the analysis and search of social network security issues. The Long Short-Term Memory (LSTM) is used to extract Weibo topic information in the memory wisdom algorithm. This is the uniqueness and novelty of research and provides an experimental reference for the subsequent improvement of the topic security and data transmission performance of social networks.

2. Recent Related Work

2.1. Research Status of DCNNs. In recent years, the classification accuracy of Deep Neural Network (DNN) and convolutional neural network (CNN) has begun to significantly improve with the increasing network structure and the comprehensive improvement of hardware performance [9], as the neural network model of the multilayer network structure. The neural network model has been widely used in many fields, such as system intrusion detection and images remote sensing recognition [10]. Rm et al. used DNN to study benchmark intrusion detection in the medical IoT environment, and the results showed that the performance of the DNN model was more accurate than the original machine learning algorithm, and the time complexity was reduced by 32% [11]. Devan and Khare used the softmax classifier to classify network intrusions under the Tensor flow framework through DNNs, and the results showed that the model was better than the shallow machine learning algorithm of the original data set [12]. Chen et al. used an improved CNN architecture to evaluate and calibrate the pollution degree of agricultural irrigation water resources and improve the accuracy of near-infrared prediction [13]. Igarashi et al. used DCNN (AlexNet) to predict and study the degree of cancer invasion on endoscopic images of the upper gastrointestinal tract. The accuracy rates of the trained and validated data sets are, respectively, 99.3% and 96.5% [14]. Kattenborn et al. used deep learning CNNs to extract vegetation attributes from remote sensing images. Those improve the spatial resolution of vegetation remote sensing data [15]. Awais et al. studied the classification of mixed and awake states of newborns based on DCNN and support vector machine algorithms. Its accuracy rate reaches 93.8% [16]. In summary, CNN and DCNN have been widely used in computer vision in various fields.

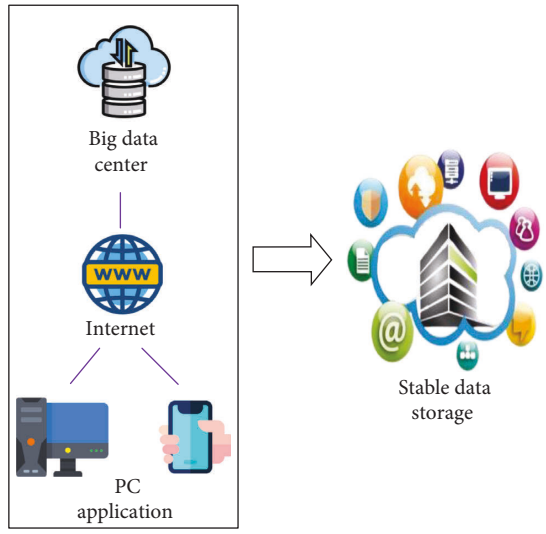


Figure 1: Mining topic data on social networks.

2.2. Status of Social Network Security Issues. With the rapid development of social networks, the personal privacy leakage of users has become more and more serious. The private information of users in social networks is illegally collected and used. Many scholars have conducted related researches to solve the security protection problems of social networks. Alshaikh et al. proposed a privacy measurement algorithm on the degree of adjacency to achieve the level of data availability and privacy protection "trade-off" based on the differential privacy model, combined with clustering and randomization algorithms [17]. Zhan and Tao studied the security protection technology and security protection mechanism under the 5th Generation Mobile Communication Technology (5G) network [18]. Maragatham et al. found that the trust-based photo-sharing component can help solve the security problems in the photo-sharing process of social platforms through simulation experiments [19]. Meanwhile, the anonymization of pictures should be represented with the edge characterized by the distributor.

The security issues of social networks are still at a basic stage by analyzing the research of the abovementioned scholars. Few scholars apply DCNNs to the topic mining and analysis of social network security. Therefore, the use of DCNN to construct social network information security and intrusion detection models has extremely important practical guiding significance for the analysis and prediction of network information security under the trend of increasing information security risks of social networks [20].

3. Weibo Security Topic Mining and Security Analysis on DCNN

3.1. Topic Mining and Demand Analysis of Social Network Security. Weibo is one of the social networking platforms that people use the most in their daily lives. The information on Weibo can not only describe people's daily life, but also the opinions and experiences of different groups of people on a certain item. The topic of Weibo has become an important basis for reflecting popular social topics [21]. This research can identify and monitor important topics that affect network security from Weibo. The most important thing is to collect all Weibo information on the topic in time after the incident. To collect topic information comprehensively, it is necessary to transform the messages sent by users into a topic model. After the feature expression of Weibo is obtained in the experiment, the DCNN model is used to train and learn the deep semantic information in Weibo, and the system will accurately identify and detect the topic feature expression of Weibo [22]. Figure 1 shows the mining of topic data on social networks. The social network data extracted from the user's mobile phone client and computer terminal are filtered by the Internet, stored in a big data center, and extracted and analyzed through the data center when needed.

3.2. Weibo Search Ranking and Topic Feature Mining and Extraction. Weibo search ranking is playing an increasingly important role in Weibo exploration and analysis [23]. As a hot research topic, Weibo search is also widely used by academic backgrounds and industries. The main content of the Weibo search is divided into two parts: extraction of Weibo features and research on Weibo topics [24]. The extraction of Weibo features refers to the analysis of the characteristics of Weibo social relations and the temporal and spatial characteristics of Weibo on the micro-meaning of Weibo text information, unifying the content of Weibo blogs, eliminating the semantic gap between media, and performing images and videos and unified extraction of multimedia information. Weibo topic research is to strengthen user search terms by inputting the user's search intent, combining semantic expansion and the expansion of the knowledge base, and implementing a classification model on Weibo features.

In the Weibo search, the most important thing is the related classification, which combines all the features extracted above, and obtains the similarity between the Weibo message to be matched and the user request according to the classification template and returns the search result [25]. The quality of the classification template will affect the user's search experience. The classification template directly determines the order in the search results to study the network, and users tend to pay special attention to the first tweet in the search results. Therefore, the question about the Weibo information at the top of the search results can satisfy the user's needs in terms of information, and the search directly determines the user's search experience. The classification model has played an important role in the Weibo research framework for many years.

Weibo information classification process: extract Weibo features from Weibo information data, select learning methods, learn classification models, and form the weight of each feature [26]. The linear combination used to generate the classification function can be expressed as follows:

$$f(x_i) = \sum_{j=1}^{|F|} w_j * t_j, \tag{1}$$

x_i represents the ith feature of t_j, w_j represents the weight value of the jth feature in the ranking function, and F is the feature set. On the ranking model, the Weibo topic information related to the query sentence will be returned to the search result. The overall feature extraction structure is shown in Figure 2.

3.3. Weibo Security Topic Detection on Hidden Dirichlet Distribution. Because Dirichlet distribution has good conjugation and aggregation, it is widely used in the field of deep learning [27]. The Dirichlet distribution is used to mine and analyze the probability distribution of document topics by using common appearance features in the text [28]. Since Dirichlet distribution is an unsupervised learning algorithm, it can easily realize vectorized representation and modeling of text. Here, the trained model is used to extract the topic detection features of the text, the Weibo detection message is used as the topic distribution vector, and the Gibbs sampling algorithm is used to sample the distribution of lexical items. The Gibbs sampling is defined as follows:

$$P\left(z_i = k \mid z_i, w\right)\theta_{m,k} * \varphi_k, \tag{2}$$

$$\varphi_{ki} = \frac{n_{k,i} + \beta_t}{\sum_{i=1}^{n}\left(n_{k,i} + \beta \mathbf{t}\right)}. \tag{3}$$

Equation (2) is the process of document generation, which means that when the mth document is generated, the topic number of the nth word of the document is first generated from the document distribution matrix. Formula (3) is the process of document word generation, which represents the process of generating the nth word of the mth document in the thesaurus.

The Dirichlet distribution is the foundation of digital image processing. In the signal analysis and processing link, Fourier Transform is used to transfer the signal from the time domain to the frequency domain. The frequency domain is used to process the signal to more accurately analyze the composition of the signal frequency, laying the foundation for the filtering operation. It is transferred from the spatial domain to the frequency domain after the Fourier transform of the image. Therefore, the image is subjected to Fourier Transform operation. This is very important in the field of digital images. It is the basis of image feature

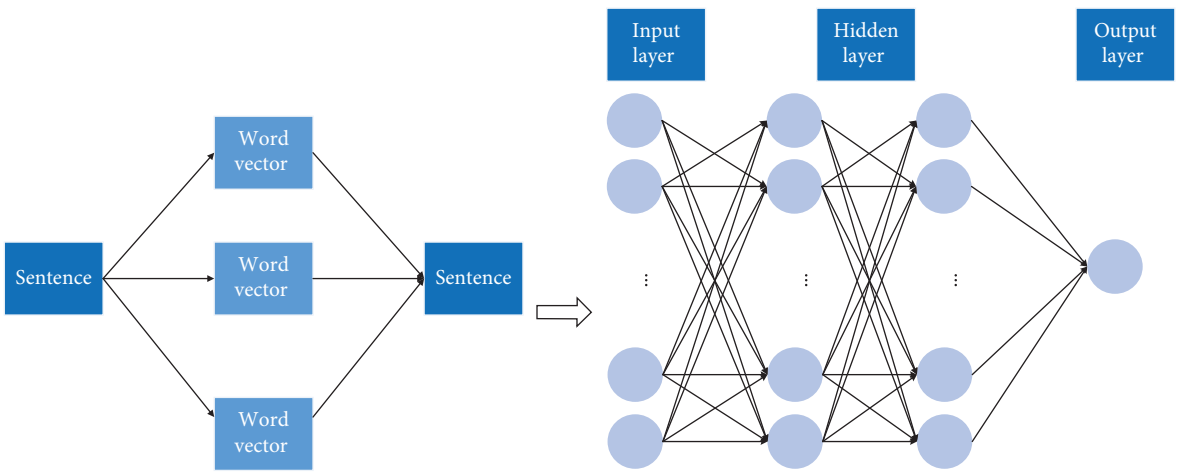


Figure 2: Weibo topic feature extraction structure diagram.

extraction and the basic necessary condition for image edge detection and filtering and noise reduction. The DNN Weibo topic detection algorithm and DCNN's Weibo security topic detection are used to verify the effectiveness of the security topic mining model.

3.4. Weibo Topic Detection on DNN. The neural network-based Weibo topic detection method requires the detection object as a classification task [29]. Different from the previous vector machine model supported by the vocabulary unit as the Weibo text vector, the neural network method will extract the latent semantic features of the Weibo text quantization process, and then use the Weibo text vector to form multiple classifiers to supplement the Weibo topic detection Model formation. The research quantifies the text through neural detection of Weibo topics, which has strong linear separability. The framework of the method for detecting Weibo topic objects on neural networks is shown in Figure 3. In the preprocessing part, the text of Weibo is mainly processed for word classification, word banning, and high-frequency word deletion.

3.5. Weibo Security Topic Detection Model on DCNN

3.5.1. Long Short-Term Memory (LSTM) Network. A recurrent neural network (RNN) uses short-term information during the training process or makes the model lose the long-term training results. To solve this problem, the research uses the LSTM network structure [30]. In the LSTM network, the threshold structure is introduced into the "forget" and "memory" functions. During training, the threshold structure selectively transmits information to the network so that implicit long-term information can be transmitted during the formation. The vector words formed by the LSTM network can better express the depth of semantic information implicit in the semantic information in the text and obtain more representative vector words of text features, as shown in the following equations.

$$i_t = \left(W_{x,i}x_t + W_{z,i}Z_{t-1} + b_i\right), \tag{4}$$

$$f_t = \left(W_{x,f^x t} + W_{z,f^z t-1} + b_f\right), \tag{5}$$

$$o_t = \left(W_{x,o^x t} + W_{z,o^z t-1} + b_o\right), \tag{6}$$

$$g_t = \tanh\left(W_{x,m^x t} + W_{z,m^z t-1} + b_g\right), \tag{7}$$

$$m_t = f_t \circ m_{t-1} + i_t \circ g_t, \tag{8}$$

$$z_t = o_t \circ \tanh m_t, \tag{9}$$

m_t represents the value m_{t-1} of the last iteration and the unit function value of the input x_t. Pass the vector o_t to all connected layers, and finally, calculate the output value with z_t. Meanwhile, m_t indicates the iteration vector representing the number of model training times, which is passed to the next training as part of the input vector for the next iteration. W is the weight matrix. b_* is the bias vector, which is the training parameter of the neural network. $u \circ v$ represents the dot product operation between the two.

3.5.2. Use the Long and Short Temporal Memory Network to Obtain the Deep Semantic Features of the Weibo Text. On the LSTM sequence memory unit, the construction of the neural network mechanism is shown in Figure 4.

To limit the transmission of the parameters of the neuron network, the maximum number of expansion steps is set as 32, which means that the neuron network will restart after the network is expanded 32 steps forward. A series of Weibo topic data is used as the input of RNN, which predicts the next word and updates the neural network settings according to the prediction accuracy. The LSTM state vector of each iteration step of the neural network is used as the corresponding word vector input to form a neural network until the network converges to the specified maximum number of iterations.

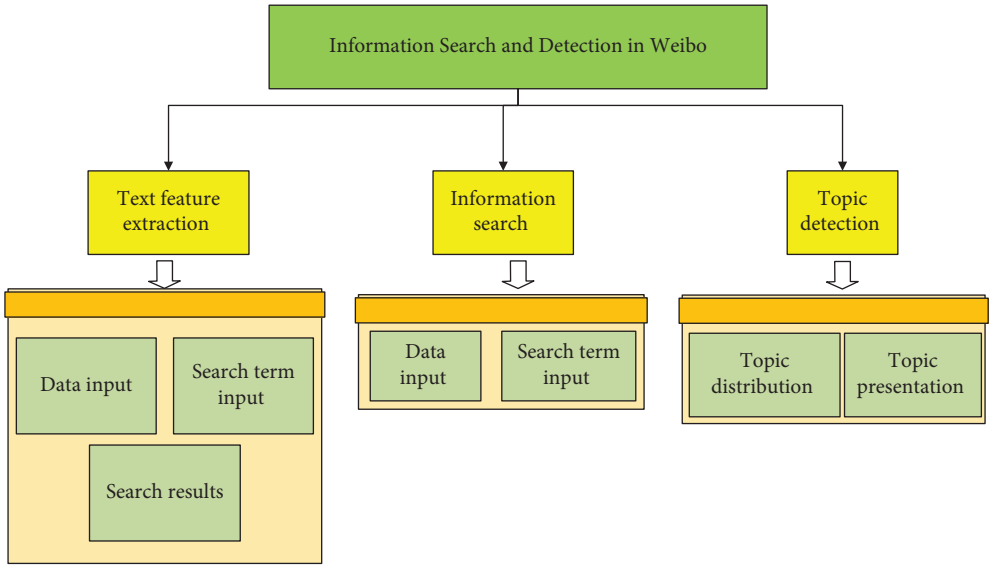


Figure 3: Framework diagram of Weibo security topic mining and detection.

Meanwhile, each Weibo has a start character and an end character. The LSTM state vector corresponding to the formation step of the neural network of the ending character can be expressed as a Weibo vector. After using text corpus data to form a RNN, the proposed method can obtain the Weibo word vector that represents the feature of the Weibo text. This vector representing the text features of Weibo can be used to represent the words in Weibo and apply them to subsequent Weibo search tasks and Weibo topic detection.

3.5.3. Vectorized Representation of Weibo Text. The first step in implementing a Weibo topic object detection model on a DCNN is to quantify Weibo text. The quantification method of Weibo text directly affects the detection effect of the neural network [31]. In the text extraction model on CNN, the common method of text direction quantification is to supplement and crop the text after preprocessing and concatenate the word vector corresponding to each word in the text to obtain each word in the assembled text to obtain the direct quantification of the text [32]. For the Weibo object detection model on convolution depth, when preprocessing the Weibo text, the text less than the threshold length is filled with zero value, and the text greater than the threshold length is cropped. And then, the word vector of each word is spliced, and the obtained Weibo text information is vectorized, and the vectorized representation process of the entire Weibo text is shown in Figure 5.

The training corpus in Figure 5 is the Weibo text corpus and real-time news corpus. Text preprocessing includes operations such as word segmentation, word extraction, and removal of low-frequency words from the text information. Filling and cropping: zero-value padding for the length less than and cropping for the part whose length are greater than the text length threshold. Then, the trained word vector is used to assign the processed text to obtain the vectorized representation of the Weibo text.

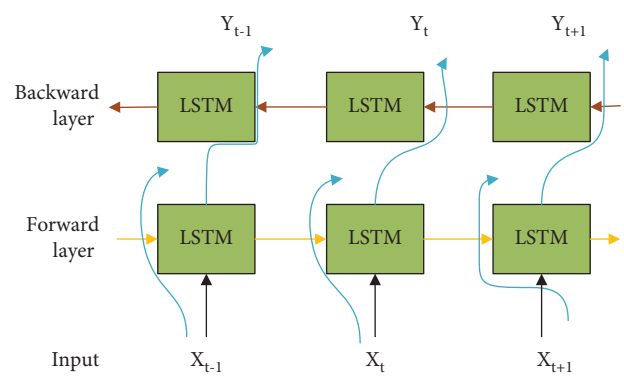


Figure 4: Neural network structure for obtaining topic depth semantics.

3.5.4. Optimized Training of the Neural Network. Dropout layer: it is the characteristic layer for overfitting in the neural network. In the process of forwarding propagation of the network, some activated neurons are randomly discarded, which can increase the necessary redundancy of the network. Meanwhile, in the case of loss of activated neurons, the model is allowed to maintain the correct classification, which reduces the problem of overadjustment so that the obtained model and training data will not be too much [33].

The formation of CNN is the principle of network training for each layer. The initial weight of the CNN training layer is random, and the backpropagation algorithm is used to adjust the corresponding weight value [34]. The reverse algorithm is divided into four parts: forward propagation, backpropagation, comparison and update of the loss function, and weight. The initial filter cannot effectively extract features. The loss function usually uses the average square error, that is, the semiaverage square error, to calculate the error between the result of the propagation phase prediction and the propagation of the true mark.

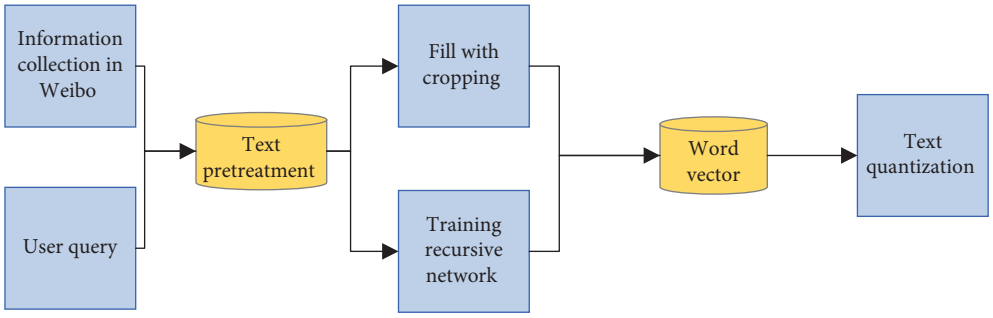


Figure 5: The vectorized representation process of Weibo text.

Table 1: Weibo topic detection on DNN.

Weibo topic detection steps on DCNN
Input: Weibo text data of the topic to be detected
(1) Text preprocessing of Weibo topic data
(2) Build a DCNN model and initialize the network
(3) Use the word vector representation of the Weibo used for training as the input of the DCNN to train the DCNN model
(4) For the Weibo to be detected, use the trained DCNN model to predict, obtain the topic label, and complete the Weibo topic detection
Output: Weibo hashtag

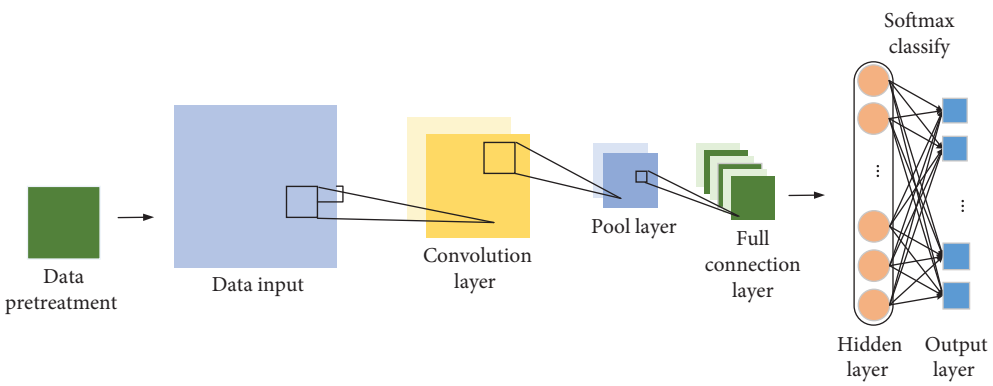


Figure 6: CNN text classification diagram.

Backpropagation determines important weights and adjusts the weights to reduce the overall error of the model. The detection process of the Weibo object detection algorithm on DCNN is shown in Table 1.

3.5.5. Implementation and Research of Weibo Security Topic Detection Model on DCNN. When building a new DCNN model, a large amount of data training set and related interference data set are usually needed for model training. It aims to detect and identify Weibo security topics. Therefore, the training data set and the interference data set are image data. The size of the image is defined as $224 * 224$ pixels. The DCNN model is designed on this basis. Information on Weibo can usually be grouped into topic templates, like the task of classifying text. The most important step is to obtain the expression characteristics of Weibo topics [35]. Traditional topic detection methods use unary language models to represent Weibo topics, often ignoring potential syntactic and semantic information. Using the DCNN model, it is possible to learn the grammatical and semantic information of Weibo more deeply, make the features of Weibo more accurate, and improve the accuracy of Weibo topic detection. The subject detection framework on DNN is shown in Figure 6.

Weibo is first preprocessed and trained through the network, and word vectors are obtained to represent each piece of Weibo information. Second, the vector matrix of Weibo keywords enters the DCNN, and the result shows that the feature vector of each Weibo can be obtained by training the DCNN. The word vector representation of Weibo is shown in the following equations:

$$P = w1 \cdots ws. \tag{10}$$

P is a Weibo sentence. The short Weibo text can be seen as a sentence. $w1 \cdots ws$ is the word vector of each word in Weibo.

The Weibo word vector matrix is input into the DCNN, and the feature map of the Weibo is obtained through the convolutional layer including the filter.

Table 2: Data security detection algorithm flow based on semantic feature extraction.

(1) Algorithm: DCNN
(2) **Input:** the raw points of $P = \{p_1, p_2, \ldots, p_n\}$
(3) **Output:** the ground surface points p_g
(4) **Parameters:** $b_0, b_k, H, \mu, \lambda, C_{th}, V_{s_n}$
(5) Mapping all the point cloud data into the polar grid $\mathrm{Grid}\ (m, k) = \mathrm{PolarGridMap}\,(P, b_0, b_k, H, \mu, \lambda)$
(6) Initialization $p_g = \varnothing$
(7) **For** $s = 0$: $m - 1$ do
(8) **For** $j = 0$: $k - 1$ do
(9) $\mathrm{Grid}\ (S_x, b_1) = \{p_i = (x_i, y_i, z_i)\}$
(10) The z-coordinate value of the point plus the laser installation $\mathrm{Grid}\ (S_x, b_1) = \{p_i = (x_i, y_i, z_i + H)\}$
(11) Order the points in the grid $\mathrm{Grid}\ (S_x, b_1)$ at height
(12) $n = 1$
(13) **IF** $p_1^2 > C_w$
(14) return
(15) **Else**
(16) **While** $\Delta z = p_{j+1} - p_i^2 < V_{ta}$
(17) $N = n + 1$;
(18) **End while**
(19) **End IF**
(20) **End For**
(21) **End For**
(22) **Return** p_g

Table 3: The text used in the experiment.

Number	Category	Quantity	Proportion (%)
1	Technology	75.5	7.55
2	Art	47.5	4.75
3	Life	65.3	6.53
4	Entertainment	138.4	13.84
5	Economy	124.1	12.41
6	Education	104.1	10.41
7	Electronics	163.2	16.32
8	Energy	104.4	10.44
9	Environment	107.6	10.76
10	History	20.2	2.02

$$c = [c_1, c_2, \ldots\ldots c_{n-h+1}]. \tag{11}$$

Ci is the feature of the Weibo topic after filtering the results.

$$ci = f\left(W \cdot X_{i.j+h-1} + b\right), \tag{12}$$

where h is the window size of the topic filter and b is the paranoid value.

This study aims to analyze and compare the performance of DCNN's Weibo security topic detection model. Irregular data detection algorithms are used in the three-dimensional data system to make judgments on topic extraction performance. Topic security is detected on the Internet. The algorithm flow of data detection and transmission based on Weibo text and semantic feature extraction is shown in Table 2.

3.6. Simulation Experiment. The experiment uses Weibo as a social network instance and uses the data crawled in Weibo as a verification data set for judging the security of social network topic content. The experimental data is divided into 10 in total. There are 1,000 documents in text language, including technology, art, life, entertainment, economy, education, electronics, energy, environment, and history. The detailed distribution is shown in Table 3. Since some categories of text are relatively small and not representative, categories with more than 100 texts are selected for experimentation. Among them, the training data is 40%, and the test data is 60%.

Table 4: Parameter settings of simulation experiment environment.

	Test environment parameters
System structure	B/S
CPU	Intel Core i5-6300HQ
Main frequency	2.3 GHz
RAM	8.00 GB
Development language	Java
Server database	Tomcat MySQL

To verify the proposed algorithm more objectively, supervised learning training is carried out in combination with the click tag of the Weibo content, and topic subtags are created for the click tag manually to further improve the accuracy of the evaluation. The algorithm uses MATLAB software to carry out data analysis and simulation experiments to verify the algorithm performance of the improved DCNN. In the experiment, the number of iterations of the neural network algorithm is 120, the simulation time is 2000 seconds, and the batch is 128. The objective function is optimized using MATLAB software. The environmental parameter settings of the software are shown in Table 4.

The constructed DCNN Weibo security topic mining and detection model is compared with neural network models proposed in other related fields, including CNN, DNN, and AlexNet, to evaluate the detection accuracy of the model. The prediction accuracy of the model is analyzed from the angles of accuracy, precision, recall, and $F1$ value.

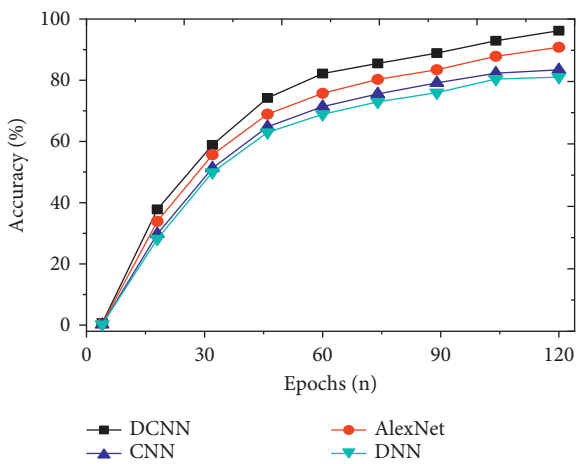


Figure 7: Comparison of the accuracy of different models.

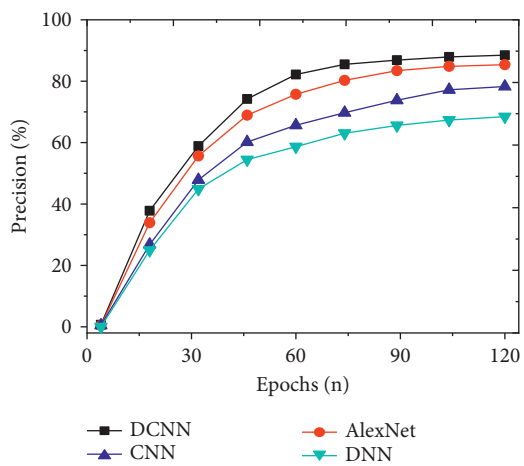


Figure 8: Comparison of precision of different models.

In addition, the above models are compared and analyzed, and the average transmission rate and average delay time of the models are used to evaluate the secure transmission performance of social network data.

4. Results and Discussion

4.1. Analysis of Detection Performance of Different Models. To study the detection effect of the improved DCNN on social network security topics, the detection results of the model are analyzed from the perspectives of accuracy, precision, recall, and $F1$ value. The indicators of the proposed model are compared with models, such as CNN, DNN, and AlexNet. And the results are shown in Figures 7–10. Further compare the training time and testing time required for each model, as shown in Figures 11 and 12.

As shown in Figures 7–10, the constructed system model is compared with the neural network model proposed by scholars in other related fields from the perspectives of accuracy, precision, recall, and $F1$ value. It is found that the recognition accuracy of the constructed model can reach 96.17% after 120 iterations of the model,

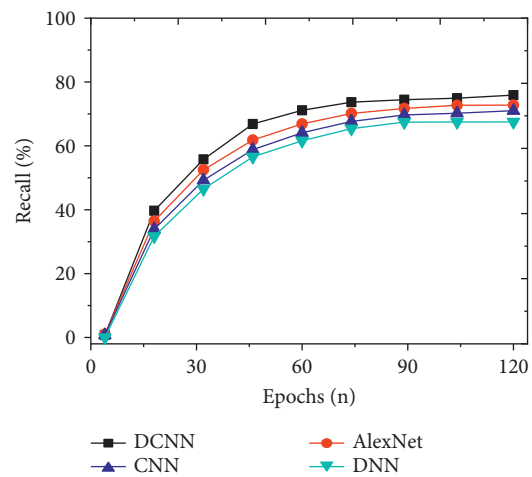


Figure 9: Comparison of recalls of different models with an increasing number of iterations.

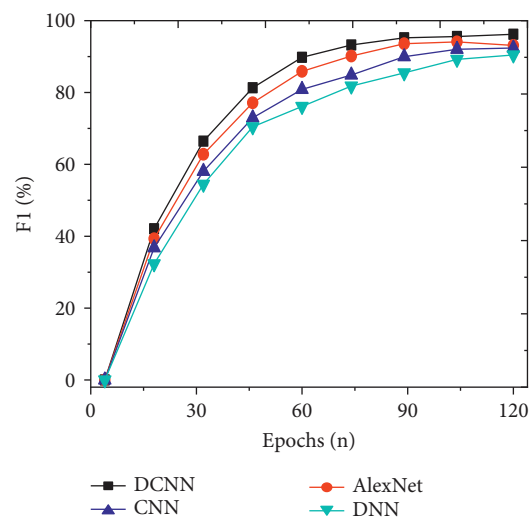


Figure 10: Comparison of $F1$ values of different models with an increasing number of iterations.

which is at least 5.4% higher than other models. Meanwhile, the accuracy, recall, and $F1$ value of the intrusion detection model are 88.57%, 75.22%, and 72.05%, respectively. Compared with other algorithms, the model's accuracy, recall, and $F1$ value are higher, at least 3.1% higher than other models. Therefore, compared with the network security detection model proposed by other scholars in related fields, the security detection model of the improved social network platform has better recognition and prediction accuracy.

This research further compares and analyses the training time and test time required for each algorithm, and the results are shown in Figures 11 and 12. As the number of iterations increases, the required training time and testing time show a trend of first decreasing and then basically stable; that is, convergence is achieved. In addition, the training time and testing time of the improved DCNN network security detection model stabilized at 65.86 s and

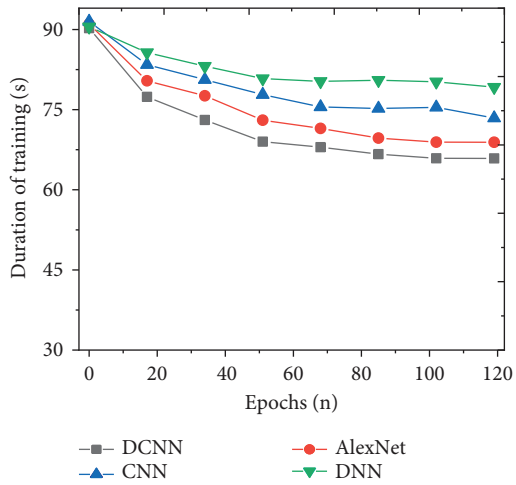


Figure 11: Comparison of training time of different models.

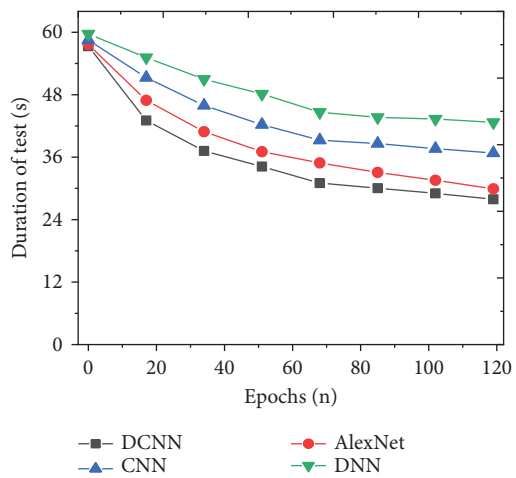


Figure 12: Comparison of test time of different models.

27.90 s, respectively. Compared with the models proposed by other scholars, the prediction time of the improved DCNN network security detection model is significantly shortened. This may be because the improved DCNN network security detection model can enhance the generalization ability and accelerate the convergence speed of the model training process. Therefore, for the mining and analysis of social network security topics, the improved DCNN network security detection model can obtain higher prediction results in a shorter time.

4.2. Data Transmission Security Performance Analysis of Different Network Models. From the aspects of the average transmission rate and data transmission delay of network data security transmission performance, the improved DCNN network security topic detection model is further compared with AlexNet, CNN, and DNN. The results are shown in Figures 13 and 14.

By comparing the network data security transmission performance of each model under different transmission

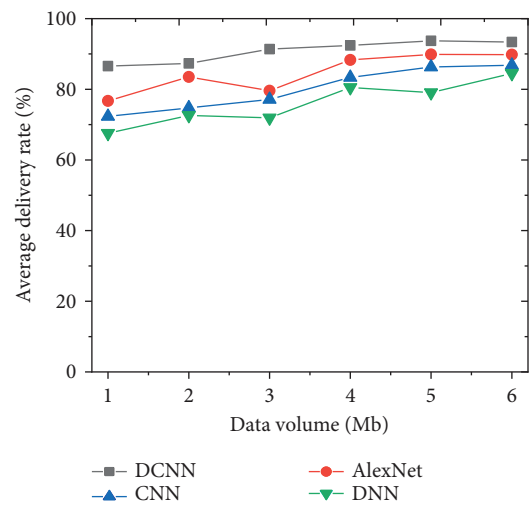


Figure 13: Comparison of the average data transmission rate of each model network under different transmission volumes.

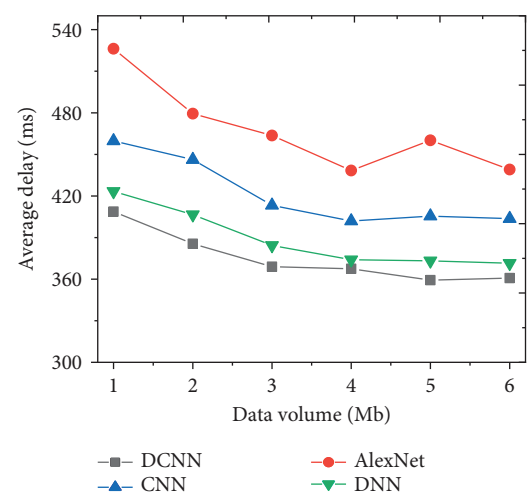


Figure 14: Comparison of average network data delay of each algorithm under different transmission volumes.

data, with the increase of transmission data, the average delivery rate of network data shows an upward trend. Figure 13 shows that the data message transmission rate of the improved DCNN network security detection model is not less than 80%. In terms of average delay, the average delay decreases with the increase of transmitted data. The average delay of the improved DCNN network security detection model is stable at about 360 ms, as shown in Figure 14. Therefore, from the perspective of different data transmission, the improved DCNN security detection model has the characteristics of lower delay and shows good network data security transmission.

4.3. Discussion of Results. The test results of the model are analyzed from the perspectives of accuracy, recall, and $F1$ value. The accuracy, recall, and $F1$ values of the intrusion detection model are 88.57%, 75.22%, and 72.05%,

respectively. Compared with other algorithms, this model has higher accuracy, recall, and $F1$ value. It is at least 3.1% higher than other models. The data message transmission rate of the improved DCNN network security detection model is not less than 80% compared with AlexNet, CNN, and DNN. The average delay decreases as the transmitted data increases. Li et al. [36] conducted gesture recognition research based on CNN, optimized the classification function of CNN, and improved the effectiveness and robustness of the entire model. Geirhos et al. [37] researched shortcut learning in DNNs. The study shows that fast learning is an important common feature of deep learning systems. Jiang et al. [38] researched the design of DNNs for photonic devices and discussed the network training process, the division of different network types and architectures, and the process of dimensionality reduction. The research has practical reference value for the simulation and design of the photonic system [39–45]. In summary, the results of CNN and DNNs have been applied in various fields in the research work of predecessors [46–49]. The difference between this study and previous studies is the use of DCNN to complete the analysis and search of social network security issues [50–52]. Meanwhile, the LSTM algorithm in the memory wisdom algorithm is used for the extraction of Weibo topic information. This is of great significance to the improvement of the security transmission performance of social network data.

5. Conclusion

With the rapid development of information technology, information security issues in social networks have become increasingly severe, and it is urgent to detect network attacks or intrusions. This study extracts the vector data of Weibo-related security topics through the research of social network security issues, combined with the mining and data analysis of network security topics. Weibo security topics are detected based on Latent Dirichlet Allocation and DNN. Meanwhile, DCNN's Weibo security topic detection model is implemented and used to improve the CNN. Combined with the DNN, the improved DCNN security detection model is implemented to ensure the safe operation of social networks. Finally, through the performance analysis of simulation experiments, the improved DCNN network security detection model predicts accuracy and precision rates of 96.17% and 88.57%, respectively, showing high prediction performance and good data transmission performance. The experimental results can be as follows: the security of social networks provides experimental evidence. However, some shortcomings still exist. Firstly, the dynamic growth of Weibo in real life is very fast, and the establishment of a fast, iterative text data representation is an important factor to be considered in the subsequent research. Secondly, there are not only texts on Weibo, but also a large amount of multimedia information, such as images and videos. How to mine the security topics of this part of the information is a difficult point in search research. Therefore, it is necessary to consider more factors in this aspect in future research.

Data Availability

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethical Approval

This article does not contain any studies with human participants or animals performed by any of the authors. Informed consent was obtained from all individual participants included in the study.

Conflicts of Interest

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

Authors' Contributions

All authors listed have made a substantial, direct, and intellectual contribution to the work and approved it for publication.

Acknowledgments

The authors acknowledge the help from the university colleagues.

References

[1] J. Yatabe, M. S. Yatabe, and A. Ichihara, "The current state and future of internet technology-based hypertension management in Japan," *Hypertension Research*, vol. 44, no. 3, pp. 276–285, 2021.

[2] D. Mulia and M. S. Shihab, "Strategy to maintain the cinema industry in the middle of development of internet technology," *Journal Manajemen*, vol. 24, no. 1, pp. 124–138, 2020.

[3] M. Valeri and R. Baggio, "Italian tourism intermediaries: a social network analysis exploration," *Current Issues in Tourism*, vol. 24, no. 9, pp. 1270–1283, 2021.

[4] S. Liu, "Computer network information security and protection measures under the background of big data," *Journal of Physics: Conference Series*, vol. 1881, no. 3, Article ID 032092, 2021.

[5] J. Guo and L. Wang, "Learning to upgrade internet information security and protection strategy in big data era," *Computer Communications*, vol. 160, pp. 150–157, 2020.

[6] P. Zhou and W. Zhang, "Research on computer network information security and protection strategy based on deep learning algorithm," in *Proceedings of the 2020 International Conference on Advance in Ambient Computing and Intelligence (ICAACI)*, pp. 181–184, Ottawa, ON, Canada, February 2020.

[7] L. Ding, Z. Wang, X. Wang, and D. Wu, "Security information transmission algorithms for IoT based on cloud computing," *Computer Communications*, vol. 155, pp. 32–39, 2020.

[8] Y. Lyu, H. Li, M. Sayagh, Z. M. J. Jiang, and A. E. Hassan, "An empirical study of the impact of data splitting decisions on the performance of AIOps solutions," *ACM Transactions on Software Engineering and Methodology*, vol. 30, no. 4, pp. 1–38, 2021.

[9] D. Bau, J. Y. Zhu, H. Strobelt, A. Lapedriza, B. Zhou, and A. Torralba, "Understanding the role of individual units in a

deep neural network," *Proceedings of the National Academy of Sciences*, vol. 117, no. 48, pp. 30071–30078, 2020.

[10] X. Li, J. Tang, Q. Zhang et al., "Power-efficient neural network with artificial dendrites," *Nature Nanotechnology*, vol. 15, no. 9, pp. 776–782, 2020.

[11] S. P. Rm, P. K. R. Maddikunta, M. Parimala et al., "An effective feature engineering for DNN using hybrid PCA-GWO for intrusion detection in IoMT architecture," *Computer Communications*, vol. 160, pp. 139–149, 2020.

[12] P. Devan and N. Khare, "An efficient XGBoost–DNN-based classification model for network intrusion detection system," *Neural Computing & Applications*, vol. 32, pp. 12499–12514, 2020.

[13] H. Chen, A. Chen, L. Xu et al., "A deep learning CNN architecture applied in smart near-infrared analysis of water pollution for agricultural irrigation resources," *Agricultural Water Management*, vol. 240, Article ID 106303, 2020.

[14] S. Igarashi, Y. Sasaki, T. Mikami, H. Sakuraba, and S. Fukuda, "Anatomical classification of upper gastrointestinal organs under various image capture conditions using AlexNet," *Computers in Biology and Medicine*, vol. 124, Article ID 103950, 2020.

[15] T. Kattenborn, J. Leitloff, F. Schiefer, and S. Hinz, "Review on convolutional neural networks (CNN) in vegetation remote sensing," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 173, pp. 24–49, 2021.

[16] M. Awais, X. Long, B. Yin et al., "A hybrid DCNN-SVM model for classifying neonatal sleep and wake states based on facial expressions in video," *IEEE Journal of Biomedical and Health Informatics*, vol. 25, no. 5, pp. 1441–1449, 2021.

[17] M. Alshaikh, M. Zohdy, R. Olawoyin, D. Debatosh, G. Zahraddeen, and A. Jalal, "Social Network Analysis and Mining: Privacy and Security on Twitter," in *Proceedings of the 2020 in 10th Annual Computing and Communication Workshop and Conference (CCWC)*, pp. 0712–0718, IEEE, Las Vegas, NV, USA, March 2020.

[18] W. Zhan and Z. Tao, "Research on 5G mobile communication network security technology," *Journal of Physics: Conference Series*, vol. 1634, no. 1, Article ID 012055, 2020.

[19] T. Maragatham, P. Yuvarani, and J. S. Shree, "Security concerns during photo sharing in social network platforms IOP conference series: materials science and engineering," *IOP Publishing*, vol. 1055, no. 1, Article ID 012084, 2021.

[20] N. Jindal and N. Jindal, "Copy move and splicing forgery detection using deep convolution neural network, and semantic segmentation," *Multimedia Tools and Applications*, vol. 80, no. 3, pp. 3571–3599, 2021.

[21] M. Ling, Q. Chen, Q. Sun, and Y. Jia, "Hybrid neural network for Sina Weibo sentiment analysis," *IEEE Transactions on Computational Social Systems*, vol. 7, no. 4, pp. 983–990, 2020.

[22] Q. Xie, X. Zhang, Y. Ding, and M. Song, "Monolingual and multilingual topic analysis using LDA and BERT embeddings," *Journal of Informetrics*, vol. 14, no. 3, Article ID 101055, 2020.

[23] Y. Liu and K. Cao, *Weibo Public Opinion Monitoring System Based on Sensitive Information Mining.Innovative Computing*, pp. 129–137, Springer, Singapore, 2020.

[24] L. Xu, L. Li, Z. Jiang et al., "A novel emotion lexicon for Chinese emotional expression analysis on Weibo: using grounded theory and semi-automatic methods," *IEEE Access*, vol. 9, pp. 92757–92768, 2021.

[25] Y. Chen, Z. Zhang, and Z. Xia, "Sentiment Assessment of Brand Advertising on Gender Issues on Social Network: A Case Study of Femvertising on Sina Weibo in China," in *Proceedings of the 2021 4th International conference on artificial intelligence and big data (ICAIBD)*, pp. 360–364, Chengdu, China, July 2021.

[26] P. Singh, Y. P. Huang, and S. I. Wu, "An intuitionistic fuzzy set approach for multi-attribute information classification and decision-making," *International Journal of Fuzzy Systems*, vol. 22, no. 5, pp. 1506–1520, 2020.

[27] X. Zhang and Y. Gao, "Retracted article: multimedia text classification algorithm using potential Dirichlet distribution in mobile cloud computing environment," *Multimedia Tools and Applications*, vol. 79, no. 13-14, pp. 9615–9627, 2020.

[28] N. Manouchehri, H. Nguyen, P. Koochemeshkian, N. Bouguila, and W. Fan, "Online Variational learning of Dirichlet process mixtures of scaled Dirichlet distributions," *Information Systems Frontiers*, vol. 22, no. 5, pp. 1085–1093, 2020.

[29] A. Wang and J. Zhang, "Topic Discovery Method Based on Topic Model Combined with Hierarchical Clustering," in *Proceedings of the 2020 IEEE 5th Information Technology and Mechatronics Engineering Conference (ITOEC)*, pp. 814–818, Chongqing, China, July 2020.

[30] M. Alhussein, K. Aurangzeb, and S. I. Haider, "Hybrid CNN-LSTM model for short-term individual household load forecasting," *IEEE Access*, vol. 8, pp. 180544–180557, 2020.

[31] R. P. Huebener, *Ginzburg–Landau Theory, Magnetic Flux Quantization, London Model.History and Theory of Superconductors*, pp. 19–24, Springer, Wiesbaden, 2021.

[32] J. Sedmidubsky, P. Budikova, V. Dohnal, and Z. Pavel, "Motion words: a text-like representation of 3D skeleton sequences," *Advances in Information Retrieval*, vol. 12035, p. 527, 2020.

[33] M. Mahmoud, I. Edo, A. H. Zadeh et al., "Tensordash: exploiting sparsity to accelerate deep neural network training," in *Proceedings of the 2020 53rd Annual IEEE/ACM International Symposium on Microarchitecture (MICRO)*, pp. 781–795, IEEE, Athens, Greece, November 2020.

[34] C. Yu, R. Han, M. Song, C. Liu, and C. I. Chang, "A simplified 2D-3D CNN architecture for hyperspectral image classification based on spatial–spectral fusion," *Ieee Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 13, pp. 2485–2501, 2020.

[35] K. Zhao, X. Gong, J. Cong, and Z. Shuhong, "Research on the phenomenon of fans "controlling comments" in cyberspace—taking sina Weibo as an example," in *Proceedings of the 2020 3rd International Seminar on Education Research and Social Science (ISERSS 2020)*, pp. 97–100, Atlantis Press, Kuala Lumpur, Malaysia, 2021.

[36] G. Li, H. Tang, Y. Sun et al., "Hand gesture recognition based on convolution neural network," *Cluster Computing*, vol. 22, no. S2, pp. 2719–2729, 2019.

[37] R. Geirhos, J. H. Jacobsen, C. Michaelis et al., "Shortcut learning in deep neural networks," *Nature Machine Intelligence*, vol. 2, no. 11, pp. 665–673, 2020.

[38] J. Jiang, M. Chen, and J. A. Fan, "Deep neural networks for the evaluation and design of photonic devices," *Nature Reviews Materials*, vol. 6, no. 8, pp. 679–700, 2020.

[39] Z. Lv, D. Chen, H. Feng, W. Wei, and H. Lv, "Artificial intelligence in underwater digital twins sensor networks," *ACM Transactions on Sensor Networks*, vol. 18, no. 3, pp. 1–27, 2022.

[40] B. Cao, J. Zhao, Z. Lv, and P. Yang, "Diversified personalized recommendation optimization based on mobile data," *IEEE Transactions on Intelligent Transportation Systems*, vol. 22, no. 4, pp. 2133–2139, 2021.

[41] R. Liu, X. Wang, H. Lu et al., "SCCGAN: style and characters inpainting based on CGAN," *Mobile Networks and Applications*, vol. 26, no. 1, pp. 3–12, 2021.

[42] W. Yang, X. Chen, Z. Xiong, Z. Xu, G. Liu, and X. Zhang, "A privacy-preserving aggregation scheme based on negative survey for vehicle fuel consumption data," *Information Sciences*, vol. 570, pp. 526–544, 2021.

[43] J. Mou, P. Duan, L. Gao, X. Liu, and J. Li, "An effective hybrid collaborative algorithm for energy-efficient distributed permutation flow-shop inverse scheduling," *Future Generation Computer Systems*, vol. 128, pp. 521–537, 2022.

[44] Y. Feng, B. Zhang, Y. Liu et al., "A 200–225-GHz manifold-coupled multiplexer utilizing metal waveguides," *IEEE Transactions on Microwave Theory and Techniques*, vol. 69, no. 12, pp. 5327–5333, 2021.

[45] C. Qin, G. Shi, J. Tao et al., "An adaptive hierarchical decomposition-based method for multi-step cutterhead torque forecast of shield machine," *Mechanical Systems and Signal Processing*, vol. 175, p. 109148, Article ID 109148.

[46] C. Qin, D. Xiao, J. Tao et al., "Concentrated velocity synchronous linear chirplet transform with application to robotic drilling chatter monitoring," *Measurement*, vol. 194, Article ID 111090, 2022.

[47] J. Yan, H. Jiao, W. Pu, C. Shi, J. Dai, and H. Liu, "Radar sensor network resource allocation for fused target tracking: a brief review," *Information Fusion*, vol. 86-87, pp. 104–115, 2022.

[48] Z. Ma, W. Zheng, X. Chen, and L. Yin, "Joint embedding VQA model based on dynamic word vector," *PeerJ Computer Science*, vol. 7, p. e353, Article ID e353, 2021.

[49] Y. Zhang, X. Shi, H. Zhang, Y. Cao, and V. Terzija, "Review on deep learning applications in frequency analysis and control of modern power system," *International Journal of Electrical Power & Energy Systems*, vol. 136, Article ID 107744, 2022.

[50] S. S. Yang, X. L. Yu, M. Q. Ding et al., "Simulating a combined lysis-cryptic and biological nitrogen removal system treating domestic wastewater at low C/N ratios using artificial neural network," *Water Research*, vol. 189, Article ID 116576, 2021.

[51] F. Zhang, J. Zhai, X. Shen, O. Mutlu, and X. Du, "POCLib: a high-performance framework for enabling near orthogonal processing on compression," *IEEE Transactions on Parallel and Distributed Systems*, vol. 33, no. 2, pp. 459–475, 2022.

[52] B. Cao, Y. Gu, Z. Lv, S. Yang, J. Zhao, and Y. Li, "RFID reader anticollision based on distributed parallel particle swarm optimization," *IEEE Internet of Things Journal*, vol. 8, no. 5, pp. 3099–3107, 2021.

Hindawi
Computational Intelligence and Neuroscience
Volume 2022, Article ID 2249417, 14 pages
https://doi.org/10.1155/2022/2249417

Research Article

Facial Emotion Recognition Using a Novel Fusion of Convolutional Neural Network and Local Binary Pattern in Crime Investigation

Dimin Zhu,[1] Yuxi Fu,[2] Xinjie Zhao,[3] Xin Wang,[4] and Hanxi Yi[5]

[1]*School of Law, Zhejiang Gongshang University, Hangzhou, Zhejiang Province 310000, China*
[2]*Department of Science and Technology, Beijing Normal University-Hong Kong Baptist University United International College, Zhuhai 519087, China*
[3]*School of Software and Microelectronics, Peking University, Beijing, China*
[4]*Behavioural Science Institute, Radboud University, Nijmegen 6525 GD, Netherlands*
[5]*Division of Biopharmaceutics and Pharmacokinetics, Xiangya School of Pharmaceutical Sciences, Central South University, Changsha 410000, China*

Correspondence should be addressed to Xin Wang; xin.wang@ru.nl

Received 11 July 2022; Revised 10 August 2022; Accepted 24 August 2022; Published 22 September 2022

Academic Editor: Ning Cao

The exploration of facial emotion recognition aims to analyze psychological characteristics of juveniles involved in crimes and promote the application of deep learning to psychological feature extraction. First, the relationship between facial emotion recognition and psychological characteristics is discussed. On this basis, a facial emotion recognition model is constructed by increasing the layers of the convolutional neural network (CNN) and integrating CNN with several neural networks such as VGGNet, AlexNet, and LeNet-5. Second, based on the feature fusion, an optimized Central Local Binary Pattern (CLBP) algorithm is introduced into the CNN to construct a CNN-CLBP algorithm for facial emotion recognition. Finally, the validity analysis is conducted on the algorithm after the preprocessing of face images and the optimization of relevant parameters. Compared with other methods, the CNN-CLBP algorithm has higher accuracy in facial expression recognition, with an average recognition rate of 88.16%. Besides, the recognition accuracy of this algorithm is improved by image preprocessing and parameter optimization, and there is no poor-fitting. Moreover, the CNN-CLBP algorithm can recognize 97% of the happy expressions and surprised expressions, but the misidentification rate of sad expressions is 22.54%. The research result provides data reference and direction for analyzing psychological characteristics of juveniles involved in crimes.

1. Introduction

The popularization of the Internet has caused immense changes in the social environment, resulting in a significant impact on the outlook on life and values of young people, followed by psychological problems that cannot be ignored [1, 2]. Therefore, it is of vital significance to seek a method that can accurately judge the psychological state and psychological characteristics of minors. In the context of the burgeoning artificial intelligence technology, the development of deep learning and natural language processing technology provides technical support for the solution to this problem [3–5]. In the field of psychological analysis, emotion recognition is also called emotion analysis, which is used to analyze complex emotional states [6]. Facial expression is an external and intuitive reflection of emotion, and facial expression recognition technology can identify human emotional state. So far, many research results have been achieved on the application of deep learning to facial emotion recognition of excellent applicability.

The conventional methods of facial emotion recognition usually collaborate with feature extraction, feature classification, and data dimension reduction algorithms and put

forward stringent requirements for the running state of algorithms at each stage [7]. Neural network models are constantly improving and developing with the improvement in computers. On the one hand, artificial intelligence technology is primarily employed for the perception of emotion. Besides, the progress of neural network models and other technologies also provides a new direction and idea for solving problems in common facial emotion recognition. On the other hand, facial emotion recognition technology has mainly experienced three stages of development. In the primary stage, the identification of facial expressions is realized through the facial action coding system. In the second stage, facial emotion recognition is manifested in systematic recognition, including three steps, namely, information collection, feature extraction, and expression recognition. The methods used in this stage primarily contain support vector machine (SVM), Gabor wavelet, and Local Binary Pattern (LBP). However, in this stage, the recognition accuracy needs to be enhanced because the facial emotion recognition excessively depends on feature algorithm classifiers and explicit expression features. In the last stage, the deep learning algorithm is the most dominant approach to facial emotion recognition. Its most predominant feature is that it can input and output relevant data through the deep learning model instead of artificial classifiers, and this method significantly alleviates local optimization. Among diversified deep learning algorithms, the convolutional neural network (CNN) algorithm can achieve excellent performance in image recognition and attains higher recognition accuracy than the conventional algorithms.

However, the current CNN model is incompetent. Therefore, the CNN model is optimized before recognizing facial expressions and analyzing facial emotions. Specifically, a facial expression recognition algorithm is proposed by integrating the CNN model and Central Local Binary Pattern (CLBP) to analyze and identify facial expressions. It is expected to provide reference for the optimization of deep learning algorithm for facial expression recognition and preliminarily explore the correlation between facial expression recognition and psychological feature analysis.

The main contributions of this paper are summarized as follows:

(1) A fusion CNN model is constructed for facial emotion recognition. The comparative analysis indicates that the overall performance of the model reported here is better than the traditional LBP model, LeNet-5 model, and VGGNet model. The highest recognition accuracy of happy expressions of the fusion CNN model attains 98.61%, which is better than 94.11% of the VGGNet model.

(2) A hybrid CNN-CLBP facial emotion recognition algorithm is proposed based on optimized CLBP algorithm. This algorithm has outstanding performance in facial expression recognition, and the average recognition rate of happy expressions and surprised expressions is 97%.

(3) The research results provide a feasible direction for the analysis of psychological characteristics based on emotion recognition.

2. Literature Review

Some scholars have made efforts into the recognition of facial expressions. For instance, Fan and Tjahjadi proposed a facial expression recognition framework by combining CNN and handcrafted features. They found that the neural network could achieve brilliant recognition effects by extracting texture information from facial patches, and the introduction of CNN had a promoting effect on facial expression recognition [8]. Reddy et al. put forward an organic combination of deep learning features and the manual facial expression recognition method. They verified the applicability of this method in the wild scenes in experiments, which revealed the effectiveness of the recognition method under the combination of deep learning and manual production [9]. Liang et al. emphasized that the traditional handcrafted facial representations could only reveal shallow features. To break out this limitation, they proposed a new facial expression recognition method based on patches of interest, the Patch Attention Layer of embedding handcrafted features, to learn the local shallow features of each patch on face images. They finally proved the effectiveness of the method in their study [10]. Jain et al. (2018) built a hybrid convolutional-recursive neural network for facial expression recognition, which extracted information on face images through the combination of convolutional layer and Recurrent Neural Network (RNN) in view of the time correlations in images [11]. Avots et al. identified human emotions through analyzing audiovisual information. Besides, they classified the emotions on face images using different datasets as test sets, Viola-Jones face recognition algorithm, and CNN (AlexNet) [12]. Li et al. constructed a two-dimensional principal component analysis network via deep learning based on L1 norm for face recognition, tested it on some facial image database, and found that the network was robust [13]. Bernhard et al. stated that emotions had a significant influence on decision-making of humans, and hence, they applied deep learning to improve emotion recognition results. They found that the performance of RNNs and transfer learning was better than traditional machine learning, which had great inspirations on emotion computing application [14]. Kumar et al. discussed the modeling of abnormal facial expressions based on computer vision tasks and emotional deviations. They found that deep CNN could play an essential role in the training and classification of facial expressions, which provided a new visual modeling method for visual surveillance systems [15]. Mishra et al. employed CNN to recognize different emotions and intensity levels of human faces, which provided the basis and support for future research on computer emotion recognition [16].

In summary, the traditional manual methods are no longer suitable for the current research on facial expression recognition, but deep learning can significantly improve the recognition effectiveness of facial expression, especially the

hybrid model. Although some achievements have been made in the previous studies of facial expression recognition, there are few studies combining it with psychological analysis.

3. Materials and Methods

3.1. CNN Model Based on Deep Learning. CNN is a typical neural network based on deep learning, which can reduce the resolution of features, thereby showing significant advantages in the extraction of image features and analyzing modal image data [17, 18]. CNNs are primarily composed of the convolution layer, the pooling layer, and the fully connected layer [19]. Among them, the convolution layer is the most significant component distinguishing the CNN from other neural networks, which can perform convolution operations on image features to obtain various forms of feature maps and all the edge information on the image simultaneously. Besides, the convolution layer of CNNs is mainly characterized by functions including local connections, receptive fields, and shared weights, ensuring the high efficiency of CNNs in feature extraction and feature calculation. Meanwhile, it enables the CNN to differentiate different images during feature extraction, thereby optimizing the global extraction effect and reducing the training parameters of CNN through shared weights [20]. This layer highlights the initial information about the image and significantly lessens the calculative burden of the model. Consequently, the matching rate of the entire training process is greatly improved [21]. The fully connected layer of the CNN can accomplish the connection to the local features of the image [22], which can be expressed as

$$l_{W,b}(x) = f\left(W^T x + b\right), \tag{1}$$

where f represents the nonlinear activation function, x denotes the input, l refers to the output, W^T signifies the connection weight, and b represents the offset.

The crucial procedure of facial emotion recognition technology based on psychoanalysis is the extraction and classification of image features corresponding to facial expressions. The hidden layer of the CNN completes the extraction of image features and classifies image features by means of classifiers. Among them, Softmax is a widely used classifier, which mainly completes the classification by calculating its probability, as shown in

$$f\left(o_j\right) = \frac{e^{o_j}}{\sum_{n=1}^{n} e^{o_k}}. \tag{2}$$

In equation (2), n denotes the number of outputs, and o_j stands for the probability of the output results.

Furthermore, in the neural network model, Dropout can realize the zero processing of some weights.

In short, the CNN is selected for feature extraction and facial emotion recognition based on psychological feature analysis here since it is superior to other neural networks.

3.2. Construction of the Facial Emotion Recognition Model Based on Hybrid CNN Network. CNN shows outstanding performance in image feature extraction, but it has some shortcomings. In concrete application scenarios, it is essential to design the CNN structure according to specific recognition tasks. Besides, the model's training is greatly affected by the number of network layers [23–25]. Specifically, the training effect is better with more network layers. As the corresponding weight parameters increase, the entire training duration enlarges. Otherwise, the training process will be shorter. For the recognition task, the deep network's recognition accuracy is better than that of the external network, while the recognition efficiency of the external network is better than that of the deep network. Hence, it is necessary to fully consider the network's recognition accuracy and efficiency for the optimal recognition effect of the neural network. Furthermore, more attention is paid to the recognition accuracy of the proposed model, which is a critical evaluation indicator. Figure 1 illustrates the structure of CNN with sentences as input.

VGGNet, AlexNet, and LeNet-5 are common CNN models. The VGGNet model uses convolution kernels of size $3*3$, so it has fewer network parameters and a stronger learning capability than other models with convolution kernels of size $5*5$. However, it has more complex structural components and is not applicable to the classification of small-scale data information [26]. The AlexNet model utilizes multiple graphics processing units (GPUs) to complete the model training, which improves the overall training speed, recognition accuracy, and recognition rate of the model. It plays an essential role in solving the gradient disappearance problem [27, 28]. The LeNet-5 model is the simplest CNN model, which is suitable for low-resolution image recognition and processing of available classification data [29].

3.3. Fused CNN Model

3.3.1. VGGNet Model. VGGNet further explains the correlation between the depth and performance of CNN. One of the most prominent contributions of VGGNet model is to reveal that the increase in layers of the CNN model can significantly reduce the error rate and achieve excellent expansion performance and generalization performance, enabling CNN to efficiently extract features from images. The VGGNet model can be regarded as a deeper version of AlexNet model, which is composed of five convolution layers, three full connection layers, and a Softmax output layer. In summary, the model has a relatively simple structure and has a small convolution kernel and pooling kernel, as well as more convolution sublayers and channels. Meanwhile, the model has deeper layers and wider feature maps.

3.3.2. AlexNet Model. The AlexNet model is also a key link to the development of CNN, which realizes the excellent performance of CNN in the field of image classification. The innovation of AlexNet model lies in the smooth application of ReLU activation function and Dropout mechanism, the use of overlapping maximum pooling, the proposal of local

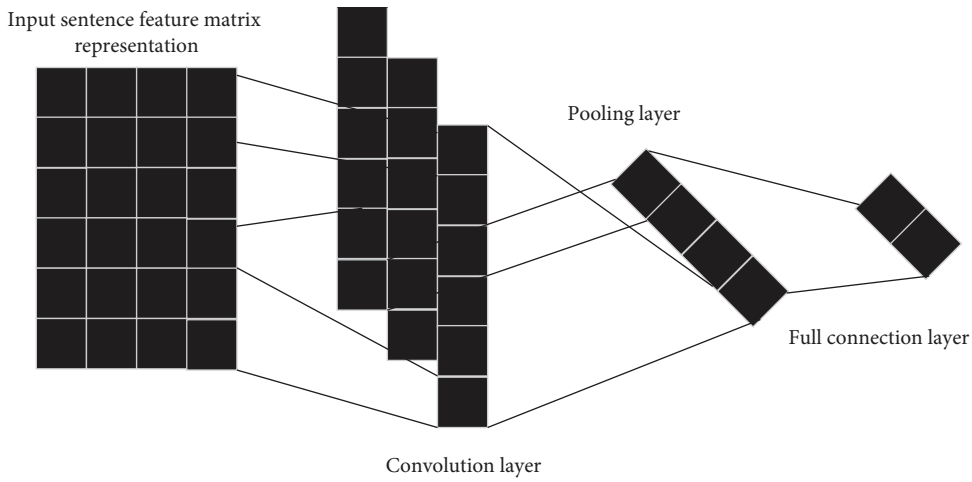


Figure 1: Structure of CNN.

Table 1: Parameter settings of the facial emotion recognition neural network based on psychological feature analysis.

Network structure	Convolution layer	Pooling layer	Fully connected layer	Output layer
Output size	1 : 128∗128∗1 2 : 64∗64∗1 3 : 32∗32∗1 4 : 14∗14∗6	1 : 28∗28∗6 2 : 10∗10∗16	1 : 1∗1∗1 2 : 1∗1∗6	1∗1∗6

response normalization, and the accelerated training of GPU. Among them, the use of ReLU can effectively avoid the occurrence of overfitting and alleviate the interdependence between different parameters to some extent.

3.3.3. LeNet-5 Model. LeNet-5 is a classical CNN model, and its unique structure of convolution-pooling alternation can effectively extract the translation invariant features of the input image. As the name suggests, the LeNet-5 model has five layers in its structure, namely, two convolution-pooling layers and three fully connected layers.

To sum up, these three models have distinctive excellent performance. Based on this, the advantages of each model are organically integrated to construct a fused CNN model.

3.4. Parameter Setting of the Hybrid Neural Network. The facial expression generally accounts for a small proportion of an image, so it is difficult to achieve the optimal effect by applying the above neural network model separately when processing images of such characteristics. As a result, the above three neural network models are combined to construct a fused CNN model for facial emotion recognition, comprising four convolution layers, two pooling layers, two fully connected layers, and a Softmax classification layer. Table 1 reveals the specific parameter settings of the model. At the initial stage of facial emotion recognition by the hybrid CNN model, the input of the grayscale feature image of facial expressions is completed in the input layer. Next, it is followed by the convolution layer, which corresponds to 1–3 in the following table. Here, the image feature extraction is completed through calculating the output of the region neuron corresponding to the input. Moreover, its back corresponds to the pooling layer 1, which is primarily responsible for the pooling processing on the point corresponding to the maximum value in the region, which is as the values after pooling process. The subsequent layers following pooling layer 1 are convolution layer 4, pooling layer 2, and finally, two fully connected layers and an output layer. Among them, the fully connected layer 1 contains 120 fully connected neurons, and the fully connected layer 2 contains 64 fully connected neurons. The output layer corresponds to the Softmax layer containing six neurons, mainly used to predict the six types of output of human facial emotions.

3.5. Facial Emotion Recognition Algorithm Based on Feature Fusion. LBP is an algorithm describing image texture features, and the operator of LBP uses each pixel in the image to calculate the binary mode. In this process, each component pixel in the image is the central pixel, and the corresponding threshold processing is completed through the eight pixels in the neighborhood. When the adjacent pixel is smaller than the central pixel, the adjacent pixel value is 0. Otherwise, the adjacent pixel value is 1 [30, 31]. The operator for image extraction can be expressed as

$$LBP\left(x_p, y_p\right) = \sum_{p=0}^{p-1} 2^p S\left(I_p - I_c\right), \tag{3}$$

where (x_p, y_p) is the information corresponding to the center position in the neighborhood, and I_p represents the pixel value of the image. Meanwhile, I_c signifies the pixel value of

other images in the region, and S denotes a symbolic function.

However, the conventional LBP operator only covers a small area within a fixed range. Therefore, the LBP operator has limitations on scenarios of high requirements for textures of different sizes or frequency. It is unsuitable for scenarios where human faces and other images are subject to change. On this basis, the LBP algorithm is improved by adding the center pixel, and the improved LBP algorithm is expressed as CLBP, the LBP operator of the center pixel. CLBP can be written as

$$F_{\mathrm{CLBP}} = \sum_{p=0}^{p-1} S(I_c, c_I), S(x, c_I). \tag{4}$$

In equation (4), F_{CLBP} refers to the CLBP operator, S signifies the symbolic function, I_c represents the central pixel of the image, and c_I denotes the average gray value information of the corresponding image.

Due to the rotation and other influencing factors of the image processing, the robustness of CNNs will decrease in the application process. The rotation-invariant characteristics of the LBP algorithm can compensate for the defects of the CNN. Therefore, a facial emotion recognition algorithm is proposed by integrating CNN features with CLBP features. In this hybrid algorithm, feature fusion is completed in the second fully connected layer of the CNN. The specific feature fusion process includes selecting a suitable feature fusion method and the dimensionality reduction operation. Here, the feature fusion method based on connection is adopted, and (5) describes the feature fusion of the two output vectors.

$$\mathrm{consat}(x, y) = (x_1, x_2, \cdots, x_n, y_1, y_2, \cdots, y_n). \tag{5}$$

Then, image dimension reduction is realized by principal components analysis to remove the redundant information during image feature fusion. The specific implementation of this step is the centralization of the sample first, as shown in

$$u_i = x_i - \overline{x}, \tag{6}$$

where $\overline{x}$ represents the average value corresponding to the sample data. The subsequent covariance matrix can be presented as

$$C = \frac{1}{N} \sum_{i=1}^{N} (x_i - \overline{x})(x_i - \overline{x})^r = MM^T, \tag{7}$$

where M represents the matrix corresponding to the sample vector that has been centralized.

Equation (8) indicates the cumulative contribution degree η corresponding to the feature value.

$$\eta = \sum_{k=1}^{i} \frac{\lambda_j}{\sum_{i=1}^{h} \lambda_j}. \tag{8}$$

In equation (8), λ_j represents the feature value of the feature matrix, and h illustrates the number of feature values. Assume that the feature vector corresponding to the feature value is

$$Z = (\eta_1, \eta_2, \cdots, \eta_t). \tag{9}$$

Then, the projection of the feature matrix in it can be calculated according to

$$H_L = H * Z, \tag{10}$$

where H_L represents the feature matrix after the dimensionality reduction.

The neural network fused with the CNN feature and the CLBP feature is represented as hybrid CNN-CLBP model. For the hybrid CNN-CLBP model, the face images of the selected data set are cropped first. Then, the face image samples are rotated and normalized. Besides, the samples are divided into a training set and a test set according to the proportion of 8 : 2. The output facial expression features are input to the Softmax layer to obtain the classification results. Figure 2 reveals the corresponding structural composition of CNN-CLBP.

3.6. Training of the Facial Emotion Recognition Model Based on a Neural Network Algorithm. Static data sets and dynamic data sets are often used to recognize facial expressions, among which JAFFE is a common static image database [32], and CK+ (Cohn-Kanade+) and Fer2013 are image data sets composed of dynamic expression sequences [33, 34]. Considering that facial expression recognition generally processes dynamic expression sequences, the CK+ and JAFFE are chosen as the research data sets. Table 2 presents the image composition of the two data sets.

In the detection and recognition of facial expressions, the JAFFE data set and the CK + data set are common. In addition, the information contained in these two databases is more comprehensive, which can effectively improve the model performance and the accuracy for facial emotion recognition. Due to facial expressions' dynamic characteristics, images of six typical expressions of the CK + data set are chosen for facial emotion recognition. Some face images of the JAFFE data set are shown in Figure 3. Some face images of the CK + data set are presented in Figure 4.

In addition to selecting data sets, positioning detection is a critical step in recognizing human facial expressions. The accuracy of positioning detection directly impacts the recognition accuracy. Therefore, it is vital to select applicable positioning detection algorithms. In other words, a positioning detection algorithm meeting the requirements needs to consider both accuracy and efficiency. Consequently, the Haar-like algorithm is adopted to describe the facial features of human faces, and the AdaBoost algorithm is adopted for classification. In the process of feature extraction using Haar-like, feature values can be expressed as

$$f_h = \left| a \sum_{(x,y)\in S_A} i(x, y) - b \sum_{(x,y)\in S_B} i(x, y) \right|, \tag{11}$$

where S_A represents the black composition in the region, and S_B denotes the white composition in the region. Besides, a

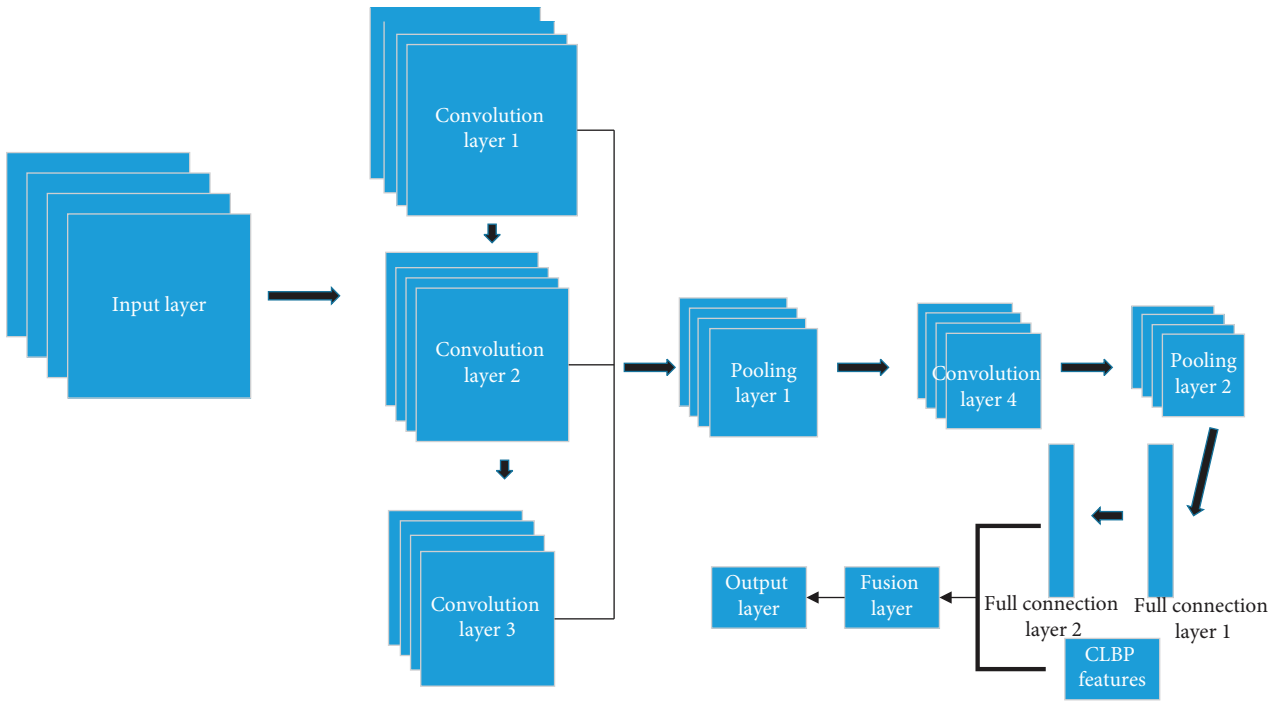


Figure 2: Structure of the hybrid CNN-CLBP model.

Table 2: Image composition of Fer2013 and CK + data sets.

Datasets	JAFFE	CK+
Image composition	Number of images: 35886 Size: 48∗48 pixel Participants: 10 Tags: happy, fear, sad, surprised, angry, disgusted, neutral	Number of images: 593 Size: 640∗480 pixel Participants: 123 Tags: happy, fear, sad, surprised, contempt, anger, disgust, neutral

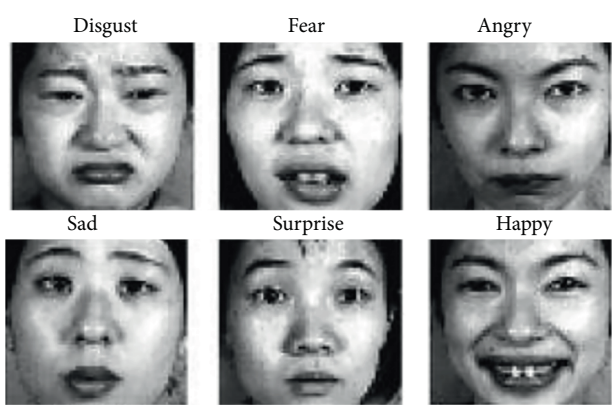


Figure 3: Human face images of the JAFFE data set (the data source: https://blog.csdn.net/akadiao/article/details/79956952).

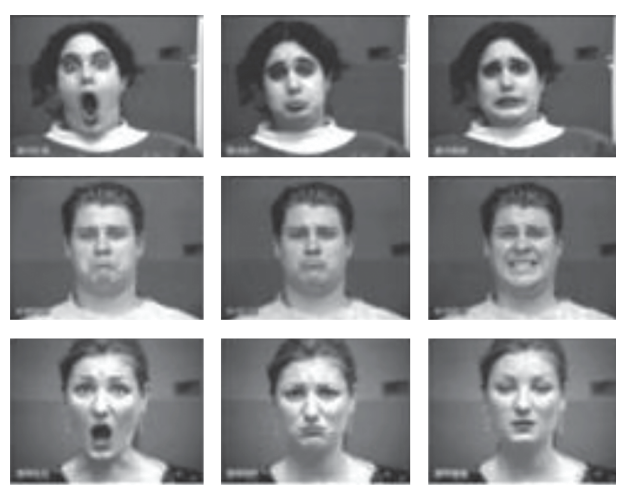


Figure 4: Human face images of the CK + data set (the data source: https://blog.csdn.net/yinghua2016/article/details/77323537).

refers to the proportion of the black area in the area, b stands for the proportion of the white area in the area, and $i\ (x, y)$ represents the corresponding pixel value in the image feature interval.

Furthermore, the feature number can be determined according to

$$\mathrm{SUM} = XY\left(W + 1 - w\frac{X+1}{2}\right)\left(H + 1 - h\frac{Y+1}{2}\right). \quad (12)$$

In (12), W represents the width of the image, H signifies the height of the image, w denotes the width of the rectangle,

and h refers to the height of the rectangle. Meanwhile, X represents the magnification factor of the rectangular feature in the horizontal direction, and Y stands for the rectangular feature magnification factor in the vertical direction.

For the AdaBoost training strong classifier, N training samples are expressed as

$$(X_1, Y_1), \cdots (X_N, Y_N), \tag{13}$$

where $Y_i = 0$ indicates negative samples of nonface data, and $Y_i = 1$ refers to positive samples of image data. Then, the weight initialization process is performed. In the case of $Y_i = 0$, the weight can be written as (4).

$$W_{i,j} = \frac{1}{2m}. \tag{14}$$

In the case of $Y_i = 1$, the weight can be expressed as

$$W_{i,j} = \frac{1}{2l}, \tag{15}$$

where m represents the number of negative samples, and l denotes the number of positive samples. The normalization of weights can be presented as

$$\frac{W_{i,j}}{\sum_{j=1}^{N} W_{i,j}} \longrightarrow W_{i,j}. \tag{16}$$

Then, it is necessary to train the model using the features by a weak classifier. The final step of classification using the AdaBoost algorithm is to update the weights. The strong classifier can be expressed as

$$h(x) = \begin{cases} 1 & \sum_{t=1}^{T} a_l h_l(x) \geq \frac{1}{2} \sum_{t=1}^{T} a_t, \\ 0 & \text{other} \end{cases} \tag{17}$$

$$a_t = \log \frac{1}{b_i}.$$

Before training the model using images, since the background information on the original image may obstruct recognizing and detecting the image, the image preprocessing mainly contains grayscale processing, cropping processing, and normalization. Generally, the CNN has an excellent performance in image processing, and it is unnecessary to preprocess images or extract features, since the fine-grained feature extraction of CNN can process the images. However, face images usually involve complicated information that is affected by multiple factors, such as the visual angle and background information. Therefore, the image information cannot be exactly exacted by a separate operation. The initial image information is processed via grayscale, cropping, and normalization. The grayscale processing is to convert the color image into a grayscale image with a single channel feature. After this operation, both the influence of light intensity and the calculation complexity in the training process decrease, improving the model's training speed. Specifically, the gray value conversion can be calculated according to

$$0.299R + 0.587G + 0.114B = Y, \tag{18}$$

where R represents the red channel in the image, G denotes the green channel in the image, B refers to the blue channel in the image, and Y stands for the gray value.

Furthermore, the image cropping is indispensable since there exists a considerable amount of disturbance information on the initial face image, which may reduce classification accuracy. Therefore, the face image is cut before expression recognition. The specific cutting method is that, in the horizontal direction, a crop factor of 0.7 is selected to complete the image cropping process. In the vertical direction, a crop factor of 0.3 is selected to complete the cropping process. After the cropping operation, the information on the image irrelevant to facial expressions is removed, and the size of the image is significantly reduced, which can significantly reduce the workload of the subsequent training. The final part after the cropping processing is the normalization of the image. Specifically, during the normalization processing, the initial image of the data sets is rotated by -30°, -15°, 15°, and 30°, respectively, considering that there may be nonfrontal facial images. Subsequently, facial emotion recognition is performed after the normalization. The image preprocessing mainly aims to reduce the impact of uneven lighting on the facial emotion recognition. In the normalization of the image, the histogram is used to equalize the image, which can be expressed as

$$s_i = H(n) = \sum_{i=0}^{k-1} \frac{n_l}{n}, \tag{19}$$

where n denotes the total number of pixels in the human face image, k represents the type of gray value, and n_l refers to the total amount corresponding to the l-th type of gray value.

It is essential to use a feasible optimization algorithm to find the model's optimal global solution to train the CNN model. Adam algorithm is an optimization algorithm developed based on a stochastic gradient algorithm. The algorithm has the characteristics of an adaptive gradient and root mean square propagation. The adaptive gradient provides the algorithm with excellent performance in computer vision, and root mean square propagation affords the algorithm excellent performance in solving intermittent problems. The mean value of the initial time gradient in the Adam algorithm can be obtained according to

$$\widehat{M_t} = \frac{M_l}{1 - b_1^l}. \tag{20}$$

The noncentral variance corresponding to the gradient at the second moment can be expressed as

$$\widehat{V_t} = \frac{V_l}{1 - b_2^l}, \tag{21}$$

where $\widehat{M_t}$ indicates the mean value of the corresponding gradient at the initial time, $b_1 = 0.9$, $\widehat{V_t}$ represents the noncentral variance of the corresponding gradient at the second time, and $b_2 = 0.999$.

This algorithm updates the parameters according to

$$\theta_{t+1} = \theta_I = \frac{\eta}{\sqrt{\widehat{V}_t + \varepsilon}}\widehat{M}_t. \tag{22}$$

The Adam algorithm has excellent convergence performance and requires a small amount of memory space. Therefore, the algorithm can solve optimization problems, including numerous data information and parameters [35]. Therefore, the algorithm is selected as a tool to optimize the neural network.

The cross-entropy loss function can measure and evaluate the difference between the probability distributions. Here, this function presented in (23) is selected as the loss function.

$$h(p,q) = -\sum_x p(x)\log q(x). \tag{23}$$

In equation (24), p refers to the correct probability distribution value, and q represents the predicted value. Furthermore, discrete variables can be calculated according to

$$h(p,q) = \sum_x p(x)\cdot\log\left(\frac{1}{q(x)}\right). \tag{24}$$

Continuous variables can be decided according to

$$-\int_x p(x)\,\mathrm{logq}(x)\,dr(x) = E_p[-\log q]. \tag{25}$$

3.7. Comparative Experiment. The CK + data set is chosen to verify the effectiveness of the hybrid CNN-CLBP model. On the premise of ensuring that all training parameters are consistent, the average and maximum values are used as comparison indicators. The hybrid CNN-CLBP model is compared with the traditional machine learning LBP model, a single LeNet-5 model, and a single VGGNet model. Table 3 provides the particular parameter settings of the LeNet-5 model and the VGGNet model.

For the facial emotion recognition algorithm based on feature fusion, after a series of preprocessing operations on the image, quantitative analysis is performed on the three indicators of accuracy, cross-entropy, and loss function to test the effectiveness of the feature fusion method.

4. Results

4.1. Comparative Analysis of Facial Emotion Recognition Results. Figure 5 indicates the facial emotion recognition results of the hybrid CNN-CLBP model, the traditional machine learning model LBP, the LeNet-5 model, and the VGGNet model.

From Figure 5(a), the average recognition rate of the traditional machine learning LBP model reaches 48.63%, that of the single LeNet-5 model is 73.22%, that of the single VGGNet network model attains 83.17%, and that of the hybrid CNN-CLPB model attains 88.16%. By the traditional machine learning model LBP, the recognition rates of the three emotions of anger, sadness, and fear are 32.31%, 28.71%, and 24.08%, respectively. In contrast, the single LeNet-5 model achieves a higher recognition rate for each expression than the traditional machine learning LBP model. On the whole, the hybrid CNN-CLPB model has a significant advantage in the recognition rate. Therefore, the hybrid CNN-CLPB model has the best effect on facial emotion recognition among the comparative models.

Figure 5(b) displays the comparison of the mean values and maximum values of facial emotion recognition for each expression by the hybrid CNN-CLPB model, the traditional machine learning LBP model, the LeNet-5 model, and the VGGNet model.

In Figure 5(b), the traditional machine learning LBP model has the highest recognition rate for the facial emotion of surprise, reaching 73.2%, and the LeNet-5 model has the highest recognition rate for happiness, reaching 86.77%. Moreover, the VGGNet model has the highest recognition rate for happiness of 94.11%. Ultimately, the hybrid CNN-CLPB model also has the highest recognition rate for happiness, reaching 98.61%.

4.2. Quantitative Analysis of the CNN-CLBP Model Based on Feature Fusion. The accuracy, cross-entropy, and loss function of the hybrid CNN-CLBP model on the training set and the test set are shown in Figure 6.

According to the analysis of the data changes in Figure 6, the accuracy rate trend of the CNN-CLBP model on the training set is almost identical to that on the test set, and the entire training process is stable. When the corresponding training steps are less than 2k, the accuracy rate of the model on the test set shows an exponential changing trend. When the corresponding training steps are within the range of 4k–8k, the accuracy rate of the model on the test set still shows relatively rapid growth. When the training steps fluctuate around 8k, the accuracy rate of the model on the test set reaches 96.8%, and then, it tends to be stable. At this time, the recognition rate of facial emotions is extremely high. There is no obvious deviation from the changing trend of cross-entropy of the training set to that of the test set. The prediction distribution of the model on the test set is closer to the actual value. When the training steps are around 8k, the corresponding value of the cross-entropy shows a small reduction. Then, the corresponding value of cross-entropy is continuously decreasing as the training step increases. For the changes in the loss function, when the training step increases, the function value of the model on the test set continues to decrease. When the number of iterations reaches 20k, the value of the function gradually stabilizes. In this quantitative analysis, both the training set and the test are set to cover all participants in the data set, but the images are chosen randomly.

4.3. Facial Emotion Recognition Analysis of the CNN-CLBP Model. Based on the CK + data set, the hybrid CNN-CLBP model recognizes typical facial expressions on face images, including anger, disgust, fear, happiness, sadness, and surprise. The results are shown in Figure 7. The facial emotion recognition effects of the hybrid CNN-CLBP model are presented in Figure 8.

Table 3: Parameter settings of LeNet-5 and VGGNet models.

Network layer		Input size	Convolution kernel size	Output size
Lenet-5	Convolution layer 1	32∗32∗1	5∗5∗1	28∗28∗4
	Lower sampling layer 1	28∗28∗4	2∗2	14∗14∗4
	Convolution layer 2	14∗14∗4	5∗5∗6	10∗10∗14
	Lower sampling layer 2	10∗10∗14	2∗2	5∗5∗14
	Convolution layer 3	5∗5∗14	5∗5∗14	1∗1∗120
	Full connection layer	1∗1∗120	120∗82	1∗1∗82
	Output layer	1∗1∗82	82∗10	1∗1∗10
VGGNet	Convolution layer 1	224∗224∗3	11∗11∗3	55∗55∗46
	Lower sampling layer 1	55∗55∗46	3∗3	27∗27∗46
	Convolution layer 2	27∗27∗46	5∗5∗46	27∗27∗128
	Lower sampling layer 2	27∗27∗126	3∗3	13∗13∗128
	Convolution layer 3	13∗13∗126	3∗3∗256	13∗13∗192
	Convolution layer 4	13∗13∗192	3∗3∗192	13∗13∗192
	Convolution layer 5	13∗13∗192	3∗3∗192	13∗13∗128

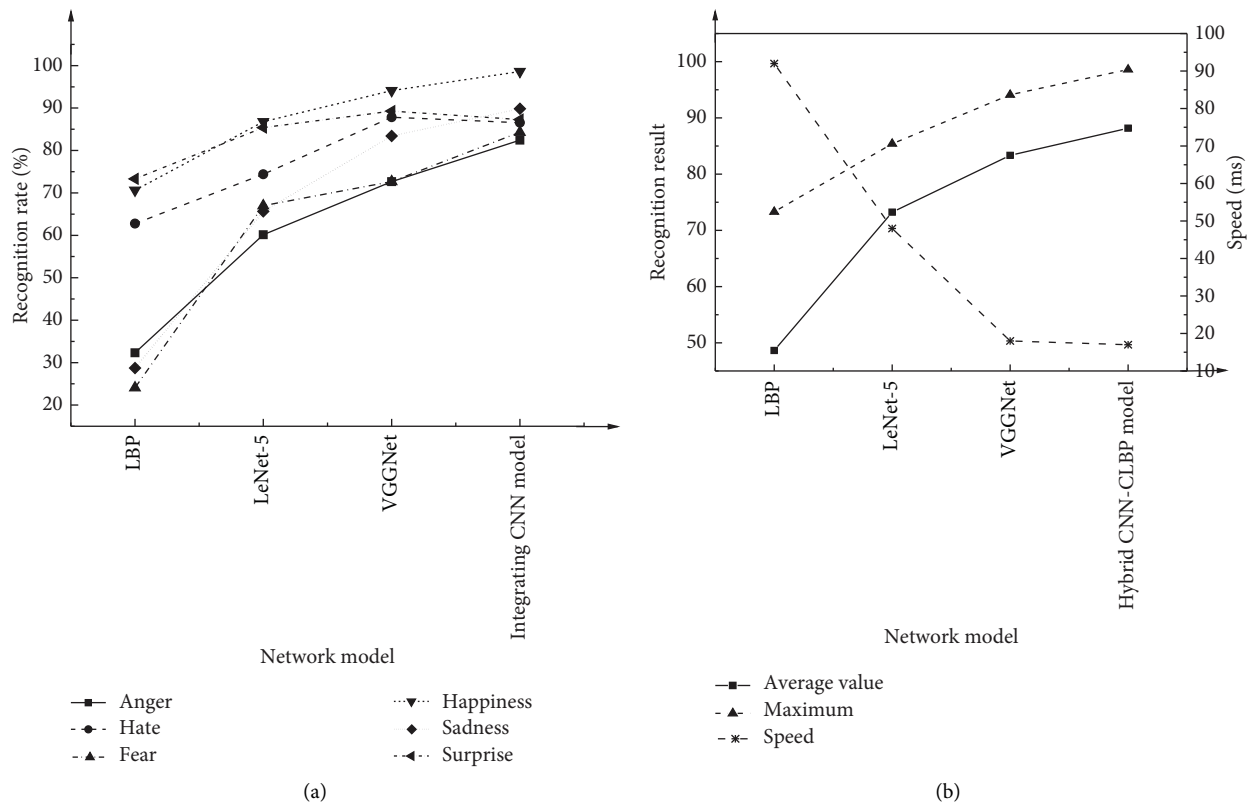


Figure 5: Comparison of recognition results of several facial emotion recognition methods: (a) recognition rate; (b) recognition results and time consumption.

Through Figure 7, the hybrid CNN-CLBP model provides a higher recognition rate of the expression of happiness and surprise than of the other expressions. Moreover, the average recognition rate of the facial expression of sadness is the lowest, reaching 77.46%. The probability of erroneous judgment as sadness is 9.50%, and the probability of erroneous judgment as disgust is 6.18%.

5. Discussion

Different from psychologically and physiologically mature adults, juveniles are in a critical developing period of physical and mental maturity. Hence, it is vital and essential to pay attention to juveniles, especially the psychology of juveniles involved in crime. Emotion recognition is a critical method based on psychoanalysis. In summary, the recognition rate of facial expressions by the traditional machine learning LBP model is significantly different from the current recognition algorithms for facial expressions. The reason is that the operator of the traditional machine learning LBP model is a matching algorithm focusing on analyzing the differences between different categories, thereby leading to the loss of detailed information on the image.

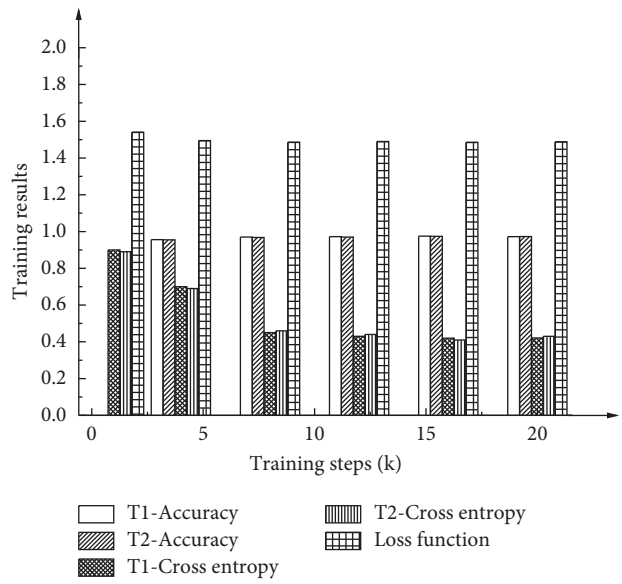


Figure 6: Quantitative analysis results of the CNN-CLBP model based on feature fusion.

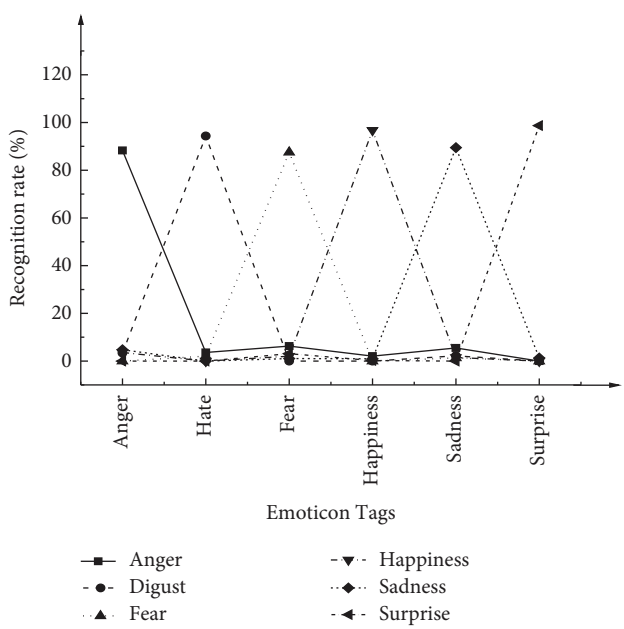


Figure 7: Facial emotion recognition results of the hybrid CNN-CLBP model.

However, facial expressions are precise reflections of the details. The image contains complex and nuance information, so the traditional machine learning LBP model has a low effect on recognizing facial expressions. The hybrid CNN-CLBP model significantly ameliorates the recognition accuracy of facial expressions. The reason is that more convolutional layers in the model can greatly improve its ability to indicate image features and reflect the detailed information on the image more accurately.

There are many research results of the application of deep learning to facial expression recognition, including the application of CNN and LBP, as shown in Table 4.

In facial emotion recognition, the hybrid CNN-CLBP model effectively avoids the problem that traditional models are easily affected by the number of network layers. Facial emotion recognition is affected by many problems caused by image rotation. The effective fusion of CNN features and CLBP features deals with such problems easily. In addition to the VGGNet model, the recognition accuracy of hybrid CNN-CLBP model is significantly better than other classical models, which may be related to the number and structure of the convolutional layers in the network. As a result, the hybrid CNN-CLBP model has achieved marvelous applicability to recognizing facial expressions, providing satisfying results of the classification of facial emotions and expressions. Meanwhile, the psychological changes of juveniles involved in crimes are also a manifestation of emotional changes closely related to their facial expressions. The hybrid multilayer CNN-CLBP model shows excellent performance in recognition of facial emotions. It provides an important development direction for analyzing psychological changes and the extraction of psychological characteristics of juveniles involved in crimes. In the recognition process of facial expressions and emotions, the training set and the test set of the hybrid CNN-CLBP model show a similar changing trend, and there is no significant difference between the changing trends of test accuracy rate and the training accuracy rate. In other words, poor-fitting never occurs, which shows that the preprocessing of face images is sufficient, and the optimization of neural network model parameters further ensures the stability of the training process. In contrast, LBP is improved by this work to expand the application of deep learning method for facial expression recognition.

In summary, it is obvious that the combination of LBP and CNN has achieved some research results in facial expression recognition, but the fusion of CNN and CLBP in facial expression recognition is still relatively scarce, which is an innovative trait of this paper. Furthermore, the hybrid CNN-CLBP model is applied here to identify facial expressions, which further expands the application field of the deep hybrid model. The effectiveness of the overall emotion recognition can be confirmed through the preprocessing and recognition of the images.

From the perspective of facial expression recognition, the hybrid CNN-CLBP model has a high recognition accuracy for the two expressions of happiness and surprise, and the probability of erroneous judgment as the expression of sadness is large. This indicates that the emotional expression of sadness in the pixel space cannot be effectively separated from other expressions. The reason may be that changes in expressions such as happiness and surprise have more significant characteristics. The amplitude of such expression changes is more extensive, which is easier to distinguish and identify. Besides, the emotional performance of happiness and surprise is different from other types of emotions. Since the connection between them is not apparent, the model shows higher recognition rates for these two emotions. In addition to the weaker changes, the expression of sadness is more likely to come up with expressions of negative emotions such as disgust and fear. Under the premise of such an obvious correlation and high

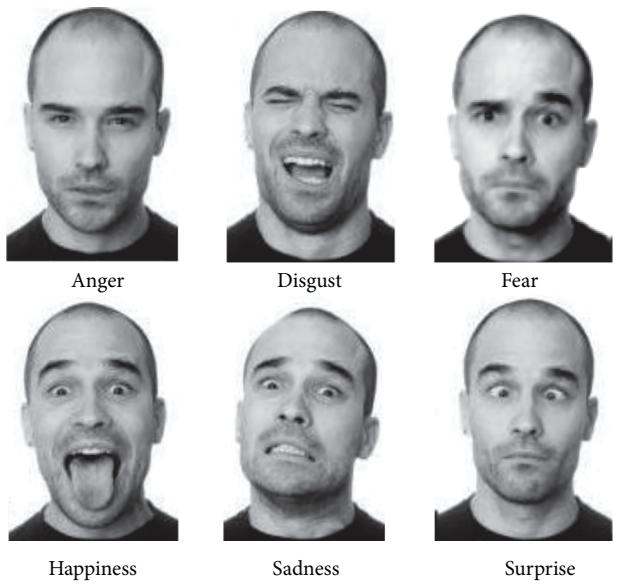


Figure 8: Facial emotion recognition effects of the hybrid CNN-CLBP model (the picture material comes from the public face recognition data set on the web page).

Table 4: Comparison of research based on algorithm recognition.

Author	Primary research contents
Jain et al. [36]	They constructed a facial expression recognition system based on a single deep CNN, including a convolution layer and a deep residual layer. Through the training of face image labels and the training of CK + dataset and JAFFE dataset, they found that the model with deep convolution layer had better recognition effect and accuracy than the traditional emotion recognition methods.
Ma and Celik [37]	They proposed a densely connected CNN structure applicable to facial expression recognition, and through this structure, the output and input of adjacent convolution layers were connected. They finally verified the effectiveness of the structure in facial expression recognition.
These works provide solid support for the research work reported here and reveal the effectiveness and necessity of this research work.	
Liu et al. [38]	They applied the fused CNN and CLBP to intelligent mining. By considering the unique visual features, they found that, in the case of applying the fused model, the accuracy of image recognition could be improved by 2% to 3% compared with the traditional methods.
Shao and Qian [34]	They proposed a two-branch CNN model. By extracting traditional LBP features and deep learning features, they found that the fusion of LBP features and CNN showed excellent performance and applicability in facial expression recognition.
Takalkar et al. [39]	They combined LBP with CNN for microexpression recognition. Through the evaluation of seven widely used microexpression databases, they found that the recognition accuracy of the proposed method had been significantly improved, and the relevant training and testing could be realized through a small number of data sets.
In contrast, this work further optimizes LBP and extends the application of deep learning method to facial expression recognition.	
Liu and Zhang [40]	They found that the recognition accuracy of CNN model could reach 78.9% after analyzing the accuracy of deep neural network in image recognition on CK + data set.
Shahid et al. [41]	They proposed a multiclass SVM and topic-related k-fold crossover method for facial expression recognition. They found that the recognition rate of this method for CK + data set could reach more than 90%, and the accuracy and calculation time were improved.
Liao et al. [42]	They introduced the conditional random forest structure to build a deep multi-instance learning model. They found that, in the field of automatic facial expression recognition, the recognition rate of the model on CK + public data set could reach more than 86%.
Miyoshi et al. [43]	They proposed an enhanced convolutional long short-term memory algorithm for automatic facial expression recognition. The test results showed that the recognition accuracy of this algorithm on CK + data set was more than 85%.
Hybrid CNN-CLBP algorithm reported here	In conclusion, although the recognition rate of the hybrid CNN-CLBP algorithm is lower than that proposed in [41], the average recognition rate for different types of expressions is still more than 88%, which is better than the current advanced algorithms. At the same time, this algorithm also considers the relationship between expression recognition and psychological analysis, which is different from other studies.

similarity, it is possible to obtain a wrong recognition result. Nevertheless, for the particular group of juveniles, the change in their psychological characteristics is closely related to facial expressions. Thus, understanding their emotions through facial expressions is crucial to exploring their psychological characteristics.

Intelligent learning methods like deep learning have great potential to explore the increasingly severe psychological problems that people are experiencing at all stages. Jiang et al. analyzed the body mass index (BMI) and realized the visual estimation of BMI based on machine learning and computer vision, thus revealing the applicability of deep learning in BMI estimation [44]. Ashiquzzaman et al. proposed an optimized deep learning convolutional neural network (DCNN). By introducing the spatial pyramid layer, the authors maintained the input dimension of DCNN at a fixed level, with faster computing speed and good performance in gesture detection and recognition. The above works laid the foundation for the application of human-computer interaction based on gesture input [44]. The potential psychological conditions or psychological problems are tapped from the perspective of facial emotion recognition here [45–48]. This method is not perfectly accurate, but emotion is an externalized expression of people's psychological states, so analyzing the psychological states through emotion is a feasible way [49–52]. This is also consistent with the above research results [53]. Besides, paying attention to juveniles' psychological states, especially those involved in crimes, and giving appropriate psychological guidance and method guidance have a positive role in helping juveniles establish correct outlooks on life and values [54–57]. The hybrid CNN-CLBP model reported here achieves excellent performance in facial expression recognition and can precisely capture different facial expressions. This is beneficial to judging changes in the psychological states of juveniles via facial expression recognition. In addition, the combination of emotion recognition through facial expressions with psychological analysis also provides a possible direction for extracting and analyzing psychological features.

6. Conclusions

By applying the hybrid CNN-CLBP model to facial emotion recognition, the essential conclusions are drawn as follows. Increasing the number of network layers of CNN can effectively improve expression recognition accuracy. Besides, image preprocessing and parameter optimization have a significant effect on enhancing the effectiveness of the model. Moreover, the hybrid CNN-CLBP model shows quite exceptional results in the fusion of facial image features and has the best accuracy in recognizing the facial expressions of happiness and surprise. The emotion recognition depending on facial expression and guided by the connection between emotion and psychology provides a direction for the simple analysis of psychological characteristics of juvenile criminals and initially combining deep learning methods with psychoanalysis.

However, several shortcomings have been identified. The hybrid CNN-CLBP model can extract relevant feature data onto expressions, but it cannot investigate particular psychological characteristics. Therefore, in the future, psychological counseling and other elements will be included to analyze and explore the psychological characteristics of juvenile criminals in a profound and detailed manner.

Data Availability

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Consent

Informed consent was obtained from all individual participants included in the study.

Conflicts of Interest

All authors declare that they have no conflicts of interest.

Authors' Contributions

All authors listed have made a substantial, direct, and intellectual contribution to the work and approved it for publication.

Acknowledgments

This work was supported by Zhejiang Provincial Social Science Fund (18NDYD46YB).

References

[1] R. W. McCormic, A. M. Pomerantz, E. Ro, and D. J. Segrist, "The "me too" decision: an analog study of therapist self-disclosure of psychological problems," *Journal of Clinical Psychology*, vol. 75, no. 4, pp. 794–800, 2019.

[2] D. Kim, H. J. Shin, S. W. Kim, J. M. Hong, K. S. Lee, and S. H. Lee, "Psychological problems of pneumothorax according to resilience, stress, and post-traumatic stress," *Psychiatry Investigation*, vol. 14, no. 6, pp. 795–800, 2017.

[3] T. Young, D. Hazarika, S. Poria, and E. Cambria, "Recent trends in deep learning based natural language processing [review article]," *IEEE Computational Intelligence Magazine*, vol. 13, no. 3, pp. 55–75, 2018.

[4] N. Fleming, "How artificial intelligence is changing drug discovery," *Nature*, vol. 557, no. 7707, pp. 55–S57, 2018.

[5] V. D. Badal, P. J. Kundrotas, and I. A. Vakser, "Natural language processing in text mining for structural modeling of protein complexes," *BMC Bioinformatics*, vol. 19, no. 1, pp. 84–10, 2018.

[6] Y. L. Huang, S. H. Chen, and H. H. Tseng, "Attachment avoidance and fearful prosodic emotion recognition predict depression maintenance," *Psychiatry Research*, vol. 272, pp. 649–654, 2019.

[7] L. Singh, S. Singh, and N. Aggarwal, "Improved TOPSIS method for peak frame selection in audio-video human emotion recognition," *Multimedia Tools and Applications*, vol. 78, no. 5, pp. 6277–6308, 2019.

[8] X. Fan and T. Tjahjadi, "Fusing dynamic deep learned features and handcrafted features for facial expression recognition,"

Journal of Visual Communication and Image Representation, vol. 65, Article ID 102659, 2019.

[9] G. Viswanatha Reddy, C. Dharma Savarni, and S. Mukherjee, "Facial expression recognition in the wild, by fusion of deep learnt and hand-crafted features," *Cognitive Systems Research*, vol. 62, pp. 23–34, 2020.

[10] X. Liang, L. Xu, J. Liu et al., "Patch attention layer of embedding handcrafted features in cnn for facial expression recognition," *Sensors*, vol. 21, no. 3, p. 833, 2021.

[11] N. Jain, S. Kumar, A. Kumar, P. Shamsolmoali, and M. Zareapoor, "Hybrid deep neural networks for face emotion recognition," *Pattern Recognition Letters*, vol. 115, pp. 101–106, 2018.

[12] E. Avots, T. Sapiński, M. Bachmann, and D. Kamińska, "Audiovisual emotion recognition in wild," *Machine Vision and Applications*, vol. 30, no. 5, pp. 975–985, 2019.

[13] Y. K. Li, X. J. Wu, and J. Kittler, "L1-2D2PCANet: a deep learning network for face recognition," *Journal of Electronic Imaging*, vol. 28, no. 02, Article ID 023016, 1 page, 2019.

[14] B. Kratzwald, S. Ilić, M. Kraus, S. Feuerriegel, and H. Prendinger, "Deep learning for affective computing: text-based emotion recognition in decision support," *Decision Support Systems*, vol. 115, pp. 24–35, 2018.

[15] R. K. Kumar, J. Garain, D. R. Kisku, and G. Sanyal, "Estimating attention of faces due to its growing level of emotions," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, Salt Lake City, UT, USA, 2018.

[16] S. Mishra, G. R. B. Prasada, R. K. Kumar, and G. Sanyal, "Emotion recognition through facial gestures - a deep learning approach," in *Proceedings of the International Conference on Mining Intelligence and Knowledge Exploration,* Springer, Cham, 2017.

[17] F. Zhang, N. Cai, J. Wu, G. Cen, H. Wang, and X. Chen, "Image denoising method based on a deep convolution neural network," *IET Image Processing*, vol. 12, no. 4, pp. 485–493, 2018.

[18] C. Min, G. Wen, Z. Yang, X. Li, and B. Li, "Non-intrusive load monitoring system based on convolution neural network and adaptive linear programming boosting," *Energies*, vol. 12, no. 15, p. 2882, 2019.

[19] H. Gao, H. Hu, Y. Zhao, and J. Li, "A real-time fiber mode demodulation method enhanced by convolution neural network," *Optical Fiber Technology*, vol. 50, pp. 139–144, 2019.

[20] Y. D. Zhang, C. Pan, X. Chen, and F. Wang, "Abnormal breast identification by nine-layer convolutional neural network with parametric rectified linear unit and rank-based stochastic pooling," *Journal of computational science*, vol. 27, pp. 57–68, 2018.

[21] G. F. Zou, G. X. Fu, M. L. Gao, J. Shen, L. J. Yin, and X. Y. Ben, "A novel construction method of convolutional neural network model based on data-driven," *Multimedia Tools and Applications*, vol. 78, no. 6, pp. 6969–6987, 2019.

[22] F. Kong, "Facial expression recognition method based on deep convolutional neural network combined with improved LBP features," *Personal and Ubiquitous Computing*, vol. 23, no. 3-4, pp. 531–539, 2019.

[23] M. Hammad and K. Wang, "Parallel score fusion of ECG and fingerprint for human authentication based on convolution neural network," *Computers & Security*, vol. 81, pp. 107–122, 2019.

[24] D. Pouliot, R. Latifovic, J. Pasher, and J. Duffe, "Landsat super-resolution enhancement using convolution neural networks and Sentinel-2 for training," *Remote Sensing*, vol. 10, no. 3, p. 394, 2018.

[25] X. Li, Q. Ding, and J. Q. Sun, "Remaining useful life estimation in prognostics using deep convolution neural networks," *Reliability Engineering & System Safety*, vol. 172, pp. 1–11, 2018.

[26] Q. Yao, R. Wang, X. Fan, J. Liu, and Y. Li, "Multi-class arrhythmia detection from 12-lead varied-length ECG using attention-based time-incremental convolutional neural network," *Information Fusion*, vol. 53, pp. 174–182, 2020.

[27] X. Jiang, B. Hu, S. Chandra Satapathy, S. H. Wang, and Y. D. Zhang, "Fingerspelling Identification for Chinese Sign Language via AlexNet-Based Transfer Learning and Adam Optimizer," *Scientific Programming*, vol. 2020, Article ID 3291426, 13 pages, 2020.

[28] T. Shanthi and R. S. Sabeenian, "Modified Alexnet architecture for classification of diabetic retinopathy images," *Computers & Electrical Engineering*, vol. 76, pp. 56–64, 2019.

[29] L. Wan, Y. Chen, H. Li, and C. Li, "Rolling-element bearing fault diagnosis using improved LeNet-5 network," *Sensors*, vol. 20, no. 6, 2020.

[30] A. Porebski, V. T. Hoang, N. Vandenbroucke, and D. Hamad, "Multi-color space local binary pattern-based feature selection for texture classification," *Journal of Electronic Imaging*, vol. 27, no. 01, 2018.

[31] J. Lu, W. S. Lee, H. Gan, and X. Hu, "Immature citrus fruit detection based on local binary pattern feature and hierarchical contour analysis," *Biosystems Engineering*, vol. 171, pp. 78–90, 2018.

[32] T. Kalsum, S. M. Anwar, M. Majid, B. Khan, and S. M. Ali, "Emotion recognition from facial expressions using hybrid feature descriptors," *IET Image Processing*, vol. 12, no. 6, pp. 1004–1012, 2018.

[33] V. V. Vu, H. Q. Do, V. T. Dang, and N. T. Do, "An efficient density-based clustering with side information and active learning: a case study for facial expression recognition task," *Intelligent Data Analysis*, vol. 23, no. 1, pp. 227–240, 2019.

[34] J. Shao and Y. Qian, "Three convolutional neural network models for facial expression recognition in the wild," *Neurocomputing*, vol. 355, pp. 82–92, 2019.

[35] C. Finlay and A. M. Oberman, "Approximate homogenization of fully nonlinear elliptic PDEs: estimates and numerical results for Pucci type equations," *Journal of Scientific Computing*, vol. 77, no. 2, pp. 936–949, 2018.

[36] D. K. Jain, P. Shamsolmoali, and P. Sehdev, "Extended deep neural network for facial emotion recognition," *Pattern Recognition Letters*, vol. 120, pp. 69–74, 2019.

[37] H. Ma and T. Celik, "FER-Net: facial expression recognition using densely connected convolutional network," *Electronics Letters*, vol. 55, no. 4, pp. 184–186, 2019.

[38] X. Liu, W. Jing, M. Zhou, and Y. Li, "Multi-scale feature fusion for coal-rock recognition based on completed local binary pattern and convolution neural network," *Entropy*, vol. 21, no. 6, 2019.

[39] M. A. Takalkar, M. Xu, and Z. Chaczko, "Manifold feature integration for micro-expression recognition," *Multimedia Systems*, vol. 26, no. 5, pp. 535–551, 2020.

[40] H. Liu and L. Zhang, "Advancing ensemble learning performance through data transformation and classifiers fusion in granular computing context," *Expert Systems with Applications*, vol. 131, pp. 20–29, 2019.

[41] A. Raza Shahid, S. Khan, and H. Yan, "Contour and region harmonic features for sub-local facial expression recognition,"

Journal of Visual Communication and Image Representation, vol. 73, Article ID 102949, 2020.

[42] H. Liao, D. Wang, P. Fan, and L. Ding, "Deep learning enhanced attributes conditional random forest for robust facial expression recognition," *Multimedia Tools and Applications*, vol. 80, no. 19, pp. 28627–28645, 2021.

[43] R. Miyoshi, N. Nagata, and M. Hashimoto, "Enhanced convolutional LSTM with spatial and temporal skip connections and temporal gates for facial expression recognition from video," *Neural Computing & Applications*, vol. 33, no. 13, pp. 7381–7392, 2021.

[44] M. Jiang, G. Guo, and G. Mu, "Visual BMI estimation from face images using a label distribution based method," *Computer Vision and Image Understanding*, vol. 197, Article ID 102985, 2020.

[45] S. Wang, H. Sheng, D. Yang, Y. Zhang, Y. Wu, and S. Wang, "Extendable multiple nodes recurrent tracking framework with RTU++," *IEEE Transactions on Image Processing, Advance online publication*, vol. 31, 2022.

[46] H. Sheng, R. Cong, D. Yang, R. Chen, S. Wang, and Z. Cui, "UrbanLF: a comprehensive light field dataset for semantic segmentation of urban scenes," *IEEE Transactions on Circuits and Systems for Video Technology, advance online publication*, vol. 34, 2022.

[47] Z. Lv, D. Chen, H. Feng, W. Wei, and H. Lv, "Artificial intelligence in underwater digital twins sensor networks," *ACM Transactions on Sensor Networks*, vol. 18, no. 3, pp. 1–27, 2022.

[48] B. Cao, Y. Gu, Z. Lv, S. Yang, J. Zhao, and Y. Li, "RFID reader anticollision based on distributed parallel particle swarm optimization," *IEEE Internet of Things Journal*, vol. 8, no. 5, pp. 3099–3107, 2021.

[49] Z. Lv, D. Chen, H. Feng, H. Zhu, and H. Lv, "Digital twins in unmanned aerial vehicles for rapid medical resource delivery in epidemics," *IEEE Transactions on Intelligent Transportation Systems, advance online publication*, vol. 45, pp. 1–9, 2021.

[50] R. Liu, X. Wang, H. Lu et al., "SCCGAN: style and characters inpainting based on CGAN," *Mobile Networks and Applications*, vol. 26, no. 1, pp. 3–12, 2021.

[51] G. Sun, Y. Cong, Q. Wang, B. Zhong, and Y. Fu, "Representative task self-selection for flexible clustered lifelong learning," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 33, no. 4, p. 1481, 2022.

[52] F. Liu, G. Zhang, and J. Lu, "Multisource heterogeneous unsupervised domain adaptation via fuzzy relation neural networks," *IEEE Transactions on Fuzzy Systems*, vol. 29, no. 11, pp. 3308–3322, 2021.

[53] Y. Feng, B. Zhang, Y. Liu et al., "A 200–225-GHz manifold-coupled multiplexer utilizing metal waveguides," *IEEE Transactions on Microwave Theory and Techniques*, vol. 69, no. 12, pp. 5327–5333, 2021.

[54] W. Zheng, X. Liu, and L. Yin, "Research on image classification method based on improved multi-scale relational network," *PeerJ Computer Science*, vol. 7, Article ID e613, 2020.

[55] B. Cao, J. Zhao, Z. Lv, and P. Yang, "Diversified personalized recommendation optimization based on mobile data," *IEEE Transactions on Intelligent Transportation Systems*, vol. 22, no. 4, pp. 2133–2139, 2021.

[56] H. Zhu, M. Xue, Y. Wang, G. Yuan, and X. Li, "Fast visual tracking with siamese oriented region proposal network," *IEEE Signal Processing Letters*, vol. 29, pp. 1437–1441, 2022.

[57] J. Li, L. Han, C. Zhang, Q. Li, and Z. Liu, "Spherical convolution empowered viewport prediction in 360 video multicast with limited FoV feedback," *ACM Transactions on Multimedia Computing, Communications, and Applications, advance online publication*, vol. 65, 2022.

Hindawi
Computational Intelligence and Neuroscience
Volume 2022, Article ID 3123448, 9 pages
https://doi.org/10.1155/2022/3123448

Research Article

Construction and Application Research of the Visual Image Obstacle Type Recognition Model Based on the Computer-Expanded Convolutional Neural Network

Yuchen Xian

School of Software Technology, Dalian University of Technology, Dalian 116000, Liaoning, China

Correspondence should be addressed to Yuchen Xian; xianyuchen528@mail.dlut.edu.cn

Received 5 July 2022; Revised 16 August 2022; Accepted 25 August 2022; Published 21 September 2022

Academic Editor: Ning Cao

Due to the development of computer vision technology and image processing technology, obstacle recognition technology has been widely used in military and scientific research fields. However, most of the existing image-based recognition technologies are easily affected by environmental factors, which makes the application scenario of this system more fixed and cannot be applied in complex environments. This paper mainly focuses on the traditional obstacle detection and type recognition method recognition accuracy, reliability and universality is difficult to meet the technical requirements of intelligent vehicles and unmanned vehicles, traditional detection equipment cost is expensive, and other problems. There are many traditional obstacle detection methods, which basically start from the color, edge, and other information of the target object to do detection and recognition research, but their recognition accuracy, reliability, and universality are difficult to meet the technical requirements of intelligent vehicles and unmanned vehicles, and the detection equipment is expensive. The dilated convolutional neural network has the ability to learn autonomously, using the original image as input, without the cumbersome preprocessing process and can extract features of the target object one by one to achieve more accurate recognition. This design will be based on the expanded convolutional neural network, design an obstacle type detection and obstacle recognition application with high recognition accuracy, and good generalization, in which this paper applies the hierarchical structure of the expanded convolutional neural network weight sharing to learn the characteristics of various types of obstacles and extract the global features with characterization significance, combined with the ROI algorithm to achieve real-time obstacle detection and high accuracy type recognition. The ROI algorithm is combined to achieve real-time obstacle detection and high-precision type recognition.

1. Introduction

In recent years, self-driving cars and autonomous mobile robots have received increasing attention from science and technology, and obstacle detection and recognition is the key to their development [1].

For example, in the military, it can replace personnel for reconnaissance, patrol, search, rescue, and other tasks to reduce unnecessary casualties on the battlefield; in civil use, it can realize automatic obstacle avoidance and unmanned driving of cars and also serve as daily traffic identification and navigation for people with visual impairment [2]; in scientific research, this technology can also be used for high-risk exploration work such as deep sea and space exploration, and for various scientific experiments to collect important information. It is foreseeable that on the basis of driverless vehicles, autonomous mobile robots and a variety of autonomous navigation systems, significant changes will take place in the future. But even so, it is a difficult technical challenge to solve [3].

Obstacle recognition technology is the advanced technology and research results of computer vision, pattern recognition, machine learning, artificial intelligence, and other multidisciplinary and multidisciplinary fields, which is the inevitable result of technological development. Humans can identify and avoid obstacles through vision, brain, and body movements during travel. Firstly, visualization methods are used to obtain traffic conditions in the direction

of vehicle travel; on this basis, the brain identifies and judges the obstacles in the line of sight through the traffic information transmitted by vision and feeds them back to the limbs; in order to achieve obstacle avoidance, the limbs will make different tendency avoidance movements according to the signals from the brain (see Figure 1).

So-called obstacle recognition is the use of computers or other computing processors and image acquisition devices to simulate the automatic recognition and feedback of any obstacle by the human brain and eyes. The recognition system has two basic elements: one is the machine vision (camera or other image capture device) and the other is the computing processor (similar to the human brain). If control elements are added, functions such as obstacle avoidance for human bodies, such as cars and automatic mobile robots, can be implemented, which are based on detection, recognition, and control. How to improve the adaptability and real-time performance of obstacles in complex environments has received increasing attention from researchers [4].

We know that 80% of human information is obtained from vision and 20% is obtained through hearing. At the same time, according to relevant studies, about one-fifth of the human population is active in image processing, so vision is very important when driving a car. In addition, machine vision-based environment perception technology is a nonintrusive method of information acquisition, which can achieve zero pollution of the road environment. Therefore, the development of vision-based obstacle detection and recognition technology will be a pioneer in the field of intelligent control of automobiles [5].

Extended convolutional neural network is an important branch in current computer vision research, which has a wide range of applications in image classification, target detection, target tracking, pose estimation, motion recognition, and scene labeling. Obstacle detection and recognition are object detection and tracking, while object distance and movement direction are estimates of object behavior and pose, and image classification and scene labeling can classify and label images. In conclusion, the extended convolutional neural network is an effective method for obstacle detection and recognition. In image recognition, the widely used convolutional neural network with multi-level feature extraction is used to detect and recognize obstacles more accurately [6].

2. Introduction to Related Technologies

2.1. Convolutional Neural Network Theory. There are many traditional obstacle detection methods, which mainly focus on object color, edge, and other information for recognition and identification, but it is difficult to meet the technical requirements of automobiles and autonomous vehicles due to recognition accuracy, reliability, and universality. In contrast, convolutional neural network is an automatic learning method, which does not need to go through complicated preprocessing and can extract the features of the target step by step, thus achieving higher accuracy. Also, the method requires only a simple test device, such as a PC and an on-board computer [7].

A video capture program can be used. Compared with traditional detection methods, convolutional neural networks have the following advantages :

(1) Hierarchical structure. The hierarchical structure of the convolutional neural network is designed to match the hierarchical distribution of image features.

The hierarchical structure of the convolutional neural network is designed to fit well with the hierarchical distribution of image features. The hierarchical structure of image features is shown in Figure 2.

In adjacent regions, a boundary line is formed by pixel information from multiple regions, and each line along different directions is combined to form a texture of an image, while a local image consists of multiple locations or key points, and the target object can be reproduced by the combination of multiple local graphics. Convolutional neural network is a layered structure consisting of convolutional and pooling layers, which can extract the image in layers and describe the image object in a holistic form [8].

(2) Connectivity. Based on the neighborhood feature similarity of the image itself, the convolutional neural network utilizes local connectivity to perform large-scale sparse of the network. Neighborhood feature similarity can also be understood as the stability of neighborhood features, i.e. [9], the statistical properties of a subregion in the original image have similar statistical properties to several of its neighboring subregions. Convolutional neural networks learn the characteristics of an image by means of local connectivity, which is achieved by combining regions with similar properties, replacing the similar properties of the whole region with one typical feature, and finally combining all the different features to form a complete target object [10].

Using the local connection method can effectively reduce the training parameters of the network. On the basis of $1000 * 1000$, the number of neurons in the next hidden layer is 1000000, and the training parameters of the model are $1000 * 1000 * 1000000 = 10^{12}$ when fully connected; however, when local connections are made between neurons, 100000 hidden layer neurons are connected with $10 * 10$ local regions, the network parameters become $10 * 10 * 1000000 = 10^{8}$, which is a reduction of 4 orders of magnitude in local connectivity compared to full connectivity and has great advantages in practical applications. Especially in larger samples, the learning efficiency of the convolutional neural network is significantly improved compared with the fully connected case [11].

(3) Weight sharing. CNN has a better feature than other neural networks, i.e., weight sharing, and the structure of weight sharing is given. Convolutional kernels of the same type, whose weights and offsets are shared, act on the original image in a series of

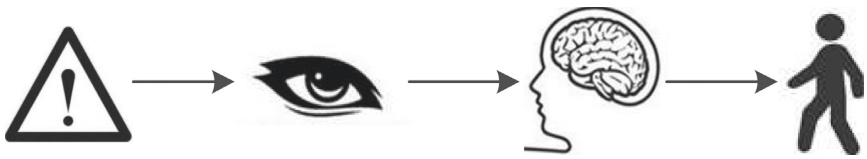


Figure 1: Diagram of the human obstacle avoidance process based on visual perception.

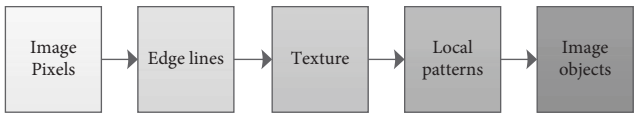


Figure 2: Hierarchical structure of image features.

covariates in sequence and finally obtain specific types of image features. Using multiple convolutional kernels, multiple feature sets can be obtained, and multiple feature sets are combined to form a complete image [12]. Weight sharing and local linking have similar functions, both significantly reduce the complexity of the network and are effective in reducing computation and accelerating learning (see Figure 3).

2.2. Dilated Convolution in Computer Vision. Dilated convolution is a kind of "ATrry convolution" which has been developed into wavelets. In comparison with pure convolutional networks, a new approach is proposed, namely, expanded convolution, which expands the network's receptive area while adding some linear parameters. Pooling operation is another method to increase the field of view (FOV) of the network in an exponential way; however, it loses much background information and requires other operations [13], such as deconvolution or up-sampling, to obtain the same input size. In contrast, extended convolution not only effectively extends the field of view (FOV) of the network but also ensures that the size of the feature map remains unchanged. To achieve this, extended convolution focuses on an application that can effectively integrate more background information into a broader view of the input. For the density prediction problem, Yu and Koltun et al. proposed a new extended convolution-based method for fusing data on multiple scales to improve current semantic segmentation methods. An et al. (39) proposed a "deep lab" system using "anomalous convolution" to control the resolution of feature maps in CNNs, and a new technique to segment semantic images in PASCALVOC2012 [14].

From an intuitive point of view, SRCNN [8] has a perceptual range of 13×13 and achieves an exciting reconstruction accuracy. The "DeepLab" system uses multiple expansion rates to extend the FOV of the network and further improve the task of semantic segmentation. Due to the results of extended convolutional techniques in semantic segmentation, a new extended convolutional neural network (DCNN) is proposed in this paper for obstacle recognition. Although the learning prior process is time-consuming, the obvious advantages of this approach are that it performs only one operation offline and can be efficiently trained in SR systems.

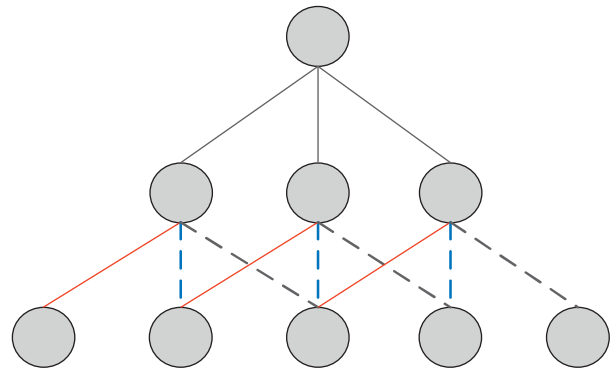


Figure 3: Weight sharing.

2.3. Theory of the ROI Extraction Algorithm. ROI is the abbreviation of region of interest, which means "region of interest," and it was first proposed by the robot in the process of studying the object. The ROI zone, which is the center of human visual attention, is the most attractive and important area of research content [15].

Since the concept of ROI was introduced, ROI-related algorithms have been widely used, and various ROI algorithms have been studied to separate the object of attention from the redundant and useless background and to focus the research on the target object. After removing the irrelevant background information [16], only the ROI region is required, which can both reduce the computation and effectively improve the efficiency of image processing, as well as eliminate the influence of non-ROI regions on the system, thus improving the real-time and accuracy of the system [17].

Researchers do not select ROIs according to certain criteria but rather according to the basic algorithm of ROI, which is due to the subjective and uncertainty of ROI.

The selection of ROI has a subjective element. There is no uniform definition of ROI (ROI) in academia, and it is also difficult to perform model calculations. Also, people have different visual and aesthetic cognitive mechanisms, and when people view the same image, the focus of attention is limited by subjective will, so the delineation of ROI is determined by the user's own perception and is subjective. If a researcher can find several representative features or small objects in a photograph, the photograph can be considered worthy of study and ROI can be extracted from it. However, there are also images with generalized ROI-free regions, such as noise images, which have no value to everyone and from which ROI cannot be extracted [18].

The choice of ROI is uncertain. When selecting ROIs, regular graphics such as positive polygons, square polygons, and rectangles can be used for ROI acquisition, while non-

regular graphics and custom graphics can also be used for intercepting ROIs. But to match the visual habits of users and to make the operation as simple as possible, regular graphics are generally used for ROI acquisition.

3. Application Method Design

3.1. Overall Scheme of Obstacle Detection and Type Recognition. In this paper, we introduce a method based on the widest difference method and morphological operations, the overall scheme is shown in Figure 4, firstly, the extraction of obstacles, and then on this basis, the classification of the target using the widest difference method and morphological operations; the second part is to use DCNN technology to extract the characteristics of the target obstacles, establish a kind of deep convolutional neural network suitable for multiple types of obstacles in the daily environment and combine the RPN network with the RPN network combined with it [19], the feature extraction of the target object and the recommendation of the target area are completed; the third stage is the obstacle recognition, in which the detection of multiple types of obstacles based on the DCNN network is achieved in real-time and the correct recognition results are outputted and calibrated using rectangular boxes in the actual vehicle driving. In particular, it is important to note that the accuracy of obstacle recognition described in the paper refers to the recognition probability of each type of obstacle output by the output layer, that is, the relationship between the properties of the obstacle and the real properties of the target by the artificial neural network [20].

In terms of obstacle feature extraction, the expanded convolutional neural network mainly accomplishes the following tasks:

(1) Extraction of features: the principle of feature extraction is that it can describe the features of the object in the real environment. Thus, the extracted features have the following characteristics: a generic feature that can be used for the same objects, that is, they can represent the same objects; it is the differential feature that distinguishes various objects, that is, the significant difference between the characteristics of such objects and other objects; it is a robust method of describing target objects, that is, the extracted features are highly resistant to interference and can effectively avoid noise and perform target recognition.

(2) By category: target object detection and recognition is based on feature extraction, using object features for classification training and determining the location of obstacles by edge regression to obtain the classification probability and specific location of the object. In this paper, we use the convolutional layer and the pooling layer to extract the target obstacles in layers and obtain the output of positive transmission through the influence of multi-layer network layers; then, the actual value and the expected value are differentiated, and a cost function is set so that the cost function tends to 0. Then, the network is updated once in the reverse direction, so that the network can be maximally close to the expected, and after repeated iterations, finally obtains with good robustness; by inputting the vehicle video into the test sample, the obstacle can be detected and recognized with high accuracy.

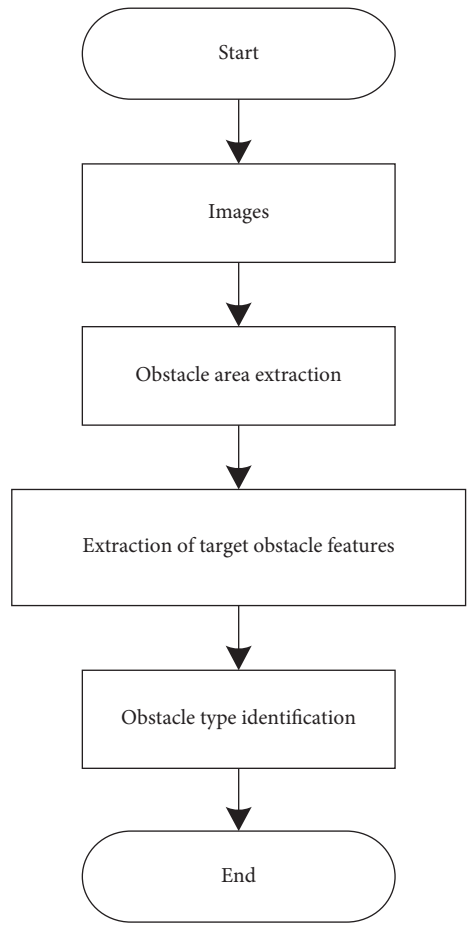


Figure 4: General scheme of obstacle detection and recognition.

3.2. Dilated Convolutional Image Super-Resolution Design. The image super-resolution technique mainly uses IR technology to reconstruct low-resolution images. Here, ILR reduces the resolution of the corresponding HR image IHR. It should be noted that the estimated ISR is in the same dimension as the corresponding IHR and is expected to be highly similar to it. ILR is generated from the IHR, which is obtained by applying a Gaussian filter to the IHR and then downsampling the image by a downsampling r-factor. In general, the image may have C color channels, and thus the ILR can be described as a real-valued tensor of size $H \times W \times C$, and the IHR as $rH \times rW \times C$ [21].

For the super-resolution problem of images, this paper proposes a 7-level extended convolutional neural network (DCNN), which is interpolated by an insertional ILR. Each layer can be described for the proposed DCNN as follows:

$$
\begin{aligned}
f^{1}(W_1, b_1) &= \sigma\left(W_1 * I^{LR} + b_1\right),\\
f^{l}(W_l, b_l) &= \sigma\left(W_l * f^{l-1}(W_{l-1}, b_{l-1}) + b_l\right),\\
f^{5}(W_5, b_5) &= \sigma\left(W_5 * \left(f^{2}(W_2, b_2) \cup f^{4}(W_4, b_4)\right) + b_5\right),\\
f^{6}(W_6, b_6) &= \sigma\left(W_6 * \left(f^{1}(W_1, b_1) \cup f^{5}(W_5, b_5)\right) + b_6\right).
\end{aligned} \tag{1}
$$

In equation (1), l [2–4, 7], $i = \{Wi, bi\}$ are the learnable network weights and offsets, where Wi is the weight of the ith convolutional layer, bi is the deviation term of each layer, and i is [1–7]. The output of layer i in DCNN is fi (Wi, bi). The weighted Wi is the 2D convolution tensor with $ni - 1 \times ni \times ki \times ki$, where ni is the number of feature maps in layer i, ki is the filter size of layer i, $n0 = C$, and bi is the vector with ni.

Figure 5 proposed dilated convolutional neural network (DCNN), which consists of seven convolutional layers with different dilation rates and two jump connections pointed by yellow arrows.

3.3. Image Preprocessing. The currently and commonly used obstacle detection methods are based on machine vision. Obstacle recognition technology based on machine vision has become a hot topic in the field of pattern recognition at present because of its advantages such as large amount of information and large detection distance.

According to the classification of visual characteristics, it can be divided into three categories: monocular vision, binocular vision, and multiocular vision. Binocular vision, also known as stereo vision, can detect the depth and distance of the target, so that the stereo information of the target can be obtained. However, the obstacle recognition algorithm based on binocular vision has defects such as high computational complexity and easy ambiguity in the case of two small visual perception areas; while the detection method of monocular vision is easier to implement, which can be achieved by the two-dimensional information obtained by sensors such as cameras, such as shape, color, and texture,, and then the detection of obstacles is achieved by separating the object from the background region through an image segmentation algorithm. In this paper, a vision-based method is applied to detect and identify obstacles, and the corresponding preprocessing and segmentation steps are given. The flow chart for the detection is shown in Figure 6.

First, a monocular vision sensor is used to acquire visual images, which is a variety of sensors (sonar, laser scanner, line and surface array, COMS camera, CCD camera, digital camera, etc.) to acquire and collect information from the surrounding environment and convert it into a variety of electronic signals that can be recognized by the machine vision system. This thesis focuses on the acquisition of video images using COMS cameras. Secondly, the acquired images are preprocessed accordingly, and the original images are filtered to remove noise such as Gaussian and pretzel while retaining important information, and the contours, edges, and textures in the images are enhanced or emphasized. Finally, by segmenting the target region (obstacle) and background, the target region is extracted and the irrelevant background is eliminated.

3.4. Obstacle Region Extraction. The flow of the algorithm based on the maximum inter-class variance method and combined with morphological operations to automatically extract ROI regions is shown in Figure 7. The specific implementation process of this algorithm is as follows: firstly, the RGB color information of the image is extracted and the components of R, G, and B color system are extracted; then, using makecform and applycform functions, the RGB color space is converted into Lab color space; based on this, the maximum variance method is used to segment the image, and according to the grey scale characteristics of the image, the graythresh function to divide the image; use the extended functions imdilate and imerode to binarise the binarised image in order to eliminate the tiny redundant targets, fill the tiny holes between regions and maintain the integrity of the ROI region as much as possible; use the fill operation to fill the gaps in the ROI region formed by the boundaries, which can be achieved by the function imfill; to remove the background, to maintain the colored information in the ROI area, to maintain the colored information in the ROI area, and to separate the ROI area from the ROI.

4. Experimental Analysis of Application Practice

4.1. Experimental Environment. In this paper, the algorithm is developed on the Windows platform using Matlab2016a, and the experimental environment and the structure of the system are given. Caffe (Caffe) is a general deep learning framework developed by BVLC (Berkeley Vision and Learning Center). The new model can be defined in textual form and its code and model can be opened up for secondary development by researchers in a specific design implementation phase, and it is fast and easy to run. A high-performance discrete graphics card of NVIDIAGeForceGTX980 Ti is chosen for better graphics processing (see Table 1).

4.2. Experimental Process. The whole experimental procedure is as follows.

The first step is to extract the obstacle. On this basis, the ROI region is detected using the maximum variance method and morphological operations, the ROI region is localized using rectangular boundaries, the ROI object region is separated from the original image, and an obstacle dataset containing 20,000 images is generated.

The second step is feature extraction. The images from the obstacle dataset, PASCALVOC 2007 and 2012 databases are applied to the learning and testing of the extended convolutional neural network, and five-level convolution and pooling operations are used to extract the obstacle features; the RPN network and the convolutional neural network together form a convolutional layer, thus providing the network with the obstacle area; ROI pooling is the target recognition and the region recommendation feature map of

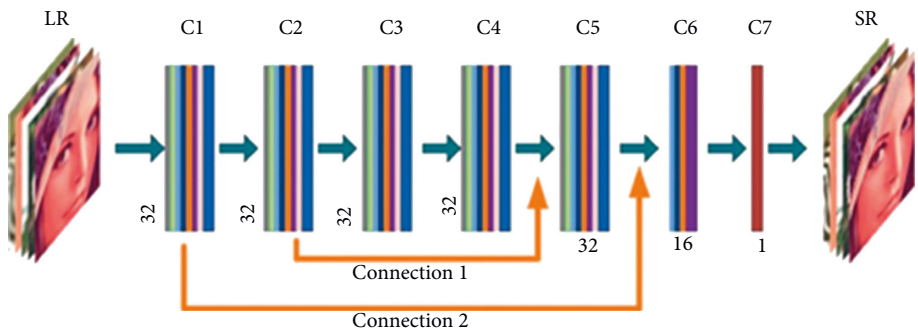


Figure 5: Dilated convolutional neural network (DCNN).

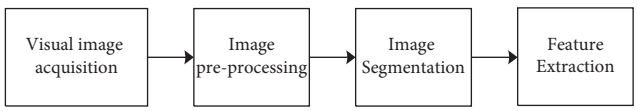


Figure 6: Flow chart of obstacle detection.

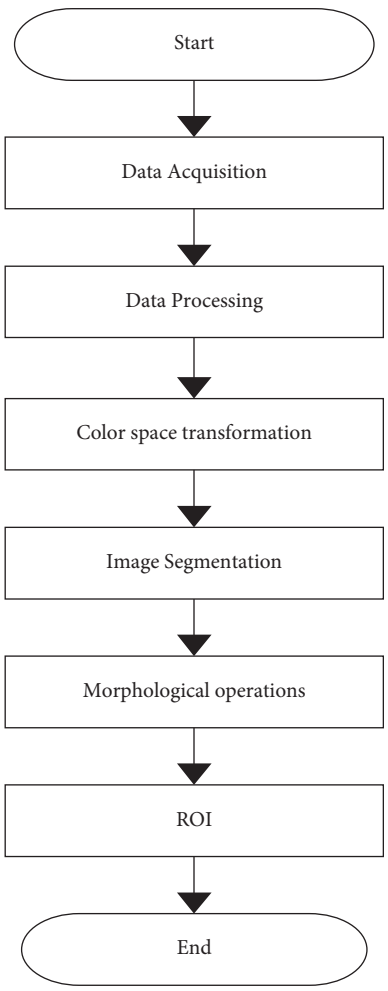


Figure 7: Flow chart of ROI extraction.

size uniformly, which is then added to the subsequent fully-connected and fully-convolutional layers for training to complete a comprehensive extraction of image features.

Table 1: Experimental environment and configuration.

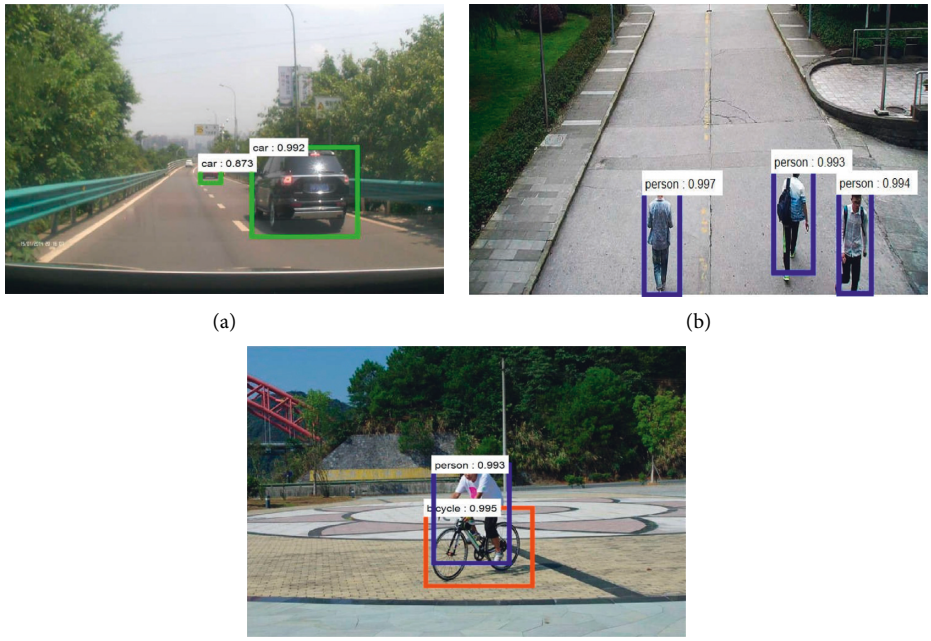
Experimental environment	Environment configuration
Operating system	Windows
CPU	8 nuclear InterI CoreI i7-6700K CPU @ 4.00 GHz
Memory	16 GB
Video card	NVIDIA GeForce GTX 980 Ti
Programming language	Matlab2016a
Database	PASCAL 2007, PASCAL2012
Deep learning framework	Caffe

The third step is to identify obstacles. Five types of obstacles are designed and identified in real-time by BoundingBox, i.e., "Detections" and "Scores."

4.3. Detection Results and Analysis. The proposed multiple types of obstacles were examined by the video of the vehicle measured in the field. In addition, the specific implementation of the software uses the original video as the test sample, but in order to better reflect the results of the test, this paper tests the video in frames and uses a single frame image as an example.

Figure 8 shows the detection and recognition of obstacles in simple traffic conditions under normal light, and it can be seen that the CNN network can detect cars and pedestrians more accurately under normal light conditions. The comparison between the 99.2% accuracy in the case of proximity and the 87.3% accuracy in the case of further distance, shown in Figure 8(a), and shows that the recognition is more accurate in the proximity of the current vehicle than in the distance, with a higher chance of recognition in general. In Figures 8(b) and 8(c), the recognition accuracy for both pedestrians and bicycles is greater than 99.3%, and in general, the DCNN network has higher than 99.3% recognition accuracy for all kinds of obstacles under general lighting conditions.

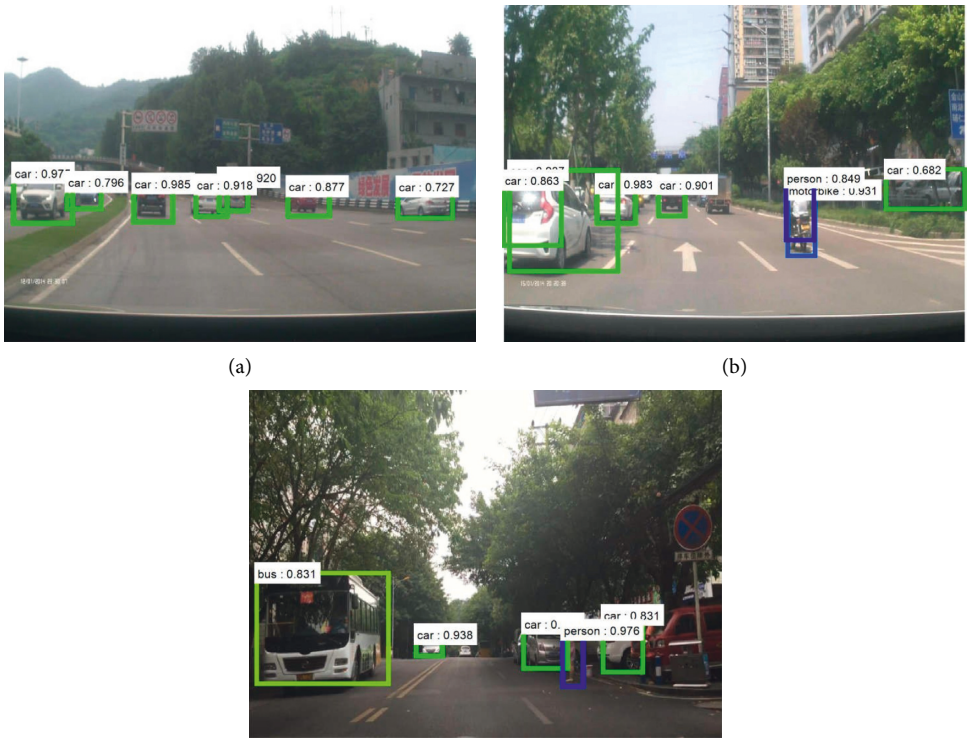
Figure 9 shows obstacle detection and recognition in complex road conditions under normal light conditions. From this figure, it can be seen that when there are various obstacles in the vehicle driving environment, such as cars, people, motorcycles, and buses, the deep convolutional neural network established by using this method can detect the obstacles in the vehicle driving environment with a high



(a) (b)

(c)

Figure 8: Obstacle detection and recognition under normal light simple traffic condition. (a) Vehicle inspection. (b) Pedestrian detection. (c) Human-vehicle interaction detection.



(a) (b)

(c)

Figure 9: Obstacle detection and recognition under normal light complex traffic conditions. (a) A wide range of vehicles in the driving environment. (b) Human-vehicle interaction in the driving environment. (c) Human-vehicle interaction in the driving environment.

Table 2: Statistics of detection results under different light and traffic conditions.

Light conditions and traffic conditions	Detection of frames	Number of obstacles	Average recognition accuracy (%)
Normal light simple traffic conditions	12000	15020	95.1
Normal light complex traffic conditions	8000	33840	92.1

recognition rate. With incomplete obstacle information, it can be detected accurately, for example, in Figure 9(a), when there are a large number of vehicles, the DCNN network can identify both the front and side obstacles of the car. For the interaction behaviors in traffic environments such as cars, people, and motorcycles, the DCNN network is able to detect and identify various obstacles with an accuracy of over 60% for all of them and 90% probability for most of them.

The detection of each frame of obstacles performed by the DCNN network used in this thesis is comparable to the time required for the classification of PASCAL VOC, which is around 0.055 seconds, ensuring real-time detection. In general, the algorithm is able to accurately detect and identify various obstacles in real-time with high accuracy, both in simple traffic environments and in complex road conditions.

To verify the effect of this method on obstacle detection under normal light conditions and road conditions, a total of 35,500 detection frames were collected, in which the number of detected obstacles reached 95,920, and the related detection results are shown in Table 2. From Table 2, it can be seen that the accuracy of obstacle recognition with the *D* CNN network under normal light conditions reached 92%, and under the dim environment, its accuracy reached 86%. At the same time, its recognition is better in simple road conditions than in complex situations.

The reliability of the DCNN network with good universality and robustness is verified by obstacle detection under various road conditions. The method can effectively solve the technical requirements of conventional obstacle detection methods in terms of accuracy, reliability, and universality that cannot be adapted to the fields of smart cars and UAVs. Moreover, the test device of this topic has a lower cost compared with the conventional test equipment, which has a great practical value.

5. Conclusion

Obstacle recognition and type recognition are currently the most basic and challenging topics in the field of computer vision, and in China, with the increasing number of cars and the continuous development of smart cars and driverless technology, road traffic safety has become a top priority. Using the weight-sharing hierarchical structure of the expanded convolutional neural network, we classify all kinds of obstacles and extract representative overall features from them and combine the ROI method with it to achieve high accuracy classification recognition. The shortcomings of traditional obstacle detection and recognition methods in terms of accuracy, reliability, and universality, as well as the high cost of traditional detection equipment, are addressed.

Data Availability

The dataset used in this paper is available from the corresponding author upon request.

Conflicts of Interest

The author declares that they have no conflicts of interest regarding this work.

References

[1] W. Zou, "The number of motor vehicles in China reaches 264 million," *Guangdong Traffic*, vol. 56, no. 1, 2015.

[2] Z. Dong, "Can Tesla leave its life in the hands of an intelligent driver after a fatal crash?" *Business School*, vol. 26, no. 8, pp. 106-107, 2016.

[3] H. Lu and Q. Zhang, *Data Acquisition and Processing*, no. 01, pp. 1–17, 2016, in Chinese.

[4] D. Li, ""13th Five-year plan" will be the best period for the development of China's automobile industry," *China Trade Guide*, no. 11, pp. 41-42, 2016.

[5] G. E. Hinton, S. Osindero, and Y. W. Teh, "A fast learning algorithm for deep belief nets," *Neural Computation*, vol. 18, no. 7, pp. 1527–1554, 2006.

[6] F. Pedregosa, G. Varoquaux, A. Gramfort, and V. Michel, "Scikit-learn: machine learning in python," *Journal of Machine Learning Research*, vol. 12, no. 10, pp. 2825–2830, 2012.

[7] M. Braverman, "Poly-logarithmic independence fools bounded-depth boolean circuits," *Communications of the ACM*, vol. 54, no. 4, pp. 108–115, 2011.

[8] G. E. Hinton and R. R. Salakhutdinov, "Reducing the dimensionality of data with neural networks," *Science*, vol. 313, no. 5786, pp. 504–507, 2006.

[9] X. Ding, *Recognition of Obstacles in Front of Vehicle Based on Convolutional Neural Network*, Dalian University of Technology, Dalian, 2015.

[10] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, 2015.

[11] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," *Advances in Neural Information Processing Systems*, vol. 25, no. 2, pp. 1097–1105, 2012.

[12] Z. Zhao, S. Yang, and Z. Ma, "Vehicle license plate character recognition based on convolutional neural network LENET-5 [J]," *Journal of System Simulation*, vol. 22, no. 3, pp. 638–641, 2010.

[13] Y. Zheng, Q. Chen, and Y. Zhang, *Journal of Image and Graphics*, vol. 19, no. 2, pp. 175–184, 2014.

[14] D. An, U. Meier, J. Masci, and J. Schmidhuber, "Special issue: multi-column deep neural network for traffic sign classification [J]," *Neural Networks*, vol. 32, no. 1, pp. 333–338, 2012.

[15] C. Ding and D. Tao, "Robust face recognition via multimodal deep face representation," *IEEE Transactions on Multimedia*, vol. 17, no. 11, pp. 2049–2058, 2015.

[16] C. Farabet, C. Couprie, L. Najman, and Y. LeCun, "Learning Hierarchical Features for Scene Labeling," *IEEE Transactions*

on Pattern Analysis and Machine Intelligence, vol. 35, no. 8, pp. 1915–1929, 2013.

[17] K. Xu, J. Ba, R. Kiros, and A. Courville, "Show, attend and tell: neural Image caption generation with visual attention," *Computer Science*, vol. 37, pp. 2048–2057, 2015.

[18] O. Vinyals, A. Toshev, S. Bengio, and D Erhan, "Show and tell: a neural image caption generator," *Computer Science*, vol. 20, no. 25, pp. 3156–3164, 2015.

[19] Y. Cui, *Nanchang*, Nanchang Hangkong University, Hangkong China, 2012.

[20] S. B. Goldberg, M. W. Maimone, and L. Matthies, "Stereo vision and rover navigation software for planetary exploration," *Proceedings, IEEE Aerospace Conference* in *Proceedings of the Aerospace Conference Proceedings*, vol. 5, pp. 2025–2036, California Institute of Technology, Big Sky, MT, USA, March 2002.

[21] D. Wu, *Visual Navigation Technology of Intelligent Mobile Robot [D]*, Zhejiang University, Hangzhou, 2001.

Hindawi
Computational Intelligence and Neuroscience
Volume 2022, Article ID 4854213, 9 pages
https://doi.org/10.1155/2022/4854213

Research Article

University Archives Autonomous Management Control System under the Internet of Things and Deep Learning Professional Certification

Yue Ma, Bing Dai, and Baorong Ding

The School of Civil Engineering, Harbin University, Harbin 150086, China

Correspondence should be addressed to Yue Ma; mayue8999@hrbu.edu.cn

Received 24 June 2022; Revised 27 August 2022; Accepted 5 September 2022; Published 21 September 2022

Academic Editor: Ning Cao

The current work aims to meet the needs of the development of archives work in colleges and universities and the modernization of management to realize the standards and standardization of all aspects of archives business construction in colleges and universities, so as to improve the political and professional quality of archives cadres. First, the radio frequency identification (RFID) technology based on the Internet of things (IoT) digitizes the university archive labels. Meanwhile, the filing cabinet's intelligent security system preserves confidential files. Second, the convolutional neural network (CNN) algorithm under deep learning is introduced and college profile information is identified. Finally, the concept of professional certification is used to clarify the purpose of the university archives automation management system. Different activation functions are used to analyze the recognition accuracy loss and recognition accuracy of university archives. The identification error of You Only Look Once (YOLO) of the ReLU-convolutional neural network (R–CNN) of college archives is analyzed. The results show that the selection of rectified linear units (ReLU) activation function for CNN can effectively reduce the loss of identification accuracy of college archives and can improve the accuracy of identification of college archives. The algorithm based on the ReLU activation function has a smaller recognition error accuracy in college archives than that of the YOLO algorithm. The recognition error of the YOLO algorithm is slightly higher than that of the R–CNN. The font recognition error of archival information based on the R–CNN is relatively large. However, the conclusion is reasonable due to the recognition difficulties of handwritten archival fonts. The file positioning recognition error rate is 19.00%, the file printing font recognition error rate is 4.75%, and the image recognition error rate is 1.90%. These results have a certain reference value for the process of identifying information in the automatic management of university archives by CNN under different activation functions.

1. Introduction

At present, there are many problems in the management of university archives. Many grassroot units have not formulated and implemented detailed rules for filing archives and lacked standardized operating procedures. The objectives of normative management of archives are not clear, the arrangement of archives lacks unified classification and enforceable standards, and the processing of archives is unscientific. These circumstances have resulted in a low degree of standardization in the generation, collection, filing, archiving, classification, grouping, and in numbering and cataloging of archives. The archived materials are not uniform, and it is difficult to fully reflect the school's teaching management level and teaching quality [1–3]. Most of the staff have not undergone professional archival knowledge training and do not understand the importance of archives, the requirements of archives management, the collection, arrangement, filing, and binding of archives materials, and their awareness of archives collection is weak. Moreover, most of the archives management personnel have several jobs and they do not devote enough energy to the archive's work. Some archived materials are not true and accurate enough, and the archived materials cannot be supplemented in a timely and accurate manner, so the archived materials cannot truly reflect the actual situation of

the individual, resulting in incomplete file management [4–6]. Some schools overemphasize the construction of scientific research, the proportion of scientific research funds is large, and they do not pay attention to the archives' work, and the funds for the archives cause are insufficient. Some colleges and universities have not included the archives work funds in the planning ranks, and even if they are included in the planning, the proportion is very small [7, 8]. In view of the abovementioned problems in the archives management work, how to do a good job in the archives management of colleges and universities is an urgent problem that needs to be solved at present [9, 10].

The original meaning of the Internet of Things (IoT) refers to connecting all items to the Internet through information sensing devices such as radio frequency identification (RFID) to achieve intelligent identification and management [11]. The Computer Systems Laboratory of Binghamton University has carried out a variety of work on wireless sensors in mobile self-organizing network protocols, application layer design of sensor network systems, etc. [12]. Using the sensor network hardware node, UbiCell integrates sensors, microprocessors, wireless transceivers, and other embedded chips and has various functions such as information acquisition, signal processing, data transmission, and real-time monitoring. Its wireless multimedia sensor network node has an image acquisition and processing capability of 300,000 pixels and 60 frames per second (FPS), which is sufficient to meet the application requirements of network monitoring and identification [13]. Researchers started a RFID-related electromagnetic theoretical research. The publication of those researches led many commercial companies to develop application equipment suitable for electronic artide surveillance (EAS) but it can only be used to identify targets. For detection, no more tags can be stored, so when there are multiple object tags present, the monitoring system cannot identify the identified item [14–16]. The United States initially applied RFID technology to the transmission industry and access control, while Europe used short-range monitoring RFID technology for animal monitoring [17]. The concept of "deep learning" first entered people's field of vision because of its powerful feature extraction capabilities and flexibility. The LeNet network architecture forms the current form of the convolutional neural network (CNN) and has achieved excellent results on the Mixed National Institute of Standards Technology (MNIST) handwritten digit recognition dataset [18].

In order to efficiently identify the information of university archives automatically and to reduce the waste of human resources, the current work digitizes college archives tags from IoT-based RFID technology. Meanwhile, it provides an intelligent security system for filing cabinets to preserve confidential archives. Second, it introduces the CNN algorithm under deep learning and provides information on college archives. Moreover, the purpose of the university archives automation management system is clarified through the concept of professional certification. Finally, for different activation functions, the university archives are analyzed for the loss of recognition accuracy and the accuracy of recognition. You Only Look Once (YOLO) of the ReLU-convolutional neural network (R–CNN) is used for error analysis of college file recognition. The current work has a certain reference value for the process of identifying information in the automatic management of university archives by CNN under different activation functions. The innovation is that the file identification is different from the traditional identification method, and the identification accuracy of college files under different activation functions is tested, which reduces the difficulty of identifying files. The combination of the rectified linear units (ReLU) activation function with the CNN structure can effectively identify different characters in the file and can identify consecutive characters.

Section 1 is the introduction, which illustrates the existing problems of archives management in colleges and universities and provides the research background. Then, this part also explains the research status of the IoT and deep learning technology, and finally, the research method process is described. Section 2 is the theoretical part, which explains the importance of the professional accreditation concept in the archives management of colleges and universities. Section 3 is the research method part, which expounds on the application of RFID technology, CNN, and contract lock in the archives management system of colleges and universities, respectively. Section 4 is the result part, which shows the recognition accuracy loss and recognition accuracy of different activation functions for college archives, as well as the analysis of the identification error of college archives by R–CNN YOLO. Section 5 is the discussion, which discusses the research results and illustrates the feasibility. Section 6 is the conclusion, describing the research results and expounding on the research shortcomings and future development directions.

2. The Concept of Professional Certification of University Archives

There are strict confidentiality management regulations for all kinds of students, teaching, personnel, and scientific research files in colleges and universities. In order to prevent the occurrence of modern signatures and leakage of information in the archiving process, contract locks are used to authenticate the real identity of the filers and ensure their signatures. Additionally, the concept of professional accreditation is used to strengthen the management control of the university archives system. The concept of professional certification is shown in Table 1.

3. Methods

3.1. The IoT-The Application of RFID in the Archives Management System of Colleges and Universities. IoT refers to various devices and technologies through various information sensing devices, such as sensors, RFID technology, a global positioning system (GPS), infrared sensors, laser scanners, and gas sensors. These devices and technologies can collect any object or process that needs to be monitored, connected, and interacted within real-time, and collects all

Table 1: Philosophy of professional certification.

Basic idea	Stress mechanism
Student center philosophy	Taking students as the center, making resource allocation and teaching arrangements around the achievement of training objectives and graduation requirements of all students, and taking the satisfaction of students and employers as an important reference for a professional evaluation
Output-oriented philosophy	Emphasizes the learning outcomes achieved by students after professional teaching design and teaching implementation, and evaluates the effectiveness of professional education against the core competencies and requirements of graduates
Continuous improvement concept	Effective quality monitoring and continuous improvement mechanism, which can continuously track the improvement effect, and use it to promote the continuous improvement of the quality of professional personnel training

kinds of information such as sound, light, heat, electricity, mechanics, chemistry, biology, and location and combines them with the Internet to form a huge Internet. Its purpose is to realize the connection between things and things, things and people, and all things and the network, which is convenient for identification, management, and control. RFID technology can support fast reading and writing, nonvisual identification, mobile identification, multitarget identification, positioning, and long-term tracking management [19]. RFID is not limited by size and shape in reading and does not need to match the fixed size and printing quality of paper for reading accuracy. RFID tags have the characteristics of miniaturization and various forms, making them applicable to different products.

RFID technology is an important part of the IoT perception layer. It is a technology that uses RF-radio frequency signals to realize contactless information transmission through spatial electromagnetic coupling and achieves object recognition through the transmitted information. Common RFID products are inductive electronic chips, inductive cards, contactless cards, electronic labels, and electronic barcodes. A complete RFID system consists of a reader and a transponder. Its action principle is that the reader transmits infinite radio wave energy of a specific frequency to the transponder, which is used to drive the transponder circuit to send out the internal ID code. The reader at this time will receive this ID Code. The principle of RFID identification is shown in Figure 1.

In Figure 1, RFID is mainly composed of an electronic label, an antenna, and a reader. There are two data areas in the electronic tag chip, namely, the identification (ID) area and the user data area. The ID area is used to store the globally unique identification code, the user identifier (UID). USD is stored in read only memory (ROM) when making chips and cannot be modified. The user data area is for users to store data, which can be read, written, and modified or added operations. Meanwhile, the antenna is a device that realizes the spatial propagation of RF-radio frequency signals and establishes a wireless communication connection. The language of communication between RFID and the reader is electromagnetic waves. Each time the reader sends a signal of a certain frequency through the antenna, after receiving the signal, the identification information stored in the electronic tag is transmitted through the wire, and the reader will receive it through the antenna. Moreover, we identify the information sent back by the electronic tag, and finally, the reader will send the identification result to the host, so that the purpose of identification is achieved. The RFID-based university archives management system is shown in Figure 2.

In Figure 2, the initialization of the RFID electronic label is completed by the RFID printer. The electronic label is gradually bound to all the data information in the electronic archives of the university as a unique identifier for the query archives. RFID readers are installed on the channels entering and exiting the archives management center of the university. When the files enter and leave the general file management center of the university, the reader will identify the file information by reading the electronic tag data and complete the rapid acquisition of the file entry and exit data. When the archives are transferred to other departments, the information of the incoming archives can be accurately proofread through the configured mobile collection equipment, and the inflow of the archives can be completed quickly and accurately. The intelligent filing cabinet manages some confidential files and realizes the functions of authorized pickup, real-time inventory, and remote monitoring. Based on this, the smart filing cabinet is processed with preventive measures. The intelligent security system of the intelligent filing cabinet is shown in Figure 3.

In Figure 3, each smart filing cabinet has an independent electronic lock control and an unauthorized personnel cannot open the filing cabinet. The whole cabinet is equipped with a video surveillance device and the school can design an alarm strategy. The school can link the security system to carry out alarm settings for mobile phone alarms, network alarms, and other alarm methods, to ensure the security of college confidential files in all-round way and improve the file prevention rate.

3.2. Positioning and Information Identification of University Archives Based on Deep Learning. Deep learning refers to a collection of algorithms that use various machine learning algorithms to solve various problems such as images and texts on multilayer neural networks. The core of deep learning is feature learning, which aims to obtain hierarchical feature information through hierarchical networks, to solve the important problems that required the manual design of features in the past. Deep learning is a framework that includes a variety of algorithms. Among them, the CNN is a kind of neural network specifically designed to process

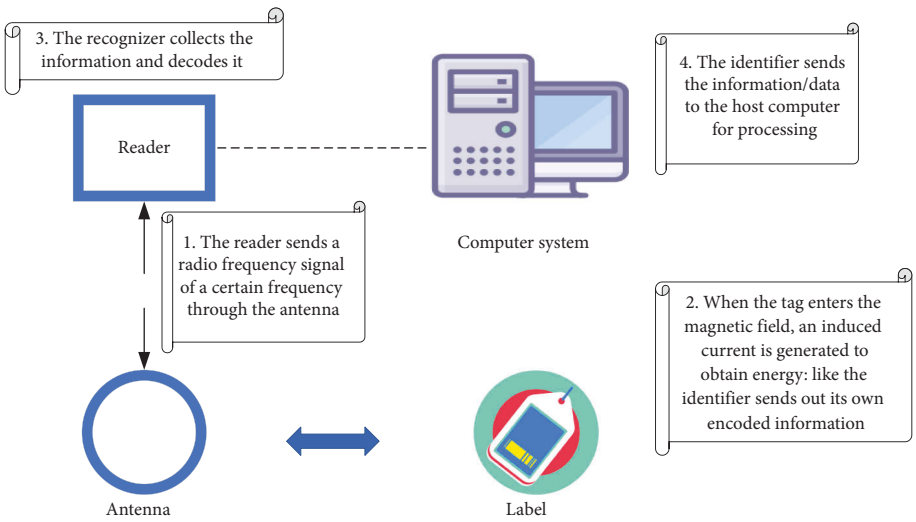


Figure 1: Identification principle of RFID.

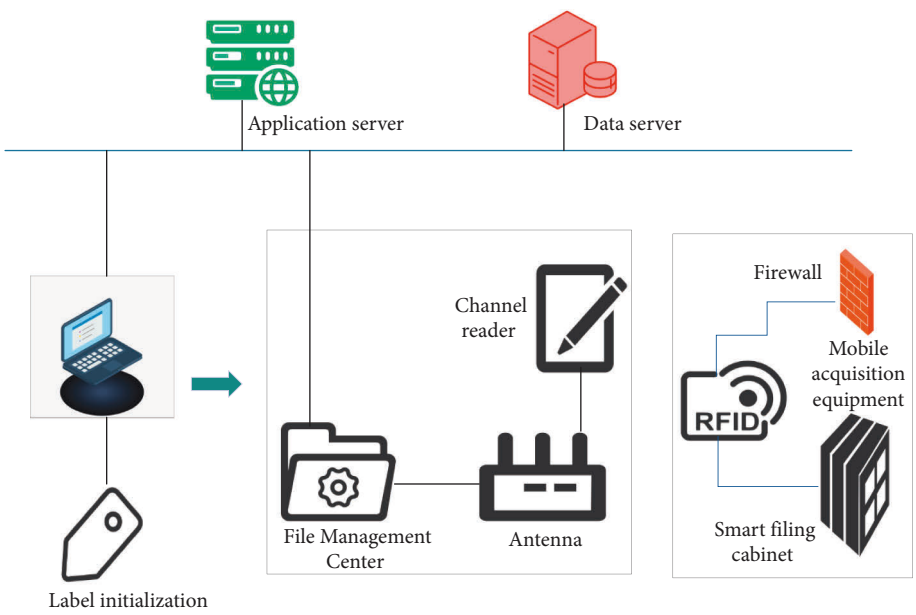


Figure 2: RFID-based university archives management system.

data with a grid-like structure. Convolutional networks are those neural networks that use convolution operations instead of normal matrix multiplication operations in at least one layer of the network. The CNN is essentially a filter. The filter is filtered by convolution. CNN work well with other types of models, such as recurrent neural networks and autoencoders, one example of which is symbolic language recognition. To some extent, the CNN tries to regularize on the basis of a feed-forward neural network (FNN) to prevent overfitting, and it can also identify the spatial relationship between data well [20]. The network structure of a simple CNN is shown in Figure 4.

In Figure 4, the CNN employs a convolutional function. The neurons between layers are not used for the connection. The convolutional layer connects the neurons in the part between the four layers. It consists of convolutional layers, pooling layers, and fully connected layers. Among them, the convolution layer cooperates with the pooling layer to form multiple convolutional groups, extract features layer by layer, and finally complete the classification through several fully connected layers. To sum up, CNN simulates feature distinction through convolution and reduces the order of magnitude of network parameters through weight sharing and pooling of convolution. Finally, traditional neural networks are used for tasks such as classification [21]. First, the image data of the archive page size A4 paper are input based on the archive image. The archive data label is shown in the following equation:

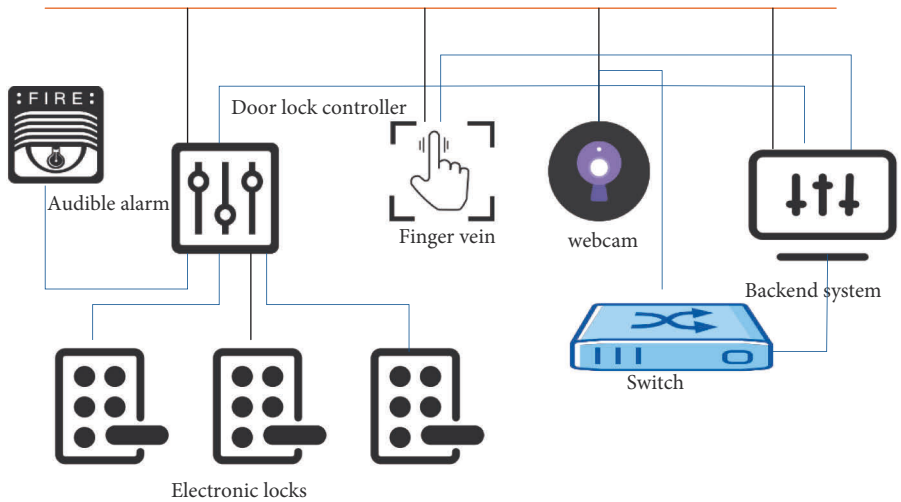


Figure 3: System for intelligent security of filing cabinets.

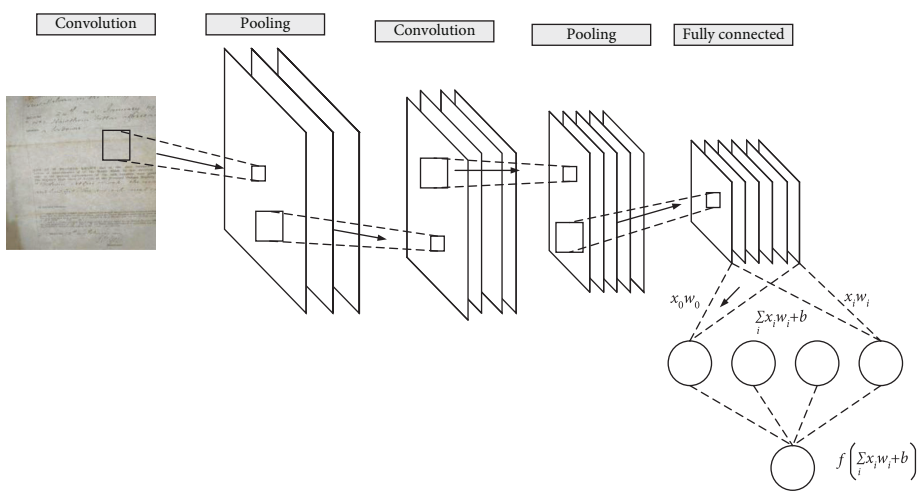


Figure 4: Structure of the CNN.

$$y = \left[P_c, b_x, b_y, b_w, b_h, c_1, c_2, c_3\right]^T. \quad (1)$$

In equation (1), P_c is 1 and there is a target (picture, file printing font, and file information font). b_x, b_y, b_w, and b_h are the horizontal and vertical coordinates of the center position of the image frame. The length and width of the border, c_1, c_2, and c_3 are the target categories. The loss function is as follows:

$$y = \left[1, b_x, b_y, b_w, b_h, 1, 0, 0\right]. \quad (2)$$

Convolution implements sliding window and realizes image window traversal. We constructed a subset of images to be recognized based on archive images and constructed a joint training set. The relationship between the size of the output layer and the size of the input layer, filter size, and step size is shown in the following equation:

$$\text{output} = \left[\frac{\text{input} - \text{filter} + 1}{\text{stride}}\right]. \quad (3)$$

In equation (3), *output* is the size of the output layer. *Input* is the size of the input layer. *Filter* is the size of the filter. *Stride* is the step size. The joint training set includes a subset of archival images to be identified and a subset of archival images. The computation of convolutional image features is shown in equations (4) and (5):

$$f = \text{ReLU}\left(\sum_{i=1}^{10} x_i w_i\right), \quad (4)$$

$$\text{ReLU} = f(x) = \begin{cases} \alpha x, & x < 0, \\ x, & x \geq 0. \end{cases} \quad (5)$$

When the convolutional image features are calculated, let the sliding window stride = 1. Re*LU* is a symbolic function. x_i is the file image feature data. w_i is the connection weight. f is the output value. α is a parameter. In addition, the activation function can also choose *sigmoid* and *tanh*, as shown in the equations (6) and (7):

$$\mathrm{f}(\mathrm{x}) = \frac{1}{1+e^{-x}}, \tag{6}$$

$$\tanh(\mathrm{x}) = \frac{\sinh\ (x)}{\cosh(x)} \quad = \frac{x^{x} - x^{-x}}{e^{x} + e^{-x}}. \tag{7}$$

The height and width of the archive image output are shown in equations (8) and (9):

$$\mathrm{h} = \frac{I\_h - c\_h}{\mathrm{stride}\_h} + 1, \tag{8}$$

$$\mathrm{w} = \frac{I\_w - c\_w}{\mathrm{stride}\_w} + 1. \tag{9}$$

The input file image size is $I\_w * I\_h = 210 * 297$. The sliding window step size is stride_w and stride_h. $c\_w * c\_h$ is the size of the convolution kernel. Multiple convolution kernels produce multiple convolutional feature maps. The output feature size is as follows:

$$M = \frac{I_w * I_h - F + 2P}{\mathrm{stride}} + 1. \tag{10}$$

M is the feature size of the output file image. F is the size of the convolution kernel. P is the pixel filled in the feature map. According to the variance of the target image, the picture area of the image to be identified, the file text printing area, and the student information text area to which the file belongs are distinguished. At this time, the classification category is $k = 3$. SoftMax is used to classify and identify different file information, and its function expression is as follows: as

$$P(y = c \mid x) = \frac{1}{1 + \sum_{k=1}^{c-1} e^{\theta_k^T x}}. \tag{11}$$

In equation (11), $c[c = 0,\ 1,\ 2]$ represents the type of classification. X is the variable value under different categories. θ is the surrogate estimated parameter. Intersection-over-Union, the concept used in IoU target detection is the overlap rate between the candidate bound and the ground truth bound, that is, the ratio of their intersection and union. The ideal situation is complete overlap, that is, a ratio of 1. Let C be the real frame of the file, G be the algorithm predicted frame, and its calculation is as follows:

$$LoU = \frac{\mathrm{area}(C) \cap \mathrm{area}(G)}{\mathrm{area}(C) \cup \mathrm{area}(G)}. \tag{12}$$

3.3. Application of Contract Lock in the University Archives Automation Management Control System. Archives management is one of the important basic tasks of colleges and universities. According to the "Administrative Measures for Archives of Colleges and Universities," colleges and universities shall file nearly ten kinds of archives materials such as students, teaching, personnel, administration, party and mass, scientific research, and equipment promptly every year. With the advancement of file management informatization, more and more colleges and universities have begun to introduce electronic signatures to promote the comprehensive digital transformation of file transfer and filing. The university archives automation management control system can be seen in Figure 5.

In Figure 5, the contract lock electronic signature system is open and integrated with the university archives management system and the school affairs management system to provide legal and effective electronic signature and electronic seal services for the transfer and filing of archives in the school [22]. Whether it is an electronic file or a paper file, the filing approval can be initiated through the process. The archivist signs and seals online, which comprehensively improves the efficiency of university file transfer and filing. When thousands of documents are signed in batches, covenant lock provides batch signing services. The archive files in the archiving stage are unified and aggregated. Archivers can add electronic signatures in batches with one click, and the system will automatically affix a special stamp for archiving. In this way, the filer can be freed from the mechanized signing work, and the intelligent conversion of document specifications can be realized. Colleges and universities have strict requirements on the specifications of archived files, and business personnel must adjust according to the requirements before they can approve the filing. To improve the efficiency of file specification conversion, contract locks can intelligently select the specification style and can automatically adjust it.

4. Results and Discussion

4.1. Loss and Accuracy Analysis of University Archives Recognition under Different Activation Functions. According to the aforementioned research methods, the identification of university archives is analyzed for identification loss and accuracy. Figure 6 shows the analysis of the accuracy loss of university file recognition under different activation functions.

As can be seen in Figure 6, the tanh activation function profile recognition accuracy loss is smaller than that of the ReLU activation function and less than the Sigmoid activation function. The tanh activation function profile has the smallest loss of recognition accuracy, with an average value of 0.16. Followed by the ReLU activation function, with an average recognition accuracy loss of 0.19. The largest loss is seen in the Sigmoid function, and the average loss of recognition accuracy in college archives is 0.38, and with the growth of the archive dataset, the degree of loss is uneven. Therefore, the ReLU activation function is chosen. Figure 7 shows the accuracy analysis of university archives recognition under different activation functions.

As shown in Figure 7, the growth of the archive dataset and the archive recognition rate under different activation functions do not have much effect. The recognition accuracy of the ReLU activation function is similar to that of the tanh function. The sigmoid function is not very accurate in identifying university archives. The average file recognition accuracy of the ReLU function is 0.95, the file recognition accuracy of the sigmoid function is 0.89, and the file

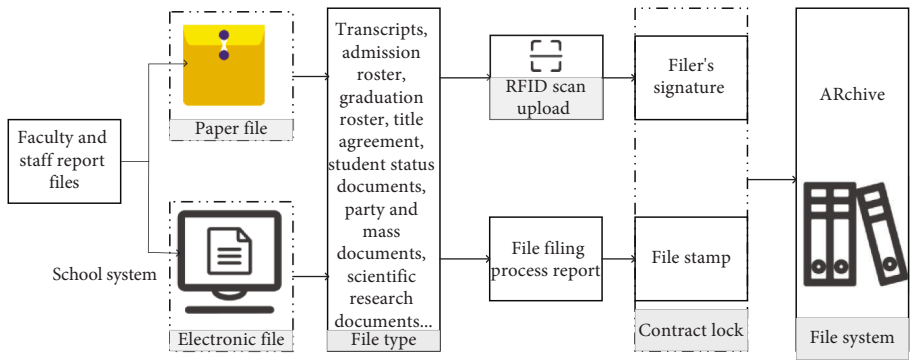


Figure 5: The control system for automatic management of university archives.

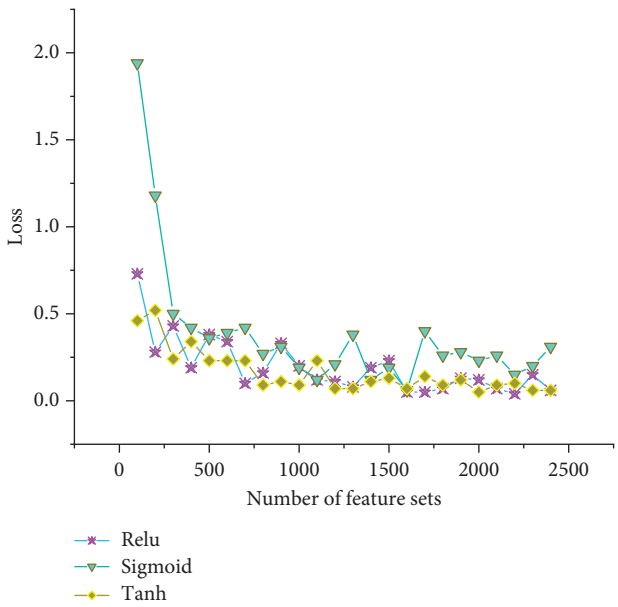


Figure 6: Analysis of the accuracy loss of college archives recognition under different activation functions.

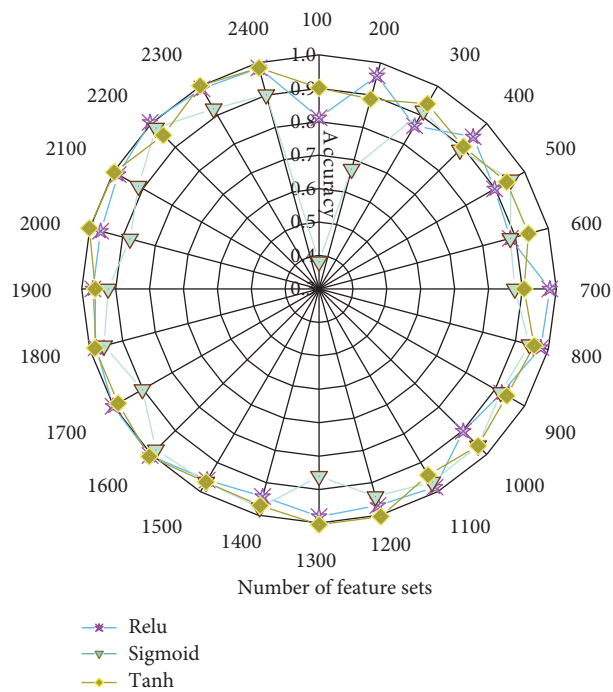


Figure 7: Analysis of the accuracy of university file recognition under different activation functions.

recognition accuracy of the tanh function is 0.96. Therefore, the recognition accuracy of ReLU and tanh functions is the best.

4.2. R-CNN and YOLO's University File Recognition Error Analysis. YOLO can use a fully connected layer at the end to predict dossier bounding boxes. Among them, the height of the bounding box is relative to the size of A4 paper, and because there are objects of different sizes and aspect ratios in each file image, it is difficult for YOLO to learn to adapt to the shapes of different objects during training, which leads to poor performance of YOLO in precise localization. YOLO draws on the prior box of faster-CNN's region proposal network (RPN). The PRN is used to convolve the feature map obtained by the CNN feature extractor to predict the bounding box and confidence (whether or not there is an object) at each location. Each location is set with a priori boxes of different sizes and scales. All RPN predicts are the offset of the bounding box relative to the prior, and using the prior makes learning easier. One of the pooling layers is removed to make the detection box resolution higher. The size of the input in the detection model is $210 * 297$. The total step size for downsampling in the YOLO model is 32. Therefore, the feature map size is $7 * 9$, and there is only one central position. The center point of some large objects falls in the center of the image, and it is relatively easy to use a center point of the feature map to predict the bounding box of the object. According to the automatic management and control system of university archives, R-CNN and YOLO are subjected to different degrees of recognition accuracy error analysis in the five categories of file printing fonts, file pictures, simulation, location, and file information fonts. Figure 8 shows the error analysis of R-CNN and YOLO's university file recognition.

As shown in Figure 8, the algorithm using the ReLU activation function has a smaller error in the recognition accuracy of university archives than that of the YOLO algorithm. The font recognition error of archive information under R-CNN is the largest, which is 71.60%. However, there

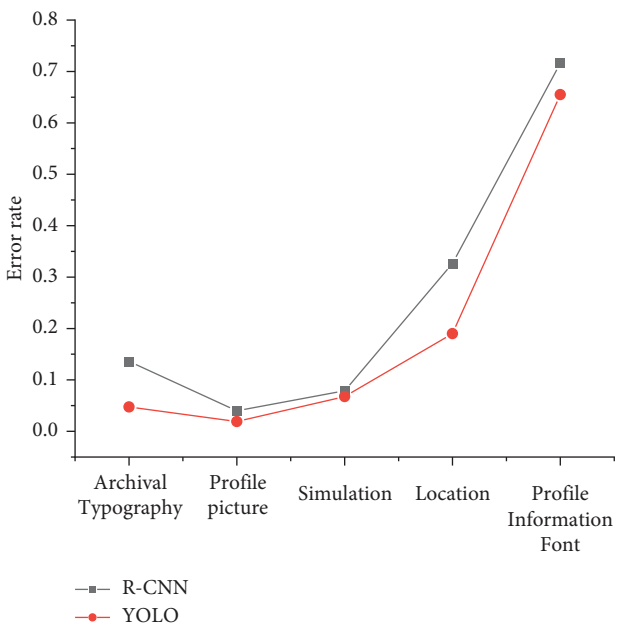


Figure 8: Analysis of university archives recognition error analysis of R-CNN and YOLO.

are difficulties in recognition based on handwritten archival fonts, so it is reasonable. In the process of simulating file recognition, the recognition error rate of R-CNN is 6.75% smaller than that of YOLO. Moreover, in the file positioning and recognition, the recognition error rate of R-CNN is 19.00% and the recognition error rate of YOLO is 32.65%. The recognition error rate of R-CNN in archivally printed fonts is 4.75% and the error rate is small. The error rate of YOLO in the file printing font is 13.60%. In the image recognition analysis, the R-CNN recognition error rate is 1.90% and the YOLO recognition error rate is 4.00%.

5. Discussion

The CNN under the ReLU activation function can effectively identify the character information of college archives. Harinahalli Lokesh and BoreGowda used deep learning technology to locate and recognize handwritten archives and found that deep learning technology could recognize continuous meteorological texts with a recognition accuracy of over 99.70% [23]. The conclusion not only shows that the use of CNN can identify university archives information but also that the identification effect is relatively consistent. This study analyzes the identification accuracy errors of different degrees in five categories of college archives' printing fonts, archive pictures, simulation, location, and archive font information, which can better reflect the effective path of identification and management of college archives.

6. Conclusion

Based on the IoT and deep learning methods, the university archives automation management and control system is studied under the concept of professional certification. The results show that the selection of the ReLU activation function for CNN can effectively reduce the loss of identification accuracy of college archives and can improve the accuracy of identifying college archives. In the analysis of the identification error of different categories of university archives, the font identification error of the archive's information under R-CNN is the largest, which is 71.60%. However, there are difficulties in recognition based on handwritten archival fonts, so it is reasonable. In the file location recognition, the recognition error rate of R-CNN is 19.00%, and the recognition error rate of YOLO is 32.65%. The recognition error rate of R-CNN in archivally printed fonts is 4.75% and the error rate is small. The error rate of YOLO in the file printing font is 13.60%. In the image recognition analysis, the R-CNN recognition error rate is 1.90% and the YOLO recognition error rate is 4.00%.

This study has a certain reference value for the process of identifying information in the automatic management of university archives under different activation functions of CNN [24–26]. However, a large amount of data during the operation of the CNN will cause hardware dependencies such as the central processing unit (CPU), trying to improve the CNN algorithm for optimization. Additionally, in the process of identifying university archives, this study is carried out under the guarantee that the fonts are not polluted, and the altered fonts are not identified and located [27–33]. It is hoped that the accuracy of CNN in image and font recognition will be increased, so as to be accurately recognized in the archives management system of colleges and universities.

Data Availability

The data supporting the current study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work was supported by Heilongjiang Higher Education Teaching Reform Research Project "Research and Practice on Grassroots Teaching Management Construction in Local Application-Oriented Colleges and Universities under the Background of Professional Certification" (No. SJGY20210520). Education Science "14th Five-Year Plan" 2021 Key Project of Heilongjiang Province, Research on Diversified Examination and Teaching Management of Engineering Majors in the Post-Epidemic Era (No. GJB1421375). Natural Science Foundation of Heilongjiang Province, "Research and Application of MEMS Gas Sensor" (No. LH2020F010). 2021 Innovation and Entrepreneurship Training Program for College Students in Heilongjiang Province, "Remote Monitoring System for Campus Lighting Based on Internet of Things" (No. 202110234005). Teacher Teaching Development Fund Project of Harbin University in 2021: Discussion and Practice of Blended Teaching Based on Knowledge Reconstruction and Interactive Learning (No. JFQJ2021009).

References

[1] T. Davenport and R. Kalakota, "The potential for artificial intelligence in healthcare," *Future healthcare journal*, vol. 6, no. 2, pp. 94–98, 2019.

[2] S. Wu, B. O. Wu, M. Liu et al., "Stroke in China: advances and challenges in epidemiology, prevention, and management," *The Lancet Neurology*, vol. 18, no. 4, pp. 394–405, 2019.

[3] O. J. Wouters, K. C. Shadlen, M. Salcher-Konrad et al., "Challenges in ensuring global access to COVID-19 vaccines: production, affordability, allocation, and deployment," *The Lancet*, vol. 397, no. 10278, pp. 1023–1034, 2021.

[4] P. J. Kroth, N. Morioka-Douglas, S. Veres et al., "Association of electronic health record design and use factors with clinician stress and burnout," *JAMA Network Open*, vol. 2, no. 8, p. e199609, 2019.

[5] Y. Chen, S. Ding, Z. Xu, H. Zheng, and S. Yang, "Blockchain-based medical records secure storage and medical service framework," *Journal of Medical Systems*, vol. 43, no. 1, pp. 5–9, 2019.

[6] T. L. Olsen and B. Tomlin, "Industry 4.0: opportunities and challenges for operations management," *Manufacturing & Service Operations Management*, vol. 22, no. 1, pp. 113–122, 2020.

[7] M. Martinus, N. T. P. Sihaloho, and D. N. Shadrina, "Archives management in support official administration in toho subdistrict head office, mempawah regency," *Ilomata International Journal of Social Science*, vol. 3, no. 1, pp. 340–346, 2022.

[8] D. Vasylenko, L. Butko, and D. Salem, "The archives governance in conditions of archive fund digitization," *Socio-Cultural Management Journal*, vol. 4, no. 1, pp. 83–100, 2021.

[9] W. Zhou, C. Wen, and L. Dai, "Collaborative construction of social-oriented family archives: a case study based on the practice of China," *Archives and Records*, vol. 41, no. 1, pp. 52–67, 2020.

[10] Y. Zhou, "The influence of emergencies on archives security management and its countermeasures," *Journal of Network and Computer Applications*, vol. 4, no. 1, pp. 29–36, 2019.

[11] Z. Lv and H. Song, "Mobile internet of things under data physical fusion technology," *IEEE Internet of Things Journal*, vol. 7, no. 5, pp. 4616–4624, 2020.

[12] J. Wang, X. Chen, and X. Gao, "Economic management teaching mode based on mobile learning and collaborative learning," *IEEE Access*, vol. 8, pp. 200589–200596, 2020.

[13] A. Genta, D. K. Lobiyal, and J. H. Abawajy, "Energy efficient multipath routing algorithm for wireless multimedia sensor network," *Sensors*, vol. 19, no. 17, p. 3642, 2019.

[14] C. Cheng, Y. Wu, Y. Qu, R. Ma, and R. Fan, "Radio-frequency negative permittivity of carbon nanotube/copper calcium titanate ceramic nanocomposites fabricated by spark plasma sintering," *Ceramics International*, vol. 46, no. 2, pp. 2261–2267, 2020.

[15] Y. Xia, W. Li, W. Clark, D. Hart, Q. Zhuang, and Z. Zhang, "Demonstration of a reconfigurable entangled radio-frequency photonic sensor network," *Physical Review Letters*, vol. 124, no. 15, p. 150502, 2020.

[16] W. E. Liu, Z. Y. Shi, J. Levinsen, and M. M. Parish, "Radio-frequency response and contact of impurities in a quantum gas," *Physical Review Letters*, vol. 125, no. 6, p. 065301, 2020.

[17] J. Ramírez-Faz, L. M. Fernández-Ahumada, E. Fernández-Ahumada, and R. Lopez-Luque, "Monitoring of temperature in retail refrigerated cabinets applying IoT over open-source hardware and software," *Sensors*, vol. 20, no. 3, p. 846, 2020.

[18] Y. F. Qin, H. Bao, F. Wang, J. Chen, Y. Li, and X. S. Miao, "Recent progress on memristive convolutional neural networks for edge intelligence," *Advanced Intelligent Systems*, vol. 2, no. 11, p. 2000114, 2020.

[19] A. V. Gotmare, S. U. Bokade, Z. Z. Inamdar, and S. Bhirud, "A systematic literature review on RFID application in manufacturing and supply chain management," *Industrial Engineering Journal*, vol. 12, no. 10, pp. 1–12, 2019.

[20] D. Durstewitz, G. Koppe, and A. Meyer-Lindenberg, "Deep neural networks in psychiatry," *Molecular Psychiatry*, vol. 24, no. 11, pp. 1583–1598, 2019.

[21] H. Wang, L. Yu, S. W. Tian, and Y. Peng, "Bidirectional LSTM Malicious webpages detection algorithm based on convolutional neural network and independent recurrent neural network," *Applied Intelligence*, vol. 49, no. 8, pp. 3016–3026, 2019.

[22] Y. Tian, T. Li, J. Xiong, M. Z. A. Bhuiyan, J. Ma, and C. Peng, "A blockchain-based machine learning framework for edge services in IIoT," *IEEE Transactions on Industrial Informatics*, vol. 18, no. 3, pp. 1918–1929, 2022.

[23] G. Harinahalli Lokesh and G. BoreGowda, "Phishing website detection based on effective machine learning approach," *Journal of Cyber Security Technology*, vol. 5, no. 1, pp. 1–14, 2021.

[24] B. Cao, J. Zhang, X. Liu et al., "Edge–cloud resource scheduling in space–air–ground-integrated networks for internet of vehicles," *IEEE Internet of Things Journal*, vol. 9, no. 8, pp. 5765–5772, 2022.

[25] G. Chen, P. Chen, W. Huang, and J. Zhai, "Continuance intention mechanism of middle school student users on online learning platform based on qualitative comparative analysis method," *Mathematical Problems in Engineering*, vol. 2022, Article ID 3215337, 12 pages, 2022.

[26] W. Zheng, X. Liu, and L. Yin, "Sentence representation method based on multi-layer semantic network," *Applied Sciences*, vol. 11, no. 3, p. 1316, 2021.

[27] W. Zheng, Y. Zhou, S. Liu, J. Tian, B. Yang, and L. Yin, "A deep fusion matching network semantic reasoning model," *Applied Sciences*, vol. 12, no. 7, p. 3416, 2022.

[28] W. Zheng and L. Yin, "Characterization inference based on joint-optimization of multi-layer semantics and deep fusion matching network," *PeerJ Computer Science*, vol. 8, p. e908, 2022.

[29] X. Liang, L. Luo, S. Hu, and Y. Li, "Mapping the knowledge frontiers and evolution of decision making based on agent-based modelling," *Knowledge-Based Systems*, vol. 250, p. 108982, 2022.

[30] S. Zhao, F. Li, H. Li et al., "Smart and practical privacy-preserving data aggregation for fog-based smart grids," *IEEE Transactions on Information Forensics and Security*, vol. 16, pp. 521–536, 2021.

[31] H. Kong, L. Lu, J. Yu, Y. Chen, and F. Tang, "Continuous authentication through finger gesture interaction for smart homes using WiFi," *IEEE Transactions on Mobile Computing*, vol. 20, no. 11, pp. 3148–3162, 2021.

[32] Y. Zhang, X. Shi, H. Zhang, Y. Cao, and V. Terzija, "Review on deep learning applications in frequency analysis and control of modern power system," *International Journal of Electrical Power & Energy Systems*, vol. 136, p. 107744, 2021.

[33] T. Zhong, M. Cheng, S. Lu, X. Dong, and Y. Li, "RCEN: a deep-learning-based background noise suppression method for DAS-VSP records," *IEEE Geoscience and Remote Sensing Letters*, vol. 19, pp. 1–5, 2022.

Hindawi
Computational Intelligence and Neuroscience
Volume 2022, Article ID 1835798, 12 pages
https://doi.org/10.1155/2022/1835798

Research Article

Predictive Control of the Mobile Robot under the Deep Long-Short Term Memory Neural Network Model

Lan Zheng

The School of Civil Engineering, Harbin University, Harbin 150086, China

Correspondence should be addressed to Lan Zheng; zhenglan@hrbu.edu.cn

Received 23 July 2022; Revised 24 August 2022; Accepted 30 August 2022; Published 21 September 2022

Academic Editor: Ning Cao

At present, there is a phenomenon of network data packet loss in the trajectory tracking control system, which will degrade or even destabilize the system's performance. Therefore, this work first explains the theory of the deep long-short term memory (LSTM) neural network model, the kinematic model of mobile robots, and the trajectory tracking error model. The reasons for data packet loss in the control system are analyzed. Second, a prediction model based on the LSTM network is designed according to the theory mentioned above. Finally, the training effect of the LSTM model and the robot trajectory tracking effect based on the model are tested by setting up simulation experiments. The research results are as follows: (1) The pose test error of the mobile robot will eventually tend to zero through the simulation curve generated by the pose parameters (x, y, θ) of the mobile robot. (2) The trajectory tracking error of the deep LSTM neural network prediction and compensation method with the packet loss rate of 5% is less than that with the packet loss rate of 10%. (3) The linear velocity v of the mobile robot based on the prediction model of the LSTM network varies greatly but is always in the interval (−2, 2). Its angular velocity ω initially fluctuates greatly but gradually tends to zero after about 13 s. (4) When the prediction model tracks the trajectory of the robot, the horizontal position x, the vertical position y, and the angle θ coincide with the reference trajectory. The exploration is conducted to provide a reference for the research on data packet loss in the networked mobile robot trajectory tracking system.

1. Introduction

1.1. Research Background and Motivations. Mobile robots can be divided into wheeled, bipedal, crawler, and crawling types according to different moving methods, and wheeled mobile robots are the most widely used [1]. The wheeled mobile robot is a typical nonholonomic mechanical system with nonlinear, underactuated, and drift-free characteristics, which brings great challenges to the research on its motion control. In particular, the trajectory tracking problem of mobile robots has become one of the key technologies [2]. Its trajectory tracking error system is usually a nonlinear strong coupling system, which does not meet the necessary conditions of Brockett. Mobile robots inevitably come into contact with the external environment during the movement process, and there are problems such as model uncertainty and external disturbances [3]. In addition, some scholars have proposed solutions to the problem of data packet loss in the networked mobile robot trajectory tracking system. For example, the virtual polling algorithm implemented at the application layer improves the network performance by reducing the degree of data conflict. However, this method does not consider the situation that the wireless channel will produce packet loss. Using the method of setting the transmission protocol to reduce the error caused by the unreliable wireless channel effectively suppresses the influence of the wireless channel, but does not compensate for the data packet loss caused by medium access [4]. Besides, there is a method to modify the controller output to improve the control performance by setting a gain schedule according to the current network conditions. Although it considers the influence of the wireless channel and medium access control, its statistical estimation algorithm for network data packet loss needs to be improved [5]. In general, most researchers are committed to optimizing existing network communication protocols, reducing the occurrence of network data packet loss and developing new network protocols to solve the impact of data packet loss on the performance of

networked mobile robots [6]. Although these two methods can improve the system performance, the changes or innovations for network protocols have an enormous workload and high work difficulty. It is effective to analyze the impact of data packet loss in the existing network communication on the performance of the control system, improve and optimize the controller and propose a good algorithm to predict and compensate for data packet loss [7].

1.2. Research Objectives. According to the content, the specific framework is given as shown in Figure 1.

This paper aims to use the LSTM neural network model to achieve high-precision prediction and control of mobile robot trajectories, thereby reducing some of the problems existing in mobile robots. This paper provides a reference value for further research on motion trajectory control.

2. Literature Review

The research on mobile robots began in the late 1960s, and Tianfu Innovation Institute was the first to begin research on mobile robots. At present, mobile robots have developed into an important branch of robotics [8]. In the aerospace field, there are lunar rovers and Mars rovers. The emergence of various reconnaissance robots and patrol cars in the military field, sweeping robots in the field of daily life, and hotel food delivery robots fully demonstrate the broad development prospects and application value of mobile robots [9]. The working conditions of the mobile robot have strong uncertainty according to the working field compared with the robot fixed in a position. Meanwhile, the mobile robot has the characteristics of unstructured, so its performance requirements are high. It needs to have the walking function, perception of the outside world, and some specific functions set by humans [10]. Researchers should conduct in-depth research on environmental perception, dynamic programming, and sensory information fusion [11].

Regarding the trajectory prediction research of mobile robots, Islam et al. combined differential flatness characteristics and integral sliding mode control and considered the dynamic characteristics of wheel actuators, which could realize trajectory tracking control of single-wheel mobile robots [12]. Barzegar-Kalashani et al. used a complete sliding mode controller for the first time in the research on trajectory tracking control of a two-wheeled mobile robot, which could effectively reduce the influence of uncertain factors and achieve a good tracking effect [13]. Yang et al. combined a neural network with a sliding mode control algorithm to ensure the stability of neural network adaptation, and they obtained appropriate equivalent control when the parameters of the robot model were unknown. This method could ensure that the output tracking error converged to zero [14]. Liang et al. designed an adaptive trajectory tracking control algorithm based on the kinematic error model of mobile robots according to the shortcomings of traditional trajectory tracking control laws. They verified that the algorithm could track the reference trajectory at a fast speed and has an excellent tracking control effect through experiments [15].

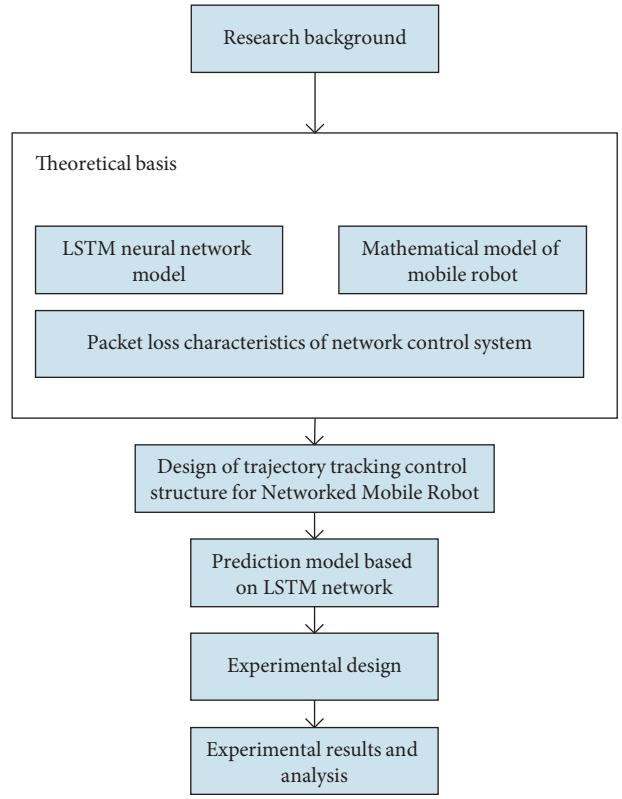


Figure 1: Structure of the work.

It is found that although the research on the trajectory tracking control of networked mobile robots with data packet loss has been developed, there are still many deficiencies by sorting out the literature. On the one hand, in the process of trajectory tracking control of mobile robots, the two sub-problems of motion planning and tracking control are usually solved separately. This results in that the external constraints are not considered when the reference trajectory of the mobile robot is given. The robot cannot reach the given input at some sampling moments, so it cannot achieve effective trajectory tracking. On the other hand, most of the current methods for improving the tracking performance of networked mobile robots are the development or optimization of network communication protocols with a heavy workload and high difficulty. It is urgent to propose an effective packet loss prediction compensation algorithm. The innovation lies in proposing a new prediction model and setting simulation experiments based on the original research on robot trajectory control. The new model is based on a long short-term memory (LSTM) neural network, which is optimized and integrated into robot trajectory prediction control. This work has a reference value in the research of robot trajectory prediction control.

3. Research Methodology

3.1. LSTM Neural Network Model

3.1.1. The Concept of the LSTM Model. With the continuous development of deep learning, the types of network architectures are also increasing, which are mainly divided into

two categories: deep discriminative models and deep generative models [16]. This work uses deep learning for network data packet loss prediction compensation, so the deep LSTM neural network in the deep discriminant model is adopted. The LSTM unit is a variant of the most widely used recurrent neural network (RNN). It inherits the characteristics of most of the RNN models and solves the problem that RNN is difficult to train due to long-term dependencies [17]. LSTM units realistically represent or simulate the cognitive processes of human behavior, logical development, and neural organization. It is suitable for handling problems that are highly related to time series, such as machine translation, dialogue generation, encoding, and decoding [18]. The LSTM unit structure is shown in Figure 2.

Figure 2 shows the overall research framework. For LSTM, the model cell adds a cell state to store the long-term state, so it can effectively solve the long-term dependency problem. The key to the LSTM unit is how to effectively control the cell state. Three gates are added: input gate, forget gate, and output gate. The essence of the gate is a fully connected layer, the input is a vector, and the output is a real vector between zero and one. The elements in the gate's output vector are in turn multiplied by the vector to be controlled. When the output of the gate equals zero, its product with any vector is zero. At this point, no information can pass through. When the output of the gate is equal to one, its product with any vector is unchanged. At this point, any information can pass [19]. LSTM cells control the cell state through input gates and forget gates. The input gate determines how much of the current input is stored in the cell. The forget gate determines how much of the cell state at the previous moment is retained in the current cell state. The output gate determines how many cell states are passed to the current output of the network [20].

3.1.2. Features of the LSTM Model. In the process of training the network model using the gradient descent method, the traditional RNN's weight update strategy is the correct direction based on the weight at the end of the output sequence. The weight changes only depend on the sequence input at the most recent moments, and the input data from a long time ago has little effect. Besides, the network has poor long-term memory function, and the training results are inclined to new input information [21].

During the training process of the LSTM model, the stochastic gradient descent method is used to optimize the network parameters. This is the standard optimization method recently adopted to train deep neural networks, and it is easy to implement [22]. At each iteration, the objective function is run on a subset of the training set, which is called the mini-batch. The gradient of the mini-batch objective function is obtained by back-propagating the neural network parameters. The step size is determined by the average value of the past step size and the current gradient, and the weight of the past step size is determined by the momentum hyperparameter. The parameters (weights and biases) are updated according to the new step size and scaled by the learning rate. The new mini-batch is randomly determined

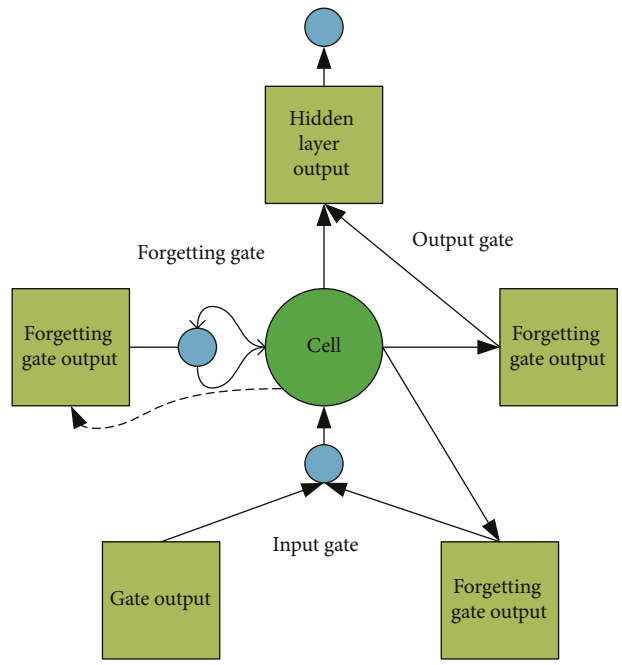


Figure 2: Schematic diagram of the LSTM unit structure.

from the training set and used for the next iteration. The learning rate is decreased during optimization [23].

3.2. Mathematical Model of the Mobile Robot

3.2.1. Kinematics Model of the Mobile Robot. The kinematic model of a mobile robot directly reflects the relationship between its pose state and control input, and it is the most intuitive mathematical model [24]. The research object is a typical two-wheel differential mobile robot. This kind of mobile robot is composed of a balance wheel and two driving wheels, and the mobile robot realizes the rotation of the car body through the differential speed of the two driving wheels. The balance wheel only has a supporting function and cannot provide power for the mobile robot. Its specific structure is displayed in Figure 3.

According to the analysis of the collected literature, the balance wheel does not provide power, so the balance wheel can be ignored when the kinematic model equation of the mobile robot is derived. Besides, its number will not affect the kinematic equation form. The kinematic model of the two-wheel differential mobile robot can be expressed as

$$\begin{bmatrix} \dot{x} \\ \dot{y} \\ \dot{\theta} \end{bmatrix} = \begin{bmatrix} \cos\theta & 0 \\ \sin\theta & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} v \\ \omega \end{bmatrix}, \tag{1}$$

(x, y) is the center position coordinate of the mobile robot. θ is the direction angle, and it is also the angle between the forward direction and the x-axis. v and ω are control input signals in the kinematic model, and their physical meanings are the linear and angular velocities of the mobile robot [25].

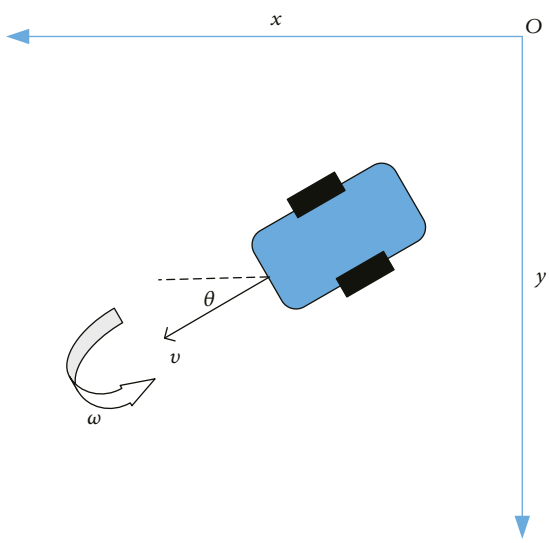


Figure 3: Schematic diagram of a two-wheel differential mobile robot.

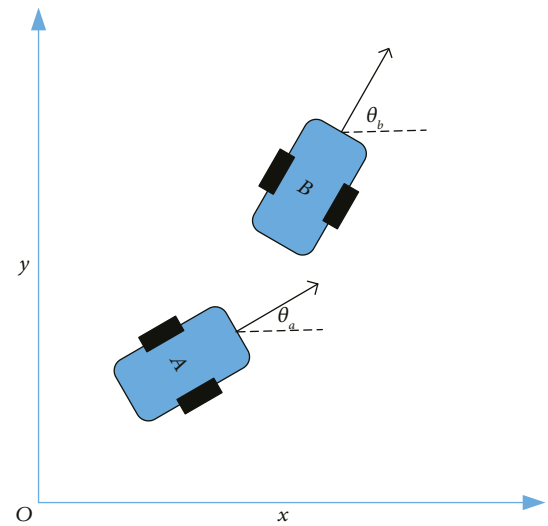


Figure 4: Schematic diagram of pose coordinate error.

3.2.2. Trajectory Tracking Error Model of the Mobile Robot. The mobile robot task trajectory tracking error model established here is demonstrated in Figure 4.

In Figure 4, point A (x_a, y_a) is the center position of the target mobile robot. θ_a is the direction angle. They together constitute the pose state vector of the target mobile robot $[x_a, y_a, \theta_a]^T$. Point B (x_b, y_b) is the center position of the reference mobile robot. θ_b is the direction angle. The pose state vector is $[x_b, y_b, \theta_b]^T$. At this time, the pose error between the target robot and the reference robot can be expressed as

$$\begin{bmatrix} x_e \\ y_e \\ \theta_e \end{bmatrix} = \begin{bmatrix} \cos\theta & \sin\theta & 0 \\ -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_b - x_a \\ y_b - y_a \\ \theta_b - \theta_a \end{bmatrix}. \tag{2}$$

Then, equation (3) can be acquired as

$$\begin{cases} x_e = (x_b - x_a)\cos\theta + (y_b - y_a)\sin\theta, \\ y_e = -(x_b - x_a)\sin\theta + (y_b - y_a)\cos\theta, \\ \theta_e = (\theta_b - \theta_a). \end{cases} \tag{3}$$

The trajectory tracking error model of the mobile robot can be obtained by derivation of (3).

$$\begin{bmatrix} \dot{x}_e \\ \dot{y}_e \\ \dot{\theta}_e \end{bmatrix} = \begin{bmatrix} v_b \cos\theta_e - v + y_e\omega \\ v_b \sin\theta_e - x_e\omega \\ \omega_b - \omega \end{bmatrix}. \tag{4}$$

The trajectory tracking control problem of the mobile robot can be transformed into the stabilization problem of the trajectory tracking error model shown in (4) based on the new pose error state vector $[x_e, y_e, \theta_e]^T$. All the items in the trajectory tracking error vector $[x_e, y_e, \theta_e]^T$ tend to be zero by designing appropriate v and ω control laws. When the system reaches $[x_e, y_e, \theta_e] = [0, 0, 0]$, $x_b = x_a$, $y_b = y_a$, and $\theta_b = \theta_a$, the mobile robot trajectory tracking control task is completed [26].

3.3. Packet Loss Characteristics of the Network Control System

3.3.1. Reasons for packet loss. In a network control system, each network node is frequently exchanging information. During network transmission, data collision or competition between nodes may occur, which will result in data loss during transmission. The network protocol has a retransmission mechanism to solve the problem of packet loss, but packet retransmission timeout may also occur when the system load is too heavy and there are many nodes in the network for data exchange. Packet loss can still happen. However, the phenomenon of packet loss sometimes occurs in the practical application of network control systems, which will reduce the system performance. Different network control systems will set a range of packet loss thresholds. Once the packet loss rate exceeds the threshold, the system will oscillate erratically, resulting in packet loss during network transmission [27]. There are three main reasons for packet loss.

A network node in the system has a communication failure. When the processor of a node in the network fails, the buffer where the data packet is located will be emptied, resulting in packet loss.

The packet transmission task in the system is too frequent. When the system communicates frequently, data conflicts inevitably occur. At this time, each node in the network competes for the right to use the network bandwidth, and the loss of data packets will occur. Although there is a retransmission mechanism, the destination node usually directly discards the packets that have not been transmitted beyond the retransmission time threshold.

There is channel interference in the network. The external environmental factors that the actual system is exposed to affect the transmission quality of the data packets in the channel. Channel interference in the network may cause disorder or loss of physical signals, resulting in distortion of data packets after reaching the destination node. At this time, valid data cannot be recovered through corresponding algorithms, and data packets are lost [28].

3.3.2. Description and Analysis of Network Packet Loss Characteristics. At present, the following three methods are mainly used to describe the data packet loss characteristics of the network control system. The first is statistical methods. For example, the probability distribution and the packet loss rate of data packets are assumed by using random system theory and switching system theory. The second is the theory of variable-delay systems. The total number of data packet losses between two sampling times needs to be given. The third is the theory of switched systems or predictive control theory. The sampling moment of packet loss in the system is regarded as the disconnection of the network transmission channel. A switching system with dynamic switches is used to represent the network control system structure of packet loss [29].

In the network control system, serial communication is used for information transmission among the controller, the controlled object, and the sensor. Each node device in the system shares the network channel bandwidth. In the case of frequent system communication, errors occur in the process of data packet transmission or reception due to the competition of each node for the right to use the network bandwidth. When the data packet does not reach the receiver within the specified time, it is called packet loss [30]. When data loss occurs in the system, the system information transmission channel is temporarily disconnected. Valid information at some sampling moments is not transmitted to the receiver in time. This will directly affect the structure and parameters of the system, causing system performance degradation. For the problem of data packet loss, although most network control systems have robustness, it is only limited to the number of data packet losses within the allowable threshold range. If this threshold is exceeded, it will seriously affect the control performance of the system and even make the system unstable [31]. Therefore, Figure 5 shows the structure obtained by simplifying the network control system with packet loss.

The switch can be opened and closed, and its state indicates whether there is a packet loss. When the switch is closed (S_2, S_4), the data packet at the sampling moment can be transmitted smoothly. When the switch is off (S_1, S_3), the data packet is lost at the sampling moment. At this time, the transmission size at the previous moment is regarded as the packet loss data at this moment or the data at this moment is directly set to zero, which are called the keep-input strategy and the zero-input strategy, respectively. The keep-input strategy is widely used, and its mathematical expression is as follows.

When the switch is located at S_1 and S_3, there i

$$\begin{cases} \overline{u}(k) = u(k-1), \\ \overline{x}(k) = x(k-1). \end{cases} \tag{5}$$

When the switch is located at S_1 and S_4, there is

$$\begin{cases} \overline{u}(k) = u(k-1), \\ \overline{x}(k) = x(k). \end{cases} \tag{6}$$

When the switch is located at S_2 and S_3, there is

$$\begin{cases} \overline{u}(k) = u(k), \\ \overline{x}(k) = x(k-1). \end{cases} \tag{7}$$

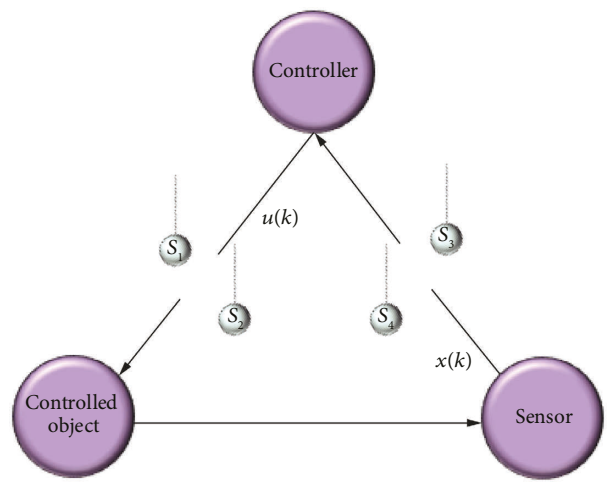


Figure 5: Simplified structure of the network control system.

When the switch is located at S_2 and S_4, there is

$$\begin{cases} \overline{u}(k) = u(k), \\ \overline{x}(k) = x(k), \end{cases} \tag{8}$$

$\overline{x}(k)$ represents the input value, and $\overline{u}(k)$ represents the output value. Besides, a new variable$z(k)$ is defined as

$$z(k) = \begin{bmatrix} x(k) \\ \overline{x}(k) \\ \overline{u}(k) \end{bmatrix}. \tag{9}$$

The network control system model with data packet loss can be expressed as

$$z(k+1) = \Phi_s z(k). \tag{10}$$

However, only adopting the keep-input strategy and the zero-input strategy to compensate for the data packet loss of the network control system cannot solve the problem of system performance degradation or even system instability caused by data packet loss. Reasonable methods need to be proposed to compensate for the negative impact of packet loss. Some scholars have proposed the method of predictive control to realize the prediction of missing data, and they use the predicted value obtained by this method to replace the lost control input value. The main problem with the predictive control method is the need to define extended vectors to assist in the design of the closed-loop system controller. The introduction of the extended vector increases the conservatism of the entire closed-loop control system, so it remains to find a good predictive compensation method to overcome this adverse effect.

3.4. Trajectory Tracking Control of the Networked Mobile Robot

3.4.1. Problem Description. The model structure of the networked mobile robot control system with the data packet loss problem studied here is shown in Figure 6.

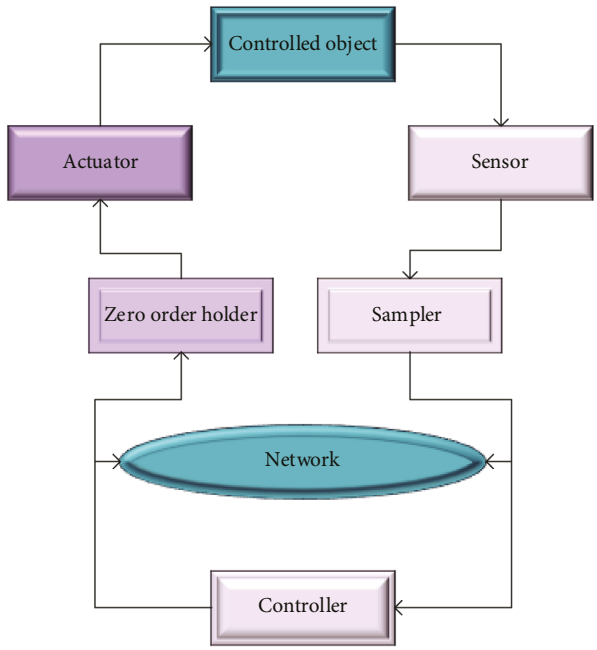


Figure 6: Network control system model with packet loss.

In this model, only the impact of packet loss on the system performance is considered during network transmission, and the sensor and controller are assumed to be time-driven and event-driven, respectively.

3.4.2. System Model. Based on the above analysis of the system packet loss characteristics and causes, the following assumptions are made to ensure the enforceability of the packet loss prediction and compensation method proposed here.

In the control loop, the sensor sends data packets to the controller with a constant sampling period. In practical industrial design and implementation, a time-driven approach is usually used to ensure a constant sampling period [32–37].

The sensors, controllers, and controlled objects in the control loop are time-synchronized through time-stamping technology to deal with the problem of possible out-of-order data packets [38–42].

In each sampling period, the deep neural network makes predictions for a single data packet obtained by the sensor and sends this data packet to the controller [43–48].

Based on the above assumptions, the structure of the networked mobile robot trajectory tracking control system under the deep neural network model reported here is revealed in Figure 7.

In Figure 7, $r(k)$ represents the reference pose of the mobile robot obtained by the partial point-taking method. $x(k)$ represents the actual motion pose of the mobile robot measured by the sensor. $e(k)$ represents the deviation between the actual pose and the reference pose of the mobile robot. This deviation serves as an input to the controller. $u(k)$ is the control amount of the mobile robot, including the linear velocity v and the angular velocity ω. $W(k)$ is the interference noise of the external environment. If no data packet loss occurs at the current sampling time, S_1 and S_2 are closed, and S_3 is disconnected. If data packet loss occurs at the current sampling time, S_1 and S_2 are disconnected, and S_3 is closed [49–54].

3.5. Prediction Model Based on the LSTM Network. According to the above LSTM model, the mathematical model of the mobile robot, and the analysis of the data packet loss characteristics, the prediction model based on the LSTM network proposed here is shown in Figure 8.

The mathematical description of the LSTM unit is as follows:

$$f_t = \sigma\left(W_f x_t + U_f h_{t-1} + b_f\right), \tag{11}$$

$$i_t = \sigma\left(W_i x_t + U_i h_{t-1} + b_i\right), \tag{12}$$

$$o_t = \sigma\left(W_o x_t + U_o h_{t-1} + b_o\right), \tag{13}$$

$$c_t = f_t \odot c_{t-1} + i_t \odot \tan h\left(W_c x_t + U_c h_{t-1} + b_c\right), \tag{14}$$

$$h_t = o_t \odot \tan h\ \left(c_t\right). \tag{15}$$

In Equations (11)–(15), x_t is the input vector at time t, f_t, i_t, o_t, c_t, and h_t represent forget gate, input gate, output gate, cell state, and hidden layer output, respectively. W_f, W_i, W_o, and W_c are input weights. U_f, U_i, U_o, and U_c are cycle weights. b_f, b_i, b_o, and b_c are biased values. σ and tanh are the *Sigmoid* and tanh activation functions, respectively. $\odot$ represents the dot multiplication operation.

4. Experimental Design and Performance Evaluation

4.1. Datasets Collection. This work will use the deep LSTM neural network toolbox is used to calculate the unmodeled parameters of the model, and 1100 groups of data with two inputs and three outputs are randomly generated. The first 1000 groups are used to train the neural network, and the last 100 groups are used to test the modeling error. The whole dataset comes from the output of the deep LSTM neural network toolbox.

4.2. Experimental Environment. To verify the prediction model of robot trajectory tracking error based on the LSTM neural network without different packet loss rates, this paper builds and simulates the model by using the toolbox in MatLab application.

4.3. Parameters Setting. In the process of modeling, the training goal of the model is set to 0.001, and other parameters are the default values.

4.4. Performance Evaluation

4.4.1. LSTM Network Model Training Effect. The prediction model based on the LSTM network is trained by setting up simulation experiments. Figure 9 reveals the results.

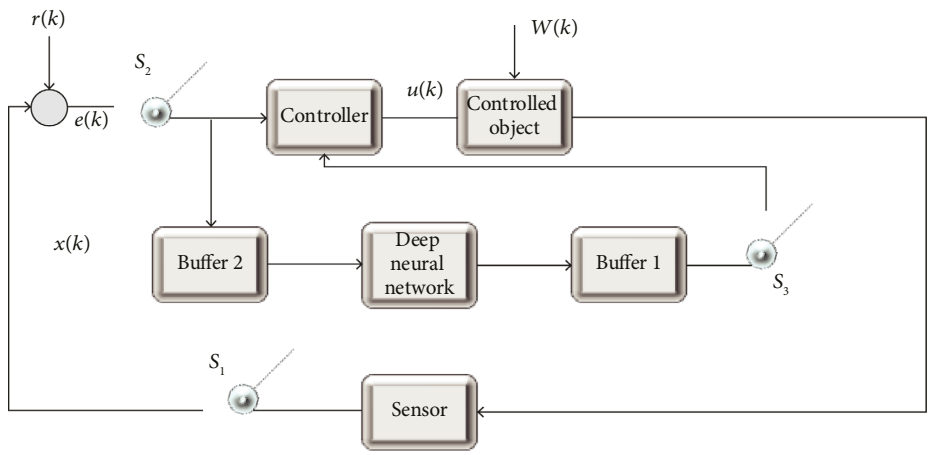


Figure 7: Structure of the mobile robot trajectory tracking control system based on the deep neural network.

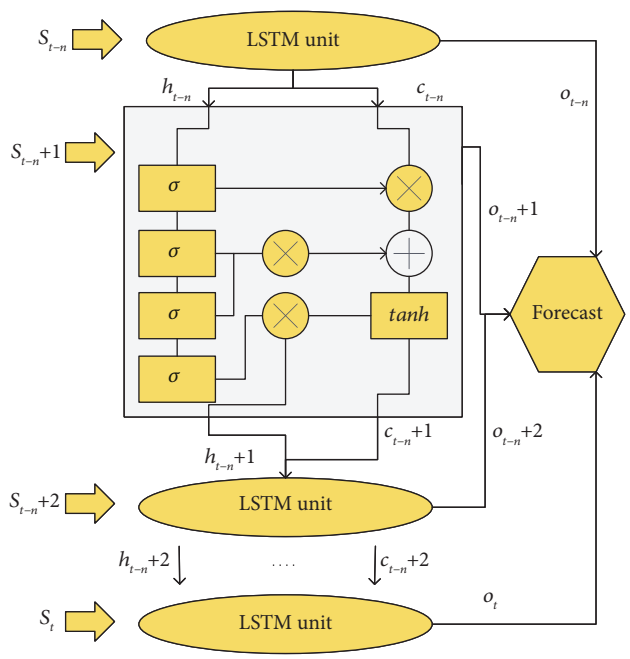


Figure 8: Mobile robot trajectory prediction model based on the LSTM network.

The meaning of the upper coordinate axis in Figure 9 refers to the serial number of the training samples collected in the experiment. In Figure 9, three lines with different colors represent the pose parameters (x, y, θ) of the mobile robot. The simulation curves generated by these three variables show that the pose test error of the mobile robot will eventually tend to zero. Therefore, it can be determined that the trained LSTM neural network can approximate the mobile robot model infinitely. The accuracy and efficiency are also high, which lays a solid foundation for the trajectory tracking controller to have a good tracking performance.

4.4.2. Robot Trajectory Tracking Control Effect Based on the LSTM Model. (1) Robot Trajectory Tracking Error Under Different Packet Loss Rates. The trajectory tracking error results of the prediction model based on the LSTM model under different packet loss rates are shown in Figure 10.

Figure 10 indicates that the trajectory tracking error of the deep LSTM neural network prediction compensation method when the packet loss rate is 5% is smaller than that of the deep LSTM neural network prediction compensation method when the packet loss rate is 10%. Generally, this method can complete the corresponding trajectory tracking task.

(2) Predictive Control Effect Of the Robot Based on the LSTM Network. Figure 11 shows the result of the change of control amount when the robot based on the LSTM network moves through the simulation experiment.

Figure 11 reveals that the linear velocity v of the mobile robot based on the prediction model of the LSTM network varies greatly but is always in the interval $(-2, 2)$. The angular velocity ω initially fluctuates greatly but gradually tends to zero after about 13 s. In conclusion, the control increment changes of the mobile robot under this model are stable. The prediction effect of the model is further verified.

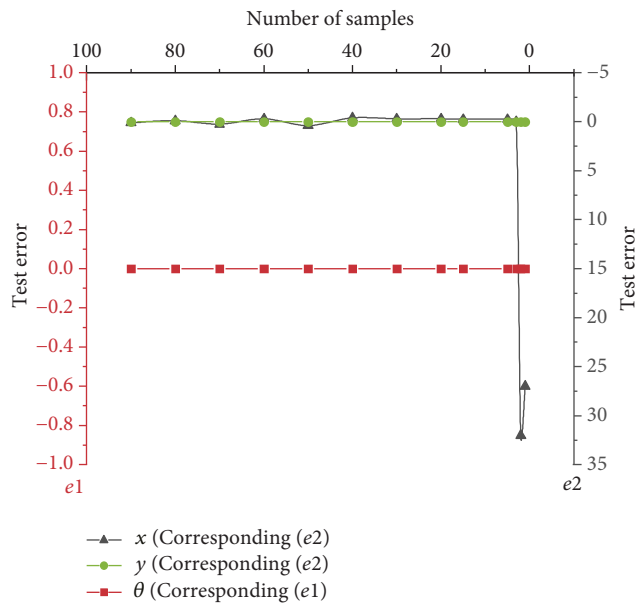


Figure 9: Prediction model training results based on the LSTM network.

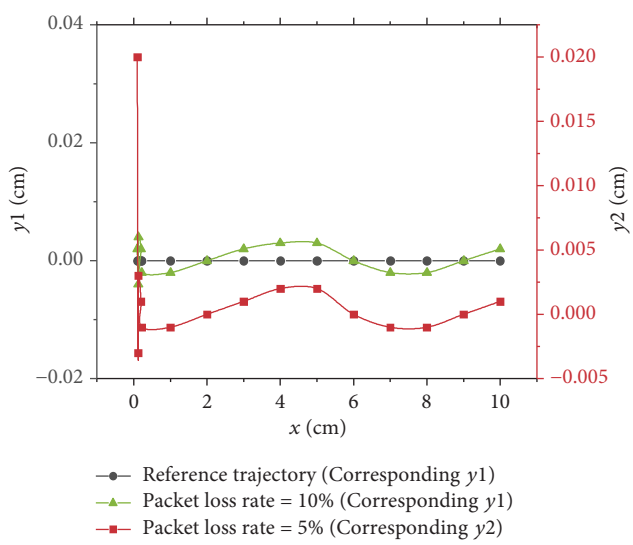


Figure 10: Trajectory tracking error results under different packet loss rates.

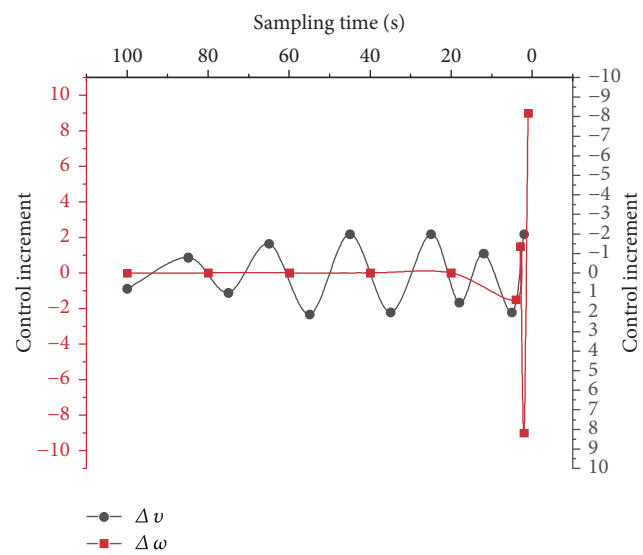


Figure 11: Control increment changes for mobile robots.

(3) Robot Linear Trajectory Tracking Effect Based on the LSTM Network. Figure 12 shows the effect of robot linear trajectory tracking based on the LSTM network through simulation experiments.

Figure 12 shows that each variable of the predicted trajectory has good prediction accuracy when the prediction model is used to track the robot's moving trajectory. It suggests that the lateral position x, longitudinal position y, and angle θ predicted by the LSTM model basically coincide with the reference trajectory provided. The above analysis shows that the model has a good prediction effect on the trajectory of the robot, and the effect of the prediction model is finally verified.

5. Discussion

On the one hand, the results show that the LSTM model is different from other network models. The training time of the model is less because of the particularity of its structure, so it has higher accuracy, which provides a reliable basis for the high-precision prediction of the robot trajectory. In addition, the model can simulate and predict the motion trajectory of the robot with high precision by reducing the packet loss rate in the network control system. On the other hand, based on the previous research on mobile robot control, this work uses the deep learning algorithm to design its control model. The selected deep learning algorithm is the LSTM algorithm, which has a good performance for

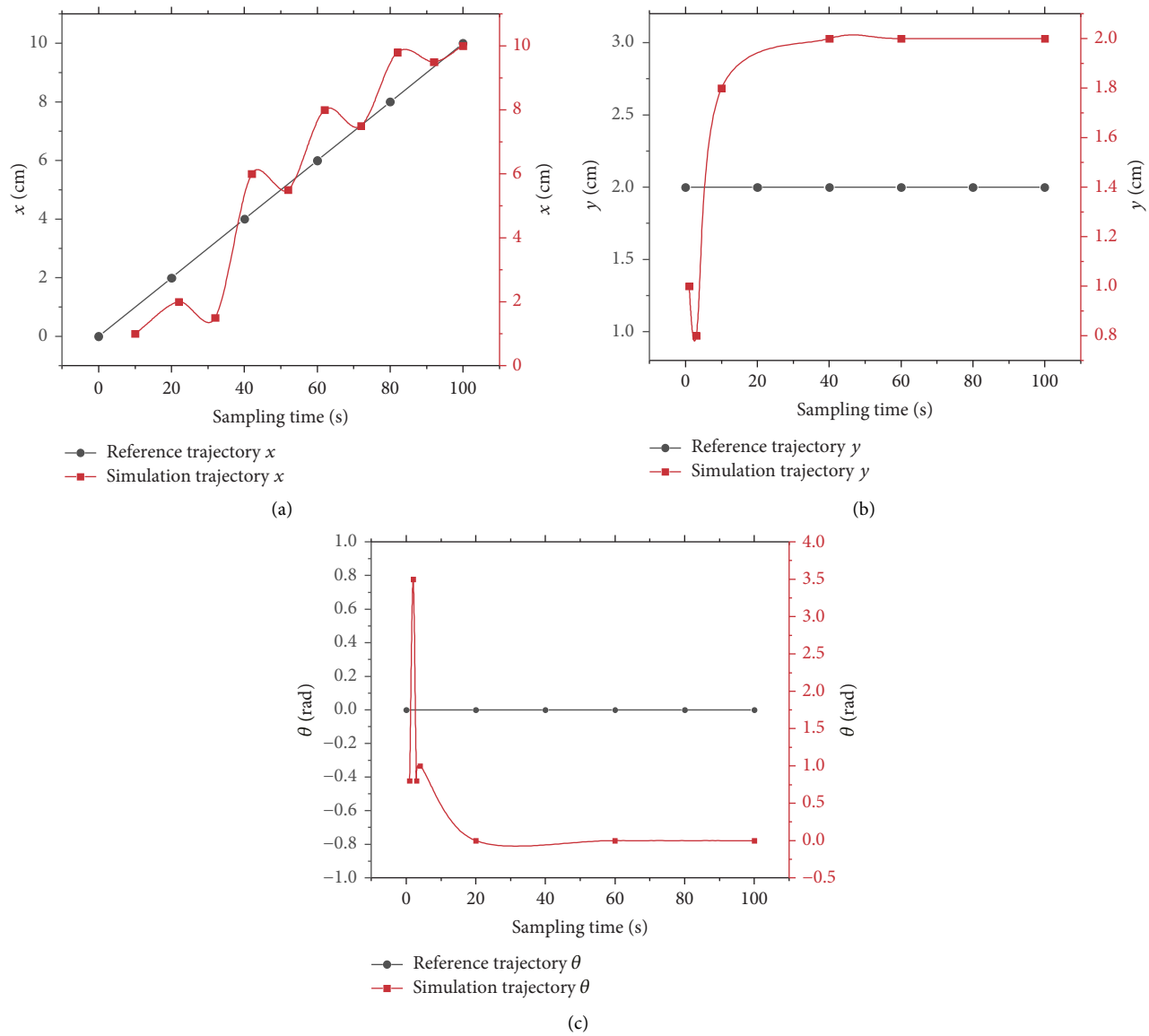


Figure 12: Mobile robot trajectory tracking status. (a) The lateral displacement tracking the state of the mobile robot; (b) the longitudinal displacement tracking the state of the mobile robot; (c) rotation angle tracking the state of the mobile robot.

predictive control. Therefore, this algorithm is selected as the theoretical support to complete the prediction of mobile robot trajectory. The basis is the existing problem of trajectory predictive control of mobile robots.

6. Conclusion

6.1. Research Contribution. This paper presents a robot trajectory prediction model based on an LSTM neural network combined with the mathematical model of mobile robots and the phenomenon of data packet loss. The research object is the existing problems of trajectory prediction control of mobile robots. The author presents a robot trajectory prediction model based on the LSTM neural network combined with the mathematical model of mobile robots and the phenomenon of data packet loss. After the model is tested by setting up simulation experiments, the following conclusions are drawn. First, the trained LSTM neural network can approximate the mobile robot model infinitely, not only with high accuracy but also with high efficiency. Second, in the case of different packet loss rates, the model can still complete the corresponding curve tracking task well. Third, the control increment changes of the mobile robot under this prediction model are stable. Fourth, when the prediction model tracks the movement trajectory of the robot, the specific variables in it are coincident with the reference trajectory.

6.2. Future Works and Research Limitations. The deficiency lies in that, on the one hand, only the straight-line form is selected for the reference trajectory when testing the model, which does not reflect the model testing effect under other reference trajectory forms. The purpose is to design a

trajectory prediction model by combining neural network technology and the kinematic model to improve the accuracy and quality of trajectory prediction. On the other hand, the effect of the LSTM prediction model proposed here has only been reflected in the simulation experiment, but it has not been applied to the mobile robot, and its effectiveness needs to be further studied. Subsequently, the reference trajectory in the form of the sine curve will be set to further verify the effect of the model. Additionally, it will be introduced into a specific mobile robot example to further verify the effect of the model.

Data Availability

The data are available from the corresponding author upon request.

Ethical Approval

This article does not contain any studies with human participants or animals performed by any of the authors.

Consent

Informed consent was obtained from all individual participants included in the study.

Conflicts of Interest

The author declares that there are no conflicts of interest.

References

[1] D. V. Shabanov, A. V. Kozlovich, R. R. Valiev, and O. V. Kochneva, "Autonomous wheeled mobile robot maneuvering in constraint environment. trajectory tracking quality criteria," *Journal of Physics: Conference Series*, vol. 1753, no. 1, Article ID 012051, 2021.

[2] C. Zhao, D. Wang, J. Hu, and Q. Pan, "Nonlinear model predictive control-based guidance algorithm for quadrotor trajectory tracking with obstacle avoidance," *Journal of Systems Science and Complexity*, vol. 34, no. 4, pp. 1379–1400, 2021.

[3] M. A. Miroliubov, A. K. Samusev, I. D. Toftul, and M. I. Petrov, "Spectral characteristics and time dynamics of tunable acoustic resonators in the strong coupling regime," *JETP Letters*, vol. 113, no. 8, pp. 547–553, 2021.

[4] V. Seedha Devi, T. Ravi, and S. B. Priya, "Cluster based data aggregation scheme for latency and packet loss reduction in WSN," *Computer Communications*, vol. 149, pp. 36–43, 2020.

[5] X. Lei, X. Chen, and S. H. Rhee, "A hybrid access method for broadcasting of safety messages in ieee 802.11p VANETs," *EURASIP Journal on Wireless Communications and Networking*, vol. 2021, no. 1, pp. 71–18, 2021.

[6] J. V. Sorribes, L. Peñalver, and J. Lloret, "A spontaneous wireless ad hoc trusted neighbor network creation protocol," *Wireless Communications and Mobile Computing*, vol. 2021, no. 2, 20 pages, Article ID 5531923, 2021.

[7] C. H. Lin and T. Y. Wang, "A novel convolutional neural network architecture of multispectral remote sensing images for automatic material classification," *Signal Processing: Image Communication*, vol. 97, no. 105, Article ID 116329, 2021.

[8] A. B. Filimonov and N. B. Filimonov, "Constructive aspects of the method of potential fields in mobile robotics," Автометрия, vol. 57, no. 4, pp. 45–53, 2021.

[9] S. V. Ivanova and O. B. Ivanov, "Education in the era of the fourth industrial revolution: development vector, prospects and challenges for Russia," *Space and Culture India*, vol. 7, no. 5, pp. 70–79, 2020.

[10] B. Debnath, A. Ghosh, and S. Deb, "Designing a rapid optical response trigger(RORT) for self-navigating and path explorer robot," *Procedia Computer Science*, vol. 167, pp. 1543–1552, 2020.

[11] M. A. Ali, M. S. Radzak, M. Mailah et al., "A novel inertia moment estimation algorithm collaborated with active force control scheme for wheeled mobile robot control in constrained environments," *Expert Systems with Applications*, vol. 183, no. 6, Article ID 115454, 2021.

[12] Y. Islam, I. Ahmad, M. Zubair, and K. Shahzad, "Double integral sliding mode control of leukemia therapy," *Biomedical Signal Processing and Control*, vol. 61, Article ID 102046, 2020.

[13] M. Barzegar-Kalashani, M. A. Mahmud, M. A. Barik, and A. M. T. Oo, "Control of arc suppression devices in compensated power distribution systems using an integral sliding mode controller for mitigating powerline bushfires," *International Journal of Electrical Power & Energy Systems*, vol. 134, no. 5, Article ID 107481, 2022.

[14] Z. Yang, D. Wang, X. Sun, and J. Wu, "Speed sensorless control of a bearingless induction motor with combined neural network and fractional sliding mode," *Mechatronics*, vol. 82, Article ID 102721, 2022.

[15] Z. Zhang, Q. Wu, X. Li, and C. Liang, "Barrier Lyapunov function-based robot control with an augmented neural network approximator," *Industrial Robot: The International Journal of Robotics Research and Application*, vol. 49, no. 2, pp. 359–367, 2022.

[16] L. Chen, A. Fan, H. Shi, and G. Chen, "Region-action LSTM for mouse interaction sequence based search satisfaction evaluation," *Information Processing & Management*, vol. 57, no. 6, Article ID 102349, 2020.

[17] Z. Chen, C. Yang, and J. Qiao, "The optimal design and application of LSTM neural network based on the hybrid coding PSO algorithm," *The Journal of Supercomputing*, vol. 78, no. 5, pp. 7227–7259, 2021.

[18] H. Zhu, C. He, Y. Fang, B. Ge, M. Xing, and W. Xiao, "Patent automatic classification based on symmetric hierarchical convolution neural network," *Symmetry*, vol. 12, no. 2, 186 pages, 2020.

[19] A. N. Moustafa and W. Gomaa, "Gate and common pathway detection in crowd scenes and anomaly detection using motion units and LSTM predictive models," *Multimedia Tools and Applications*, vol. 79, no. 29-30, pp. 20689–20728, 2020.

[20] Y. Sun and X. Xu, "Calibration of MEMS triaxial accelerometers based on the maximum likelihood estimation method," *Mathematical Problems in Engineering*, vol. 2020, no. 2, pp. 1–10, 2020.

[21] J. S. Lin and B. H. She, "A bci system with motor imagery based on bidirectional long-short term memory," *IOP Conference Series: Materials Science and Engineering*, vol. 719, no. 1, Article ID 012026, 2020.

[22] X. Han, Z. Wang, and H. J. Xu, "Time-weighted collaborative filtering algorithm based on improved mini batch k-means clustering," *Advances in Science and Technology*, vol. 105, pp. 309–317, 2021.

[23] V. N. Sizykh, S. B. Antoshkin, R. A. Daneev, M. V. Bakanov, A. V. Livshits, and A. A. Aleksandrov, "Analytical design of control system mathematical models for mobile robots based on the methods of inverse problems of dynamics and modal PID controllers," *IOP Conference Series: Materials Science and Engineering*, vol. 760, no. 1, Article ID 012053, 2020.

[24] S. Glowinski, T. Krzyzynski, A. Bryndal, and I. Maciejewski, "A kinematic model of a humanoid lower limb exoskeleton with hydraulic actuators," *Sensors*, vol. 20, no. 21, 6116 pages, 2020.

[25] U. Zangina, S. Buyamin, M. Abidin, S. Azimi, and H. S. Hasan, "Non-linear PID controller for trajectory tracking of a differential drive mobile robot," *Journal of Mechanical Engineering Research and Developments*, vol. 43, no. 1, pp. 255–270, 2020.

[26] E. Wiyadi, R. N. Setiadi, and L. Umar, "Effect of vegetation profile and air data rate on packet loss performance of lora e32-30dbm 433 MHz as a wireless data transmission," *Journal of Physics: Conference Series*, vol. 1655, no. 1, Article ID 012015, 2020.

[27] J. Guo, J. Cheng, and J. D. Diao, "System identification with binary-valued output observations under either-or communication and data packet dropout," *Systems & Control Letters*, vol. 156, no. 2, Article ID 105010, 2021.

[28] H. Wang, Y. Li, and J. Qian, "Self-adaptive resource allocation in underwater acoustic interference channel: a reinforcement learning approach," *IEEE Internet of Things Journal*, vol. 7, no. 4, pp. 2816–2827, 2020.

[29] S. Nurkholiq, R. K. Putra, A. Khumaidi, F. Fatimah, and E. Kusuma, "Redesign integrated control system GTG and HRSG to reduce loss of electrical production at combined cycle power plant muara karang," *IOP Conference Series: Materials Science and Engineering*, vol. 1096, no. 1, Article ID 012087, 2021.

[30] Y. Zhou, J. Zhao, and X. Liu, "Loss minimization control of dual three-phase linear induction motor with system-level loss model," *IEEJ Transactions on Electrical and Electronic Engineering*, vol. 16, no. 2, pp. 324–331, 2021.

[31] A. Shan, X. Fan, C. Wu, X. Zhang, and S. Fan, "Quantitative study on the impact of energy consumption based dynamic selfishness in MANETs," *Sensors*, vol. 21, no. 3, 716 pages, 2021.

[32] L. Cheng, L. Zhang, J. Li, and Y. Ke, "Measurement, identification, and compensation of pose errors for six-axis gantry automated fiber placement machine," *International Journal of Advanced Manufacturing Technology*, vol. 120, no. 3-4, pp. 2259–2276, 2022.

[33] W. Hu, S. Zhang, and L. Niu, "Self-adaptive PID control based on RBF network for trajectory tracking of dual-mass servo system," *Journal of Physics: Conference Series*, vol. 1871, no. 1, Article ID 012113, 2021.

[34] G. Sun, Y. Cong, Q. Wang, B. Zhong, and Y. Fu, "Representative task self-selection for flexible clustered lifelong learning," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 33, no. 4, 1481 pages, 2020.

[35] F. Liu, G. Zhang, and J. Lu, "Multisource heterogeneous unsupervised domain adaptation via fuzzy relation neural networks," *IEEE Transactions on Fuzzy Systems*, vol. 29, no. 11, pp. 3308–3322, 2021.

[36] D. Li, S. S. Ge, and T. H. Lee, "Simultaneous arrival to origin convergence: sliding-mode control through the norm-normalized sign function," *IEEE Transactions on Automatic Control*, vol. 67, no. 4, pp. 1966–1972, 2022.

[37] D. Li, S. S. Ge, and T. H. Lee, "Fixed-time-synchronized consensus control of multiagent systems," *IEEE Transactions on Control of Network Systems*, vol. 8, no. 1, pp. 89–98, 2021.

[38] L. Zhang, T. Gao, G. Cai, and K. L. Hai, "Research on electric vehicle charging safety warning model based on back propagation neural network optimized by improved gray wolf algorithm," *Journal of Energy Storage*, vol. 49, Article ID 104092, 2022.

[39] W. Yang, X. Chen, Z. Xiong, Z. Xu, G. Liu, and X. Zhang, "A privacy-preserving aggregation scheme based on negative survey for vehicle fuel consumption data," *Information Sciences*, vol. 570, pp. 526–544, 2021.

[40] R. Sun, J. Wang, Q. Cheng, Y. Mao, and W. Y. Ochieng, "A new IMU-aided multiple GNSS fault detection and exclusion algorithm for integrated navigation in urban environments," *GPS Solutions*, vol. 25, no. 4, pp. 147–217, 2021.

[41] R. Sun, Z. Zhang, Q. Cheng, and W. Y. Ochieng, "Pseudorange error prediction for adaptive tightly coupled GNSS/IMU navigation in urban areas," *GPS Solutions*, vol. 26, no. 1, pp. 28–13, 2022.

[42] Y. Wang, H. Wang, B. Zhou, and H. Fu, "Multi-dimensional prediction method based on Bi-LSTMC for ship roll," *Ocean Engineering*, vol. 242, Article ID 110106, 2021.

[43] T. Ni, D. Liu, Q. Xu, Z. Huang, H. Liang, and A. Yan, "Architecture of cobweb-based redundant TSV for clustered faults," *IEEE Transactions on Very Large Scale Integration Systems*, vol. 28, no. 7, pp. 1736–1739, 2020.

[44] Z. Wang, R. Ramamoorthy, X. Xi, and H. Namazi, "Synchronization of the neurons coupled with sequential developing electrical and chemical synapses," *Mathematical Biosciences and Engineering*, vol. 19, no. 2, pp. 1877–1890, 2021.

[45] Z. Zhang, J. Tian, W. Huang, L. Yin, W. Zheng, and S. Liu, "A haze prediction method based on one-dimensional convolutional neural network," *Atmosphere*, vol. 12, no. 10, Article ID 1327, 2021.

[46] K. Shang, Z. Chen, Z. Liu et al., "Haze prediction model using deep recurrent neural network," *Atmosphere*, vol. 12, no. 12, Article ID 1625, 2021.

[47] L. Yin, L. Wang, W. Huang et al., "Haze grading using the convolutional neural networks," *Atmosphere*, vol. 13, no. 4, Article ID 522, 2022.

[48] Y. Liu, Z. Zhang, X. Liu, L. Wang, and X. Xia, "Efficient image segmentation based on deep learning for mineral image classification," *Advanced Powder Technology*, vol. 32, no. 10, pp. 3885–3903, 2021.

[49] Y. Liu, Z. Zhang, X. Liu, L. Wang, and X. Xia, "Ore image classification based on small deep learning model: evaluation and optimization of model depth, model structure and data size," *Minerals Engineering*, vol. 172, Article ID 107020, 2021.

[50] W. Lei, Z. Hui, L. Xiang, Z. Zelin, X. Xu-Hui, and S. Evans, "Optimal remanufacturing service resource allocation for generalized growth of retired mechanical products: maximizing matching efficiency," *IEEE Access*, vol. 9, pp. 89655–89674, 2021.

[51] L. Guo, C. Ye, Y. Ding, and P. Wang, "Allocation of centrally switched fault current limiters enabled by 5G in transmission system," *IEEE Transactions on Power Delivery*, vol. 36, no. 5, pp. 3231–3241, 2021.

[52] J. Zhang, C. Zhu, L. Zheng, and K. Xu, "ROSEFusion: random optimization for online dense reconstruction under fast camera motion," *ACM Transactions on Graphics*, vol. 40, no. 4, pp. 1–17, 2021.

[53] S. Zhao, F. Li, H. Li et al., "Smart and practical privacy-preserving data aggregation for fog-based smart grids," *IEEE Transactions on Information Forensics and Security*, vol. 16, pp. 521–536, 2021.

[54] J. Li, K. Xu, S. Chaudhuri, E. Yumer, H. Zhang, and L. Guibas, "Grass: generative recursive autoencoders for shape structures," *ACM Transactions on Graphics*, vol. 36, no. 4, pp. 1–14, 2017.

Hindawi
Computational Intelligence and Neuroscience
Volume 2022, Article ID 2538896, 12 pages
https://doi.org/10.1155/2022/2538896

Research Article

The Fusion Application of Deep Learning Biological Image Visualization Technology and Human-Computer Interaction Intelligent Robot in Dance Movements

Nian Jin,[1] Lan Wen,[2,3] and Kun Xie[4]

[1]*College of Music and Dance, Guangzhou University, Guangzhou 510006, China*
[2]*South China Business College, Guangdong University of Foreign Studies, Guangzhou 510545, China*
[3]*Faculty of Education, Guangxi Normal University, Guilin 535400, China*
[4]*Chongqing College of Humanities Science and Technology, Chongqing, China*

Correspondence should be addressed to Kun Xie; 15020240133@xs.hnit.edu.cn

Received 3 August 2022; Accepted 6 September 2022; Published 20 September 2022

Academic Editor: Ning Cao

The paper aims to apply the deep learning-based image visualization technology to extract, recognize, and analyze human skeleton movements and evaluate the effect of the deep learning-based human-computer interaction (HCI) system. Dance education is researched. Firstly, the Visual Geometry Group Network (VGGNet) is optimized using Convolutional Neural Network (CNN). Then, the VGGNet extracts the human skeleton movements in the OpenPose database. Secondly, the Long Short-Term Memory (LSTM) network is optimized and recognizes human skeleton movements. Finally, an HCI system for dance education is designed based on the extraction and recognition methods of human skeleton movements. Results demonstrate that the highest extraction accuracy is 96%, and the average recognition accuracy of different dance movements is stable. The effectiveness of the proposed model is verified. The recognition accuracy of the optimized F-Multiple LSTMs is increased to 88.9%, suitable for recognizing human skeleton movements. The dance education HCI system's interactive accuracy built by deep learning-based visualization technology reaches 92%; the overall response time is distributed between 5.1 s and 5.9 s. Hence, the proposed model has excellent instantaneity. Therefore, the deep learning-based image visualization technology has enormous potential in human movement recognition, and combining deep learning and HCI plays a significant role.

1. Introduction

Modern technologies, such as the Internet and multimedia technology, have developed rapidly. Multimedia systems based on computer information technology have been applied in many fields. The intelligent interactive multimedia is a new platform that develops under the foundation of computer technology [1, 2]. However, the applications of traditional multimedia systems are often independent and mechanized, which are inadequate to meet people's needs. Consequently, the human-computer interaction (HCI) technology emerged. People can interact with the multimedia engine and obtain the required media information quickly and efficiently via HCI. Besides, HCI technology can promote the accurate transmission of information and improve work efficiency [3–5], which has triggered a research boom. In daily life, people can directly express their thoughts or emotions through movements. Therefore, movement recognition and analysis have become a critical direction in the field of HCI and attracted widespread attention, which leads to the wide popularity of human movement-based recognition technology [6, 7].

With the advancement of social informatization, human beings have an increasing requirement in the intelligence level of computers. HCI no longer only depends on the original hardware-based interaction, and some relatively more intelligent interaction methods gradually appear in mass life. The face recognition, gesture recognition, and speech recognition

systems constructed by machine learning technology have established a bridge between humans and computers [8]. The emergence of these convenient interaction modes has become a major development trend in the field of HCI. The development of HCI mode aims to enable the computer to serve and adapt to human needs well, so HCI focuses on humans instead of adapting to the computer. Therefore, the friendly interaction between robots and humans is extremely vital in the research of machine learning and HCI. Some scholars focus on the importance of emotional factors related to the interaction between people and computer systems, when exploring the people-centered interaction systems [9]. Motion recognition technology is essentially a classification problem close to machine learning [10].

The above research results imply that the development of the Internet and multimedia technology has made multimedia systems successfully applied to many fields. The friendly interaction between robots and human beings plays an extremely important role in the study of machine learning and HCI. Deep learning shows excellent application potential in function extraction and HCI. A combination of deep learning and HCI is innovatively proposed to extract and identify human skeleton operations to expand the application field of HCI. The ultimate research purpose is to achieve a significant reduction in time costs and dependence on traditional equipment and facilities. The innovative ideas can also achieve the purpose of improving human-computer collaboration and interaction. Moreover, combined with the image visualization technology based on deep learning and HCI system, it is envisaged that the visual geometric group network (VGGNet) and long short-term memory (LSTM) can be optimized. The final HCI system and the research results of the recognition and analysis of human dance provide a reference value.

The contributions based on the extraction and recognition of human dance movements are as follows:

(1) An optimized VGGNet human skeleton movement extraction algorithm is proposed. Its extraction accuracy reaches 96%, which is significantly better than traditional algorithms.

(2) An optimized multiple LSTM human skeleton movement recognition algorithm is proposed. Its recognition accuracy reaches 88.9%, which is significantly better than traditional LSTMs.

(3) An HCI system based on image visualization is designed, and the interaction accuracy rate reaches 92%.

(4) A reference is provided for more in-depth human movement extraction and recognition, and deep learning methods' application range in HCI systems is expanded.

2. Literature Review

2.1. Current Situation of Deep Learning in Dance Education. Dance is an important intangible cultural heritage. Dimitropoulos et al. introduced a research project (i-juries) of intangible cultural heritage, emphasizing the importance of 3D dance interaction [11]. Grammalidis et al. introduced an intangible cultural heritage dataset, i-treasure, including audio and other data information [12]. Doulamis et al. considered that intangible cultural heritage was an important source of cultural diversity, but there were few electronic documents of intangible cultural heritage. According to the "Terpsichore" project funded by the Horizon 2020 of the European Union, they proposed a high-level method based on the digitization of cultural assets [13]. Doulamis et al. discussed the digitization of tangible and intangible cultural heritage and proposed that 3D digital assets would develop into a part of augmented, virtual, and mixed reality experience [14]. Lv studied the application of virtual reality (VR) in 3D environment and HCI system and revealed the excellent performance of VR technology in 3D digitization [15]. The digitization of intangible cultural heritage has become an inevitable development trend, so has dance.

On the recognition and extraction of dance movements, Rallis et al. proposed a dance summarization method based on 3D capture data of the Vicon motion capture system. They analyzed and studied the automatic extraction of dance patterns. This method was a hierarchical scheme based on the temporal and spatial changes of dance characteristics [16]. Aiming at the preservation and dissemination of dance performance, Aristidou et al. proposed a dance action recognition framework based on Laban analysi which used feature space to capture different dance action components and pointed out a new direction for dance evaluation [17]. In terms of editing and synthesis of dance movements, Aristidou et al. used Laban analysis, radial basis function regression, and interpolation methods to map the movement features and emotional features in two directions and realized the stylization of high dynamic dance movements [18]. To sum up, there is a difference between the research of human action recognition and HCI, and there is little research on action recognition in dance education.

2.2. Research Progress of HCI. Experts and scholars have made great efforts on deep learning and HCI. Bhardwaj et al. applied support vector machine and artificial neural network classifier to fingerprint recognition. By integrating the relevant dynamic information from hundreds of biometric scanning sample datasets, they found that the accuracy of fingerprint dynamic recognition by fusing the deep learning method was improved by 5.3% [19]. Israelsen and Ahmed analyzed the influence of artificial intelligence (AI) agent in HCI and machine learning based on the research of algorithm-guaranteed AI agent and discussed the advantages and disadvantages of different methods [20]. Based on similarity embedding, Spathis et al. proposed an interactive dimension reduction framework (iSP). In this framework, user interaction formed different goals. Gradient descent was used for learning, and an end-to-end composition structure could be trained. By evaluating the framework in two interaction

scenarios, they found that the framework could be applied to semisupervised learning, transfer learning, and adaptive learning in interaction field [21]. Using interactive machine learning, Wu et al. studied local decision-making in feature selection of emotion classification task and analyzed the influence of interactive machine learning tools on feature selection results [22]. To improve the performance of multimodal image retrieval by using unmarked and marked multimodal web objects, Xu et al. proposed a semisupervised multiconcept retrieval method based on deep learning (SMRDL). Different from the traditional method of using multiple independent concepts in multiconcept semantic query, the proposed method regarded multiple concepts as a whole scene, which was used for multiconcept scene learning of unimodal retrieval. The comprehensive experimental results on two datasets of MIR flickr2011 and NUS-WIDE indicated that the proposed method was superior to some of the latest methods [23]. Long and Zhao held that intelligent teaching mode overcame the shortcomings of traditional online and offline teaching. However, there were some shortcomings in the real-time feature extraction of teachers and students. In view of this, they used particle swarm image recognition and deep learning technology to process the video teaching image of intelligent classroom. To overcome the shortcomings of premature convergence of standard particle swarm optimization (PSO) algorithm, they proposed an improved multi PSO algorithm strategy. Moreover, to improve the premature problem of PSO in search performance, they combined the algorithm with the useful attributes of other algorithms to improve the diversity of particles in the algorithm, enhance the global search ability of particles, and achieve effective feature extraction [24]. To sum up, there are many research results on the application of deep learning in HCI, but few studies on the combination of the two for dance action extraction.

3. Methods

In computer vision and image processing, movement recognition is a crucial component. However, some problems are found in its research and applications. For example, when extracting and recognizing human skeleton movements, bone modeling is challenging, movement amplitude can affect the extraction results, and feature extraction can be insufficient, increasing the difficulty in analyzing and classifying human movements. Deep learning has developed rapidly. CNN shows excellent performance in feature extraction, while LSTM has significant performance in processing time sequence problems. Therefore, CNN and LSTM are introduced to extract and recognize human skeleton movements. However, traditional CNN models have lots of parameters, using a large convolution kernel to extract features. Traditional LSTM models never consider the connection of multiple different movement times in a long time. Hence, the CNN-based VGGNet is introduced and optimized in parallel. In the meantime, LSTM is improved and optimized before extracting and recognizing human skeleton movements.

3.1. Optimization of VGGNet CNN Model. Cat's visual cortex theory inspires the deep learning-based CNN. Compared with the traditional neural network, CNN extracts the object's local feature information through the convolution layer, a critical CNN component that contains multiple convolution kernels [25]. VGGNet is a typical CNN. Unlike traditional CNNs that employ big convolution kernels to extract features, VGGNet utilizes several $3 * 3$ small convolution kernels for feature extraction. Hence, VGGNet can extract richer features and reduce the calculation amount significantly [26–28].

The features extracted by the convolution layer are integrated to improve the accuracy of VGGNet, i.e., the parallel CNN [29–31].

Extractions of input image features before fusion are as follows:

$$y_1^A = F_1^A(x), \tag{1}$$

$$y_1^B = F_1^B(x). \tag{2}$$

In (1) and (2), F_1^A and F_1^B represent features. The feature information extracted by the two small convolution kernels is fused via the feature fusion module. The convolution operation is denoted as G. The feature map after fusion processing can be written as follows:

$$y_1^c = G_1\left(y_1^A, y_1^B\right) = G_1\left(F_1^A(x), F_1^B(x)\right). \tag{3}$$

The process of fusion of the above feature maps y_1^c, y_1^A, and y_1^B can be expressed as follows:

$$y_1^c = merge\left(y_1^A, y_1^B\right). \tag{4}$$

The above fusion processing can enrich and diversify the extracted features. Graphics Processing Unit (GPU) processing is utilized for training VGGNet to compare the performance of the CNN-based VGGNet before and after optimization. Images in the training set are taken by the Kinect camera and the host computer program. The selected human movements include clapping, slapping, standing, picking up objects, and sitting down.

Movement capture includes the following steps: (1) the demonstrator makes different movements in front of the Kinect camera and (2) Kinect is utilized for evaluating human skeleton changes in real-time. Several demonstrators complete the collection of the entire training set. One thousand images are collected for each movement. Finally, a total of 5,000 human skeleton images under different movements are obtained. The skeleton images affected by the environment are removed, and the remaining human skeleton images are retained. These images train the VGGNet before and after feature fusion. Accuracy and loss rates are taken as evaluation indicators [32, 33]. Parameter settings of the entire training process are shown in Table 1.

3.2. Extraction Algorithm of Human Skeleton Movements. Traditional human pose estimation algorithms extract human skeleton features via the bottom-up manner. Each skeleton extraction object requires a detector, and each

Table 1: VGGNet training parameter settings.

Parameters	Training times	Learning rate	Number of images read	Optimizer
Corresponding value	500	10^{-5}	32	Adam

movement is estimated separately. Therefore, traditional algorithms have many problems, such as false detection, long-running time, and poor instantaneity, which cannot meet the demands. Based on the OpenPose open-source database [34], the optimized VGGNet is the network architecture, and the histogram equalization [35, 36] is introduced to suppress noises, thereby extracting the 2D features of the human skeleton.

OpenPose is an open-source database released in 2017 based on skeleton extraction. Unlike traditional pose estimation algorithms, OpenPose uses a bottom-up method. The joint points of all human body parts are detected first. Then, the nodes are connected to obtain the skeleton, thereby significantly reducing the running time. Also, OpenPose can improve detection accuracy and shorten the running time. Figure 1 illustrates the video information processing by OpenPose.

The unique convolution kernel structure in the CNN can learn spatial information in human actions, and more useful information can be obtained by different convolution kernels. Compared with traditional machine learning methods, CNN is more systematic and comprehensive in task learning with better performances. Unlike traditional CNN models, the VGGNet model extracts features by massive small convolution kernels as a typical CNN model. It can extract more features and reduce calculation amount with satisfactory generalization performance. The optimized VGGNet consists of three parts. The first part processes the image data via the input layer and employs CNN to extract the feature values of body parts. Then, the extracted feature values enter the other two parts for critical point positioning and the body-based 2D vector field positioning. The input to output via the neural network spends a total of k periods, and the information input to the current period is the output feature value obtained through the learning process of $k-1$. The optimized VGGNet's output is formed by a 2D vector field of crucial body parts and a confidence map. As the calculations increase, the candidate human body parts and the corresponding structure division become apparent via this cyclic process. Here, CNN's first convolutional layer is a double convolutional layer, and each contains 64 convolution kernels in the size of $4 * 4$. Simultaneously, an activation layer and a normalization layer are added after each convolutional layer to process the nonlinear data. A pooling layer is added after the normalization layer to reduce dimensionality and prevent overfitting, located between the two convolutional layers. The Dropout layer comes after the second pooling layer. The Part Affinity Fields (PAFs) [37, 38] are adopted to predict all the human body key points in the images.

In summary, extracting human skeleton information includes the following two processes: first, adding the corresponding image data to the input layer of VGGNet and, second, learning the feature value F according to the body parts. The 2D vector field of output corresponding to the human body in the $k=1$ period is

$$S^t = \rho^t\left(F, S^{t-1}, L^{t-1}\right), \forall t \geq 2, \tag{5}$$

$$L^t = \phi^t\left(F, S^{t-1}, L^{t-1}\right), \forall t \geq 2. \tag{6}$$

In (5) and (6), S represents the set of 2D position confidence maps, ρ and ϕ denote the set parameters, t refers to the period corresponding to the feature value, and L signifies the set of 2D vector fields.

The solution to the confidence in the confidence map can be presented as follows:

$$S^*_{j,k}(p) = \exp\left(-\frac{\left\|p - x_{j,k}\right\|_2^2}{\sigma^2}\right), \tag{7}$$

$$S^*_j(p) = \max_k S^*_{j,k}(p). \tag{8}$$

In (7) and (8), S represents the position confidence atlas and p denotes the output image in the corresponding period. Meanwhile, k refers to the number of people in the input image, j stands for the body part's serial number, and σ is a constant.

The joint point position in the 2D vector field is judged according to

$$L^*_{c,k}(p) = \begin{cases} v, \\ 0, \end{cases} \tag{9}$$

$$v = \frac{\left(x_{j2,k} - x_{j1,k}\right)}{\left\|x_{j2,k} - x_{j1,k}\right\|_2}. \tag{10}$$

In (9) and (10), p represents the pixel of the prejudgment part and v denotes the unit vector. On this basis, the average value of the 2D vector field can be written as follows:

$$L^*_c(p) = \frac{1}{n_{c(p)}} \sum_k L^*_{c.k}(p). \tag{11}$$

In (11), $n_{c(p)}$ represents the number of all points of the pixel p on the link c. After testing, candidate positions on PAFs should be determined first. Then, all connected line segments are determined.

The OpenPose open-source library can achieve excellent results of skeleton extraction. However, the image noise limits feature extraction. Therefore, histogram equalization is introduced, which enhances the contrast and reduces the noise by stretching the distribution range of pixel intensity. Videos based on image visualization are processed by Compute Unified Device Architecture (CUDA) to ensure the instantaneity of information extraction. Eighteen key part points are chosen as the input of skeleton movement

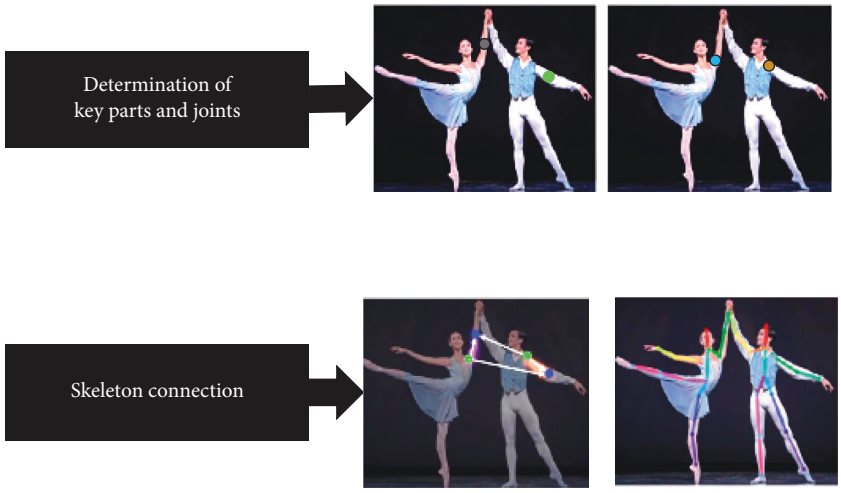


Figure 1: Video information extraction and processing based on open pose (Image URL: https://img-blog.csdnimg.cn/20200910162313905png?x-oss-processimage/watermark, type _ZmFuZ3poZW5naGVpdGk, shadow_10, text_aHR0cHM6Ly9ibG9nLmNzZG4ubmV0L3N1bm55YmxvZ3M, size_16, color_FFFFFF, t_70. Copyright statement URL: https://wwwcsdnnet/company/indexhtml#statement).

extraction while utilizing the OpenPose open-source library. A variable-view movement database containing 40 kinds of aerobic exercises is chosen for analyzing algorithm extraction effects. Eight different movements are chosen for analysis, with the classification accuracy as the primary evaluation indicator.

Here, the optimized 3D CNN (O-3DCNN) algorithm, Spatial-Temporal CNN(ST-CNN) algorithm, and optimized Deformable Part Model CNN (ODPM-CNN) are compared with the optimized VGGNet to prove its effectiveness.

3.3. Skeleton Movement Recognition Based on Optimized LSTM. Traditional neural networks have major limitations in practical application. For example, in time series processing, traditional methods perform well only in short-time series processing. In the separate data processing, the good learning and understanding abilities enable CNN to be applied in practice. However, CNN has limitations in the sequence problem processing related to time correlation. LSTM is a unique Recurrent Neural Network (RNN). LSTM can solve the long-term dependence problem in RNN applications, which has an inseparable relationship with the particular gate structure of LSTM, explicitly referring to input gates, forget gates, and output gates. The input data are calculated according to the following equation:

$$f_t = \sigma\left(w_f\left[h_{t-1}, x_t\right] + b_f\right). \tag{12}$$

In (12), w represents the weight, b corresponds to the deviation, and h_{t-1} denotes the output value corresponding to the time $t-1$. Meanwhile, x_t refers to the input value, σ represents the activation function, and f stands for the forget gate. Moreover, the memory information c_t can be displayed as follows:

$$c_t = f_t c_{t-1} + j_t \tanh\left(w_c \cdot \left[h_{t-1}, x_t\right] + b_c\right). \tag{13}$$

In (13), c_{t-1} represents deciding whether to memorize the information at the time $t-1$ and j_t means the input gate.

Finally, the output gate o_t can be expressed as follows:

$$o_t = \sigma\left(w_o \cdot \left[h_{t-1}, x_t\right] + b_o\right). \tag{14}$$

Although LSTM has many excellent performances, LSTM does not consider the correlation and feature influence between different skeleton movements over a long time. Hence, the LSTM model only depends on the human skeleton joints while recognizing human skeleton movements, resulting in limitations to recognizing human skeleton movements. Therefore, the idea of time integral is introduced. First, the pre-acquired skeleton sequence information is transformed, such as translation and rotation. In this way, all movements can obtain their relative coordinates. If the human skeleton movement has differences due to different times, a multiple LSTM model is used to extract and fuse features [39]. Finally, multiple types of movements are captured by integrating multiple LSTMs. Figure 2 reveals the overall implementation framework of the optimized multi-LSTM human skeleton movement recognition.

Extraction accuracy and loss entropy of various LSTMs are compared to verify the effectiveness of the optimized multi-LSTM human skeleton movement recognition algorithm. Specifically, algorithms selected for comparison include the single-LSTM and double-LSTM. A skeleton sequence input into the optimized F-Multi-LSTM contains 24 frames, among which each frame consists of multiple 2D skeleton points. During analysis, the Adam optimization algorithm is used as the optimization tool, and the initial learning rate is set to 10^{-4}, in an effort to achieve the model's global optimization. The single-LSTM has one input layer, while the double-LSTM has two input layers. The input is assumed as a sentence. In double-LSTM, one side of the input corresponds to the word at the beginning of the sentence and the other side corresponds to the word at the end of the sentence.

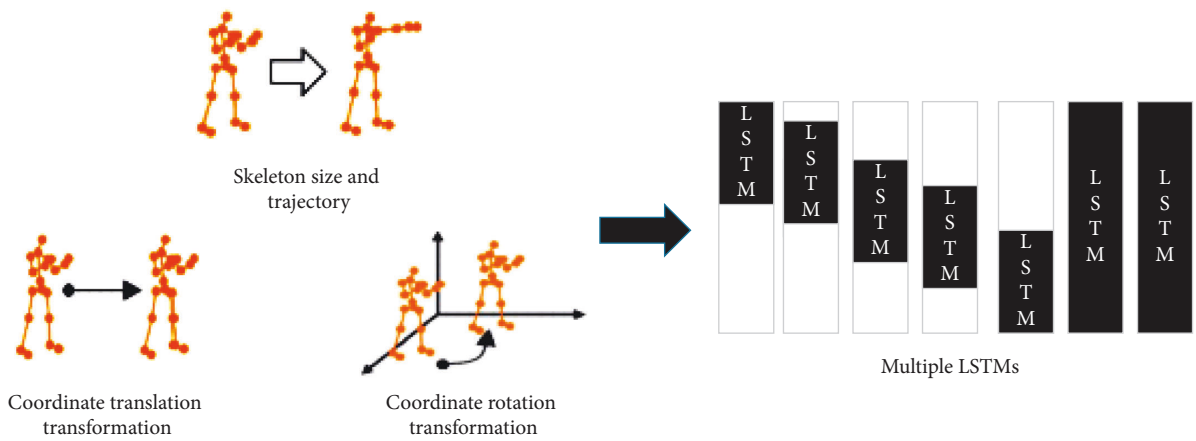


Figure 2: Overall implementation framework of multi-LSTM human skeleton movement recognition.

3.4. Design of HCI System Based on Dance Education. Dance education based on physical education helps improve students' physical fitness and transforms traditional sports teaching. According to the above image visualization-based extraction and recognition method of human skeleton movements, the Web3D engine-oriented deep movement recognition system's functional modules are shown in Figure 3.

The system based on dance education and dance movement recognition consists of the front-end interactive function module and the back-end recognition function module. The former is a 3D world built on Web Graphics Library (WebGL) technology, including data processing of video images, 3D processing, and the HCI submodule. The latter consists of two subfunction modules, namely, node recognition and classification of human dance movements.

In this HCI system, the OpenPose open-source database and optimized VGGNet model can estimate facial expressions, positioning of limbs and trunk, and people's feature information. This human skeleton extraction method can identify the critical points of the human body, thereby employing the optimized F-Multi-LSTM skeleton movement recognition network to determine the classification and label attribution of human dance movements. The designed system is based on recognizing and analyzing dance movements. Eight types of dance movements are analyzed and discussed, including stepping and knee lift (S), crouching (C), reaching out and jumping (R), turning and clapping (T), straight punch (B), arm circles (A), jumping (J), and high knee (H).

In the HCI system, the dance pose estimation module and dance movement classification module in the background recognition module are the keys. Accuracy and response time are evaluation indicators to analyze the chosen dance movements, thereby testing the feasibility of the HCI system based on dance education and movement analysis and recognition.

3.5. Data Preprocessing. The image is preprocessed as follows to better meet the needs of behavior recognition: first, the image is uniformly scaled to 432×368 based on the center point; second, image denoising. Noises are common in images, in which Gaussian noise is the most common one. The Gaussian filter is used for processing to effectively suppress the Gaussian noise in the image. The one-dimensional Gaussian distribution and two-dimensional Gaussian distribution are shown in (15) and (16), respectively. The Gaussian filter function in open-source computer vision library (Open CV) is used to realize image denoising, and the relevant parameters are optimized.

$$G(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{x^2}{2\sigma^2}}, \tag{15}$$

$$G(x, y) = \frac{1}{\sqrt{2\pi}\sigma^2} e^{-\frac{x^2 + y^2}{2\sigma^2}}. \tag{16}$$

4. Results

This section analyzes the optimized VGGNet algorithm's performance through comparison with several human skeleton movement extraction algorithms. The accuracy of the VGGNet algorithm in human skeleton movement extraction is analyzed and optimized on this basis. The effectiveness of the optimized model is verified. Besides, comparative analysis is conducted on the performance of the LSTM model, the single-LSTM model, and the double-LSTM model. Finally, the interaction accuracy and system real-time performance shall prevail to verify the HCI dance education system's performance.

4.1. Performance Comparison of Skeleton Movement Extraction Algorithms. Table 2 presents the comparison result of the extraction accuracy of human movements by several algorithms, including the original and optimized VGGNet.

Table 2 suggests that the optimized VGGNet algorithm presents the best performance in extracting human movements, with the highest accuracy of 98.2%, showing apparent superiority in performance over traditional VGGNet algorithms. The 3D CNN model can only extract a

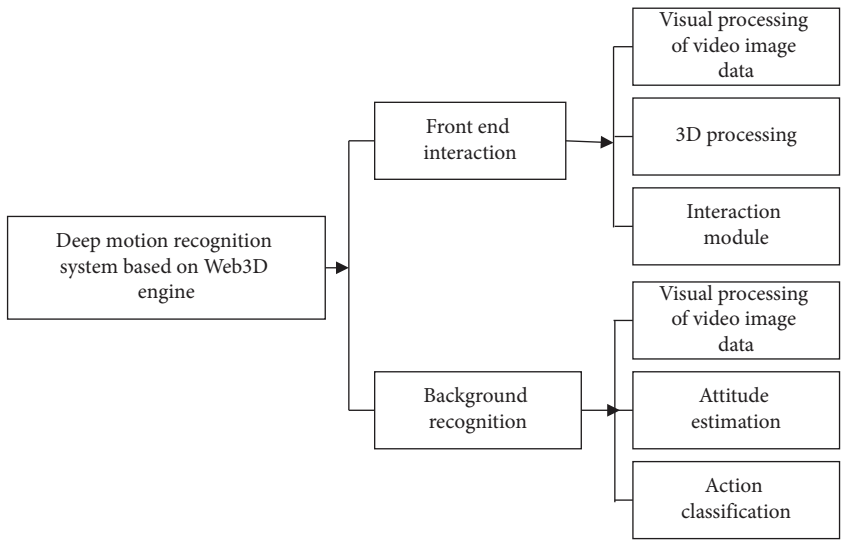


Figure 3: Functional modules of the deep motion recognition system.

Table 2: Comparison of several algorithms' extraction accuracy.

Algorithms	Original VGGNet	Optimized VGGNet	O-3DCNN	ST-CNN	ODPM-CNN
Extraction accuracy (%)	96.9	98.2	91.2	90.5	97.08

type of features from a three-dimensional space because the weights of the convolution kernel are the same in the whole space; that is, the weights are shared by the same convolution kernel, so the extraction accuracy of 3D CNN is only 91.2%. The spatial invariance of ST-CNN refers to the invariance of spatial transformation of images such as rotation, translation, and scaling. Even if the input is transformed or slightly modified, the model can recognize and extract features. ST-CNN is the most time-consuming and error-prone place in debugging interpolation and image index, so the extraction accuracy of ST-CNN is only 90.5%. ODPM-CNN model is a variability network and ODPM-CNN just the opposite, and its recognition accuracy reached 97.08%. The optimized VGGNet is also superior to other human movement extraction algorithms. In this way, the effectiveness of the proposed skeleton extraction algorithm is verified preliminarily.

4.2. Extraction Results of Human Skeleton Movements. The accuracy distribution of the eight human skeleton movements' extraction results by optimized VGGNet on OpenPose open-source database is shown in Figure 4.

This collection of 100 dance pictures is seen as a total sample, and each picture contains eight parts of the action changes. *S* represents the step and knee lifting head, shoulders, elbows, wrists, hips, knees, ankle bone node extraction accuracy; other *C*, *R*, *T*, *B*, *A*, *J*, and *H* dataset content for the above eight parts of the extraction accuracy changes under the action of the title annotation. The extraction accuracy of the head is the highest, reaching 96%, and 100 images are correctly extracted. The extraction accuracy of the shoulder reaches 84.8%, with 90 pictures extracted correctly. The extraction accuracy of the elbow

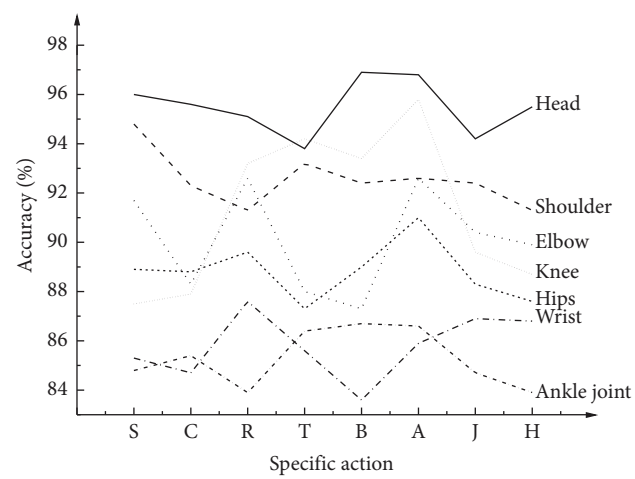


Figure 4: Accuracy distribution of movement extraction results of the human skeleton. (Labels on the *x*-axis represent dance movements. *S* represents stepping and knee lift, *C* represents crouching, *R* stands for reaching out and jumping, *T* stands for turning and clapping, *B* denotes straight punch, *A* denotes arm circles, *J* refers to jumping, and *H* refers to high knee).

reaches 92.6%, with 89 pictures extracted correctly. The extraction accuracy of the wrist reaches 87.6%, with 86 pictures correctly extracted. The extraction accuracy of the hip reaches 91.0%, with 100 pictures extracted correctly. The extraction accuracy of the knee reaches 95.8%, with 90 pictures extracted correctly. The extraction accuracy of the ankle reaches 86.7%, with 88 pictures extracted correctly. Figure 4 signifies that the extraction accuracy of bone nodes in eight body parts is different, and the proportion of sample number is also different. Moreover, Figure 4 implies that the

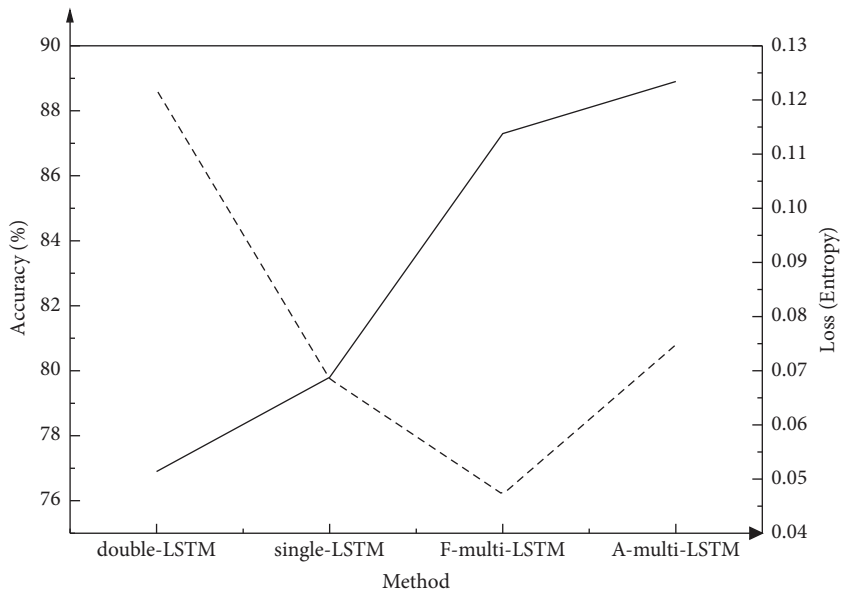


Figure 5: Comparison results of LSTM models.

proportion of accurate number extracted from the large part of the space occupied by the body parts will be significantly higher.

4.3. Skeleton Movement Recognition Results of Multiple LSTMs. The single-LSTM, double-LSTM, F-multi-LSTM, and A-multi-LSTM are compared. The results are shown in Figure 5.

The parameters represented by the abscissa in Figure 5 are different neural network models. The corresponding left-axis variables refer to the accuracy, and the corresponding right-axis variables stand for loss rates. Single-LSTM is a sequence that supports one-way variable input and output, while double-LSTM is a sequence that supports two-way input and output. Multi-LSTM is a multidimensional LSTM for high-frequency time series, which supports multiple parallel input sequences with multiple inputs, rather than the planar structure of multiple inputs in other models. F-Multi-LSTM is an optimized multidimensional LSTM, and A-Multi-LSTM is expressed as a pair of optimized multidimensional LSTM. The double-LSTM has higher accuracy than the single-LSTM according to the comparison results of loss rate and accuracy of single-LSTM and multi-LSTM. The recognition accuracy reaches 79.8%, and the loss rate is 0.0685. Compared with the single-LSTM model, the difference is 43.8%; overall, the recognition accuracy and loss rate of the proposed multi-LSTM model are the best. Specifically, the single-LSTM model's recognition accuracy reaches 88.9%, and the loss rate is 0.0748, which is the best among the comparative algorithms. Compared with the traditional LSTM model before improvement, the optimized LSTM model has higher recognition accuracy. The optimized LSTM model has the best applicability in recognizing human skeleton movements.

4.4. HCI System Performance Based on Dance Education and Movement Analysis. The eight dance movements are chosen as the benchmark. According to the indicators of interaction accuracy and system instantaneity, the HCI system's performance for dance education is shown in Figure 6.

In the dance education HCI system, the eight dance movements' overall interaction accuracy is above 70%. The interaction accuracy of movement B is the highest, reaching 92%. The overall accuracy of interactive recognition is distributed in the range of 72%–92%, with a large span. The overall response time corresponding to the eight dance movements is distributed within 5.1 seconds to 5.9 seconds, showing that the dance education HCI system has a high instantaneity.

5. Discussion

The above results indicate changes in the OpenPose open-source database's recognition accuracy and the optimized VGGNet model. The reason is that the head has almost no changes in coordinates or rotation angle. Besides, the movement range of the head is small. Therefore, the accuracy of classification and recognition of the head is the highest. In contrast, the shoulders are greatly affected by external factors, such as rotation angle and abscissa among different movements. Hence, classification and recognition accuracy of the shoulders are relatively low. The elbow movements and the wrist movements are affected by changes in moving speed and longitudinal coordinates. If the human body's moving speed is slow and the position between the arm and the camera is not parallel, the classification and recognition accuracy will be high. The hips are easily affected by changes in the leg movements. The overall accuracy of classification and recognition corresponding to the knees is high, but movements with large fluctuations, such as movement *H*, can significantly affect classification and recognition accuracy. Therefore, the accuracy is low. The ankles and other

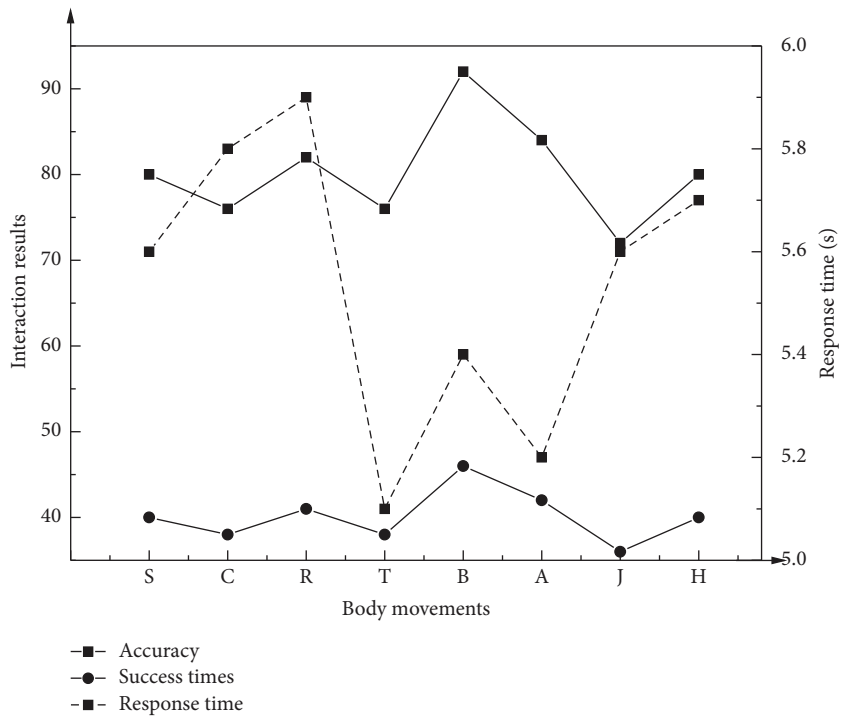


Figure 6: HCI system performance in terms of accuracy and instantaneity.

parts' classification and recognition accuracies are low, probably because of external factors such as clothes and shoes. Although the classification and recognition accuracy of different dance movements are mainly different, the average accuracy is high, confirming the proposed algorithm's effectiveness.

The multiple LSTM model is also advantageous in skeleton movement recognition. Because the optimized LSTM model is robust, its learning and classification abilities are increased, thereby increasing its accuracy in recognizing different dance movements. The distribution changes of the interaction accuracy corresponding to the dance education HCI system reveal that the interaction accuracy corresponding to different movements has a large span. Compared with the model training process, the actual interaction will be affected by sophisticated environmental conditions, such as different lighting, the restraint between different dance movements, and the conversion frequency of various dance movements. Under sophisticated environmental conditions, the interaction accuracy of the HCI system drops. Hence, attention should also be paid to improve datasets in actual HCI applications.

Meanwhile, the proposed algorithm is compared with the methods proposed by other scholars [40–49] to verify its superiority. For the training, the input image size is set to 432×368, the number of cycles is set to 50, the batch size is set to 16, and the initial learning rate is set to 0.001. Table 3 reflects the results. Table 3 demonstrates that the proposed multi-LSTM model has the highest accuracy in bone motion recognition, and the recognition accuracy has been improved by 27.79%, 17.69%, and 27.62%, respectively, compared with the comparative methods.

Table 3: Comparison of experimental results of different models.

Algorithms	Accuracy (%)
Method 1 [40]	61.11
Method 2 [41]	71.21
Method 3 [42]	61.28
Multi-LSTM model	88.9

6. Conclusions

For the dance education HCI system, the CNN-based VGGNet model is optimized and applied to extract human skeleton movements based on the OpenPose open-source database and histogram equalization. The proposed extraction algorithm for human skeleton movements shows intentional performance in extracting eight different dance movements, with the highest accuracy rate reaching 96%. From the comparison results of loss value and accuracy between a single-LSTM model and a multi-LSTM model, the accuracy of bone motion recognition by the multi-LSTM model is 79.8%, which is higher than that by a single-LSTM model. The optimized multi-LSTM model has higher accuracy in recognizing human skeleton movements than the traditional LSTM models. The constructed HCI system has an interaction accuracy of 92%. This work achieves the extension of application range of deep learning in skeleton movement recognition and the organic combination of deep learning and HCI.

The contributions based on the extraction and recognition of human dance movements are as follows:

(1) An optimized VGGNet human skeleton movement extraction algorithm is proposed, which achieves a

better extraction accuracy than traditional algorithms, attaining 96%.

(2) An optimized multiple LSTM human skeleton movement recognition algorithm is proposed. Its recognition accuracy reaches 88.9%, which is significantly better than traditional LSTMs.

(3) A HCI system based on image visualization is designed, with the interaction accuracy rate of 92%.

(4) A reference is provided for more in-depth human movement extraction and recognition, and deep learning strengthens the applicability to the HCI system.

Due to computational resource limitations, other larger and more complex datasets are considered in this experiment [50–55]. In addition, the algorithm can meet the real-time requirements, the recognition speed is still very slow. In view of the above problems, it is worth further expanding the datasets in complex scenes in the subsequent work and further optimizing the model to improve the detection speed [56].

Limited by the computing resources, other larger and more complex datasets are not explored [57, 58]. In addition, the recognition speed of the algorithm is slow although it can meet the real-time performance [59–61]. In view of the above problems, the dataset will be further expanded, especially in complex scenarios, which further optimizes the model to improve the speed of detection.

Data Availability

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethical Approval

This article does not contain any studies with human participants or animals performed by any of the authors.

Consent

Informed consent was obtained from all individual participants included in the study.

Conflicts of Interest

All authors declare that they have no conflicts of interest.

Authors' Contributions

All authors listed have made a substantial, direct, and intellectual contribution to the work and approved it for publication.

References

[1] S. Kohek, D. Strnad, B. Zalik, and S. Kolmanic, "Interactive synthesis and visualization of self-organizing trees for large-scale forest succession simulation," *Multimedia Systems*, vol. 25, no. 3, pp. 213–227, 2019.

[2] A. Canovas, J. M. Jimenez, O. Romero, and J. Lloret, "Multimedia data flow traffic classification using intelligent models based on traffic patterns," *IEEE Network*, vol. 32, no. 6, pp. 100–107, 2018.

[3] A. Rapp, "Design fictions for learning: a method for supporting students in reflecting on technology in Human-Computer Interaction courses," *Computers & Education*, vol. 145, Article ID 103725, 2020.

[4] M. Hibbeln, J. L. Jenkins, C. Schneider, J. S. Valacich, and M. Weinmann, "How is your user feeling? Inferring emotion through human-computer interaction devices," *MIS Quarterly*, vol. 41, no. 1, pp. 1–21, 2017.

[5] J. Bergstrm and K. Hornbk, "Human--Computer interaction on the skin," *ACM Computing Surveys*, vol. 52, no. 4, pp. 1–14, 2019.

[6] M. R. Malgireddy, I. Nwogu, and V. Govindaraju, "Language-motivated approaches to action recognition," *Journal of Machine Learning Research*, vol. 14, no. 1, pp. 2189–2212, 2017.

[7] S. Singh, C. Arora, and C. V. Jawahar, "Trajectory aligned features for first person action recognition," *Pattern Recognition*, vol. 62, pp. 45–55, 2017.

[8] D. J. Kim, W. K. Song, J. S. Han, and Z. Z. Bien, "Soft computing based intention reading techniques as a means of human-robot interaction for human centered system," *Soft Computing - A Fusion of Foundations, Methodologies and Applications*, vol. 7, no. 3, pp. 160–166, 2003.

[9] L. Y. Mano, B. S. Faiçal, V. P. Gonçalves et al., "An intelligent and generic approach for detecting human emotions: a case study with facial expressions," *Soft Computing*, vol. 24, no. 11, pp. 8467–8479, 2020.

[10] U. Erkan, "A Precise and Stable Machine Learning Algorithm: Eigenvalue Classification (EigenClass)," *Neural Computing and Applications*, vol. 33, no. 10, pp. 5381–5392, 2020.

[11] K. Dimitropoulos, S. Manitsaris, F. Tsalakanidou, and N Spiros, "Capturing the Intangible: An Introduction to the I-Treasures Project," in *Proceedings of the 9th International Conference on Computer Vision Theory and Applications (VISAPP2014)*, IEEE, Lisbon, Portugal, October 2014.

[12] N. Grammalidis, K. Dimitropoulos, F. Tsalakanidou, and A Kitsikidis, "The I-Treasures Intangible Cultural Heritage Dataset," in *Proceedings of the 3rd International Symposium on Movement and Computing*, July 2016.

[13] A. Doulamis, A. Voulodimos, N. Doulamis, and S Soile, "Transforming Intangible Folkloric Performing Arts into Tangible Choreographic Digital Objects: The Terpsichore Approach," in *Proceedings of the 12th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications (VISIGRAPP) 2017, Special Session on Computer Vision, Imaging and Computer Graphics for Cultural Applications*, Porto, Portugal, February 2017.

[14] N. Doulamis, A. Doulamis, C. Ioannidis, and M Klein, "Modelling of Static and Moving Objects: Digitizing Tangible and Intangible Cultural Heritage," *Mixed Reality and Gamification for Cultural Heritage*, Springer International Publishing, Berlin, Germany, 2017.

[15] Z. Lv, "Virtual reality in the context of Internet of things," *Neural Computing & Applications*, vol. 32, no. 13, pp. 9593–9602, 2019.

[16] I. Rallis, N. Doulamis, A. Doulamis, A. Voulodimos, and V. Vescoukis, "Spatio-temporal summarization of dance choreographies," *Computers & Graphics*, vol. 73, pp. 88–101, 2018.

[17] A. Aristidou, E. Stavrakis, P. Charalambous, Y. Chrysanthou, and S. L. Himona, "Folk dance evaluation using laban movement analysis," *Journal on Computing and Cultural Heritage*, vol. 8, no. 4, pp. 1–19, 2015.

[18] A. Aristidou, Q. Zeng, E. Stavrakis, and K Yin, "Emotion control of unstructured dance movements," in *Proceedings of the Acm Siggraph*, pp. 1–10, ACM, Los Angeles, California, July 2017.

[19] I. Bhardwaj, N. D. Londhe, and S. K. Kopparapu, "Performance evaluation of fingerprint dynamics in machine learning and score level fusion framework," *IETE Technical Review*, vol. 36, no. 2, pp. 178–189, 2019.

[20] B. W. Israelsen and N. R. Ahmed, "Dave. I can assure you. That it's going to be all right A definition, case for, and survey of algorithmic assurances in human-autonomy trust relationships," *ACM Computing Surveys*, vol. 51, no. 6, pp. 1–37, 2019.

[21] D. Spathis, N. Passalis, and A. Tefas, "Interactive dimensionality reduction using similarity projections," *Knowledge-Based Systems*, vol. 165, pp. 77–91, 2019.

[22] T. Wu, D. S. Weld, and J. Heer, "Local decision pitfalls in interactive machine learning: an investigation into feature selection in sentiment analysis," *ACM Transactions on Computer-Human Interaction*, vol. 26, no. 4, pp. 1–27, 2019.

[23] H. Xu, C. Huang, and D. Wang, "Enhancing semantic image retrieval with limited labeled examples via deep learning," *Knowledge-Based Systems*, vol. 163, no. JAN.1, pp. 252–266, 2019.

[24] S. Long and X. Zhao, "Smart teaching mode based on particle swarm image recognition and human-computer interaction deep learning," *Journal of Intelligent and Fuzzy Systems*, vol. 39, no. 4, pp. 5699–5711, 2020.

[25] F. Zhang, N. Cai, J. Wu, G. Cen, H. Wang, and X. Chen, "Image denoising method based on a deep convolution neural network," *IET Image Processing*, vol. 12, no. 4, pp. 485–493, 2018.

[26] D. Singh, M Erinc, H Sten, and K Johannes, "Convolutional and recurrent neural networks for activity recognition in smart environment," in *Towards Integrative Machine Learning and Knowledge Extraction: BIRS Workshop, Banff, AB, Canada, July 24-26, 2015, Revised Selected Papers*, A. Holzinger, R. Goebel, M. Ferri, and V. Palade, Eds., Springer International Publishing, Berlin, Germany, pp. 194–205, 2017.

[27] C. C. Wong, Y. Gan, and C. M. Vong, "Efficient outdoor video semantic segmentation using feedback-based fully convolution neural network," *IEEE Transactions on Industrial Informatics*, vol. 16, no. 8, pp. 5128–5136, 2020.

[28] M. Tahir, H. Tayara, and K. T. Chong, "iRNA-PseKNC(2-methyl): identify RNA 2'-O-methylation sites by convolution neural network and Chou's pseudo components," *Journal of Theoretical Biology*, vol. 465, pp. 1–6, 2019.

[29] Q. Yao, R. Wang, X. Fan, J. Liu, and Y. Li, "Multi-class Arrhythmia detection from 12-lead varied-length ECG using attention-based time-incremental convolutional neural network," *Information Fusion*, vol. 53, pp. 174–182, 2020.

[30] P. M. Cheng and H. S. Malhi, "Transfer learning with convolutional neural networks for classification of abdominal ultrasound images," *Journal of Digital Imaging*, vol. 30, no. 2, pp. 234–243, 2017.

[31] A. Ardakani, C. Condo, and W. J. Gross, "Fast and efficient convolutional accelerator for edge computing," *IEEE Transactions on Computers*, vol. 69, no. 1, pp. 138–152, 2020.

[32] M. Hammad and K. Wang, "Parallel score fusion of ECG and fingerprint for human authentication based on convolution neural network," *Computers & Security*, vol. 81, pp. 107–122, 2019.

[33] P. Yao, H. Wu, B. Gao et al., "Fully hardware-implemented memristor convolutional neural network," *Nature*, vol. 577, no. 7792, pp. 641–646, 2020.

[34] Z. Dong, X. Du, and Y. Liu, "Automatic segmentation of left ventricle using parallel end-end deep convolutional neural networks framework," *Knowledge-Based Systems*, vol. 204, Article ID 106210, 2020.

[35] A. Yasoubi, R. Hojabr, and M. Modarressi, "Power-efficient accelerator design for neural networks using computation reuse," *IEEE Computer Architecture Letters*, vol. 16, no. 1, pp. 72–75, 2017.

[36] A. Sellami and H. Hwang, "A robust deep convolutional neural network with batch-weighted loss for heartbeat classification," *Expert Systems with Applications*, vol. 122, pp. 75–84, 2019.

[37] A. Holzinger, R. Goebel, and M. Ferri, "Towards integrative machine learning and knowledge extraction: BIRS workshop, banff, AB, Canada, july 24-26, 2015, revised selected papers," *Lecture Notes in Computer Science*, Berlin, Germany, 2017.

[38] E. Mueggler, H. Rebecq, G. Gallego, T. Delbruck, and D. Scaramuzza, "The event-camera dataset and simulator: event-based data for pose estimation, visual odometry, and SLAM," *The International Journal of Robotics Research*, vol. 36, no. 2, pp. 142–149, 2017.

[39] M. Shakeri, M. H. Dezfoulian, H. Khotanlou, A. Barati, and Y. Masoumi, "Image contrast enhancement using fuzzy clustering with adaptive cluster parameter and sub-histogram equalization," *Digital Signal Processing*, vol. 62, pp. 224–237, 2017.

[40] Y. Shi, Y. Wei, D. Pan et al., "Student body gesture recognition based on Fisher broad learning system," *International Journal of Wavelets, Multiresolution and Information Processing*, vol. 17, no. 01, Article ID 1950001, 2019.

[41] W. Huang, N. Li, Z. ., J. Qiu, N. Jiang, B. Wu, and B. Liu, "An automatic recognition method for students' classroom behaviors based on image processing," *Traitement du Signal*, vol. 37, no. 3, pp. 503–509, 2020.

[42] Y. Y. Cheng, Z. J. Dai, Y. Ji, and L Simin, "Student action recognition based on deep convolutional generative adversarial network," in *Proceedings of the 2020 Chinese control and decision conference (CCDC)*, vol. 35, no. 6, pp. 128–133, IEEE, Hefei, China, August 2020.

[43] S. Memiş, S. Enginoğlu, and U. Erkan, "Numerical data classification via distance-based similarity measures of fuzzy parameterized fuzzy soft matrices," *IEEE Access*, vol. 9, pp. 88583–88601, 2021.

[44] R. Liu, X. Wang, H. Lu et al., "SCCGAN: style and characters inpainting based on CGAN," *Mobile Networks and Applications*, vol. 26, no. 1, pp. 3–12, 2021.

[45] W. Zhou, J. Liu, J. Lei, L. Yu, and J. N. Hwang, "GMNet: graded-feature multilabel-learning network for RGB-thermal urban scene semantic segmentation," *IEEE Transactions on Image Processing*, vol. 30, pp. 7790–7802, 2021.

[46] G. Sun, Y. Cong, Q. Wang, B. Zhong, and Y. Fu, "Representative task self-selection for flexible clustered lifelong learning," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 33, no. 4, 1481 pages, 2020.

[47] F. Liu, G. Zhang, and J. Lu, "Multisource heterogeneous unsupervised domain adaptation via fuzzy relation neural

networks," *IEEE Transactions on Fuzzy Systems*, vol. 29, no. 11, pp. 3308–3322, 2021.

[48] W. Yang, X. Chen, Z. Xiong, Z. Xu, G. Liu, and X. Zhang, "A privacy-preserving aggregation scheme based on negative survey for vehicle fuel consumption data," *Information sciences*, vol. 570, pp. 526–544, 2021.

[49] D. Li, S. S. Ge, and T. H. Lee, "Simultaneous arrival to origin convergence: sliding-mode control through the norm-normalized sign function," *IEEE Transactions on Automatic Control*, vol. 67, no. 4, pp. 1966–1972, 2022.

[50] B. Zhu, Q. Zhong, Y. Chen et al., "A novel reconstruction method for temperature distribution measurement based on ultrasonic tomography," *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 69, no. 7, pp. 2352–2370, 2022.

[51] W. Zheng, X. Tian, B. Yang et al., "A few shot classification methods based on multiscale relational networks," *Applied Sciences*, vol. 12, no. 8, p. 4059, 2022.

[52] X. Wu, W. Zheng, X. Chen, Y. Zhao, T. Yu, and D. Mu, "Improving high-impact bug report prediction with combination of interactive machine learning and active learning," *Information and Software Technology*, vol. 133, Article ID 106530, 2021.

[53] Y. Wang, H. Wang, B. Zhou, and H. Fu, "Multi-dimensional prediction method based on Bi-LSTMC for ship roll," *Ocean Engineering*, vol. 242, Article ID 110106, 2021.

[54] Y. Ban, M. Liu, P. Wu et al., "Depth estimation method for monocular camera defocus images in microscopic scenes," *Electronics*, vol. 11, no. 13, p. 2012, 2022.

[55] W. Zheng, X. Liu, and L. Yin, "Research on image classification method based on improved multi-scale relational network," *PeerJ Computer Science*, vol. 7, p. e613, 2021.

[56] J. Wang, J. Tian, X. Zhang et al., "Control of time delay force feedback teleoperation system with finite time convergence," *Frontiers in Neurorobotics*, vol. 16, Article ID 877069, 2022.

[57] J. Li, K. Xu, S. Chaudhuri, E. Yumer, H. Zhang, and L. Guibas, "Grass: generative recursive autoencoders for shape structures," *ACM Transactions on Graphics*, vol. 36, no. 4, pp. 1–14, 2017.

[58] M. Qi, S. Cui, X. Chang et al., "Multi-region Nonuniform Brightness Correction Algorithm Based on L-Channel Gamma Transform," *Security and Communication Networks*, vol. 2022, Article ID 2675950, 2022.

[59] W. Zheng and L. Yin, "Characterization inference based on joint-optimization of multi-layer semantics and deep fusion matching network," *PeerJ Computer Science*, vol. 8, p. e908, 2022.

[60] W. Wang, Z. Chen, and X. Yuan, "Simple low-light image enhancement based on Weber-Fechner law in logarithmic space," *Signal Processing: Image Communication*, vol. 106, p. 116742, 2022.

[61] W. Zhou, L. Yu, Y. Zhou, W. Qiu, M. W. Wu, and T. Luo, "Local and global feature learning for blind quality evaluation of screen content and natural scene images," *IEEE Transactions on Image Processing*, vol. 27, no. 5, pp. 2086–2095, 2018.

Hindawi
Computational Intelligence and Neuroscience
Volume 2022, Article ID 9591781, 8 pages
https://doi.org/10.1155/2022/9591781

Research Article

Detection of Breast Cancer Lump and BRCA1/2 Genetic Mutation under Deep Learning

Yue Miao and Siyuan Tang

Department of Computer Science and Technology, Baotou Medical College, Inner Mongolia University of Science and Technology, Baotou 014160, China

Correspondence should be addressed to Yue Miao; 102007036@btmc.edu.cn

Received 10 July 2022; Revised 16 August 2022; Accepted 23 August 2022; Published 19 September 2022

Academic Editor: Ning Cao

To diagnose and cure breast cancer early, thus reducing the mortality of patients with breast cancer, a method was provided to judge threshold of image segmentation by wavelet transform (WT). It was used to obtain information about the general area of breast lumps by making a rough segmentation of the suspected area of the lump on mammogram. The boundary signal of the lump was obtained by region growth calculation or contour model of local activity. Meanwhile, multiplex polymerase chain reaction (mPCR) and mPCR-next-generation sequencing (mPCR-NGS) were used to detect BRCA1/2 genome. Sanger test was used for newly high virulent mutations to verify the correctness of mutagenic sites. The results were compared with the information marked by experts in the database. According to Daubechies wavelet coefficients, the average measurement accuracy was 92.9% and the average false positive rate of each image was 86%. According to mPCR-NGS, there was no pathogenic mutation in the 7 patients with high-risk BRCA1/2 genetic mutations. Single nucleotide polymorphism (SNP) in nonsynonymous coding region was detected, which was consistent with the Sanger test results. This method effectively isolated the lump area of human mammogram, and mPCR-NGS had high specificity and sensitivity in detecting BRCA1/2 genetic mutation sites. Compared with traditional Sanger test and target sequence capture test, it also had such advantages as easy operation, short duration, and low cost of consumables, which was worthy of further promotion and adoption.

1. Introduction

The incidence of breast cancer has been on the rise since the second half of the 20th century, and it has become a major health problem for women [1]. That the breast lesions are detected, discovered, and diagnosed as early as possible is the most effective way to save the patient's life [2]. Mammography is one of the most important methods for early detection of breast cancer worldwide [3]. With the advance of time and the rapid development of modern science and technology, the computer-assisted testing technology of breast cancer has become a hot subject that attracts more scholars' attention [4]. Computer-assisted examination technology can not only save much human labor, but also help to understand and analyze mammogram images [5]. The causes of breast cancer are complex. Generally speaking, it is mainly because the cells and tissues in the breast change their normal characteristics under the combined influence of internal and external carcinogenic factors, thus forming abnormal cell proliferation phenomenon. When these lesions reach their maximum recovery rate, breast cancer occurs [6–8]. Breast cancer is the most deadly gynecological malignant tumor at present. According to the document research, BRCA1/2 genetic mutation exists in more than 80% of patients with high genetic breast cancer, and the incidences of lifetime breast cancer and breast cancer with this genetic mutation are much higher than the general population [9, 10]. Hence, it is necessary to monitor BRCA1/2 genetic mutation and give the corresponding intervention in high-risk groups.

At present, the screening and early treatment of breast cancer mostly rely on modern techniques such as

mammography, magnetic resonance imaging (MRI), and ultrasonic examination [11]. With the advantages of simple operation, high cost performance, and noninvasive surgery, mammography has become the preferred photographic method in breast cancer screening [12]. With the rapid development of medical technology and modern scientific and technological means, the early treatment of breast cancer also plays an increasingly great role in women's health care [13]. In the design of computer-aided diagnosis (CAD) system, how to reasonably classify the range of lesions is a key skill [14]. Since the middle of the twentieth century, the concept of image segmentation has emerged in the field of computer image processing. In the following decades, image segmentation has become a hot topic in the field of digital image processing [15, 16]. Image subdivision method is classifying the image into a considerable number of regions with unique properties and extracting the interest region of the technical method and process [15]. In the field of biomedical application, image segmentation technology plays a crucial role in medical image. It can help to measure and locate the tissue structure of cancer and other diseases and make it easier to measure the tissue volume [17].

To sum up, at present, how to diagnose breast cancer in initial stage is a major problem that needs to be paid attention to in clinic. An innovative separation method based on wavelet transform (WT) was proposed. The algorithm effectively classified the suspicious area of the breast image. Then, according to the characteristic parameters of the tumor area, it also found the tumor tissue in the breast accurately, thereby realizing the automatic detection of the breast tumor. Meanwhile, multiplex polymerase chain reaction-next-generation sequencing (mPCR-NGS) was performed on eligible high-risk group with BRCA1/2 genome mutation to test whether there was BRCA1/2 genome mutation, and satisfactory conclusions were obtained.

2. Materials and Methods

2.1. Objects of Study. Seven patients who were the suspected high-risk BRCA1/2 genetic mutation carriers of XX Hospital from June to December 2020 were included. Among them, there were 4 patients who were diagnosed with breast cancer and 3 patients who were asymptomatic high-risk carriers. All the patients signed the informed consent, and the research was approved by the ethics committee.

The inclusion criteria were as follows: I. patients whose immediate family members had breast cancer lesions; II. patients whose family member had breast cancer; and III. patients whose onset age of early-onset breast cancer was less than or equal to 36 years old. If one of the above three conditions was satisfied, the research was implemented. The exclusion criteria were as follows: I. patients who refused to sign or did not sign the informed consent; and II. patients who had the medical history of trauma, cardiovascular disease, and metabolic disease.

2.2. The Serial Number of Double-Blind Method. The peripheral blood samples of the research objects and the control group were coded as $P1$, $P2$, $P3$, $P4$, $P5$, $P6$, $P7$, and $P8$. After the sequencing, the third party finished the unblinding task.

2.3. Extraction of Deoxyribonucleic Acid (DNA) of Peripheral Blood. At the temperature of 200°C to 250°C, the samples were extracted out and placed on the centrifugal tube plate rack, and they were thawed at room temperature. 20 μL proteinase K and 200 μL anticoagulant peripheral venous blood solution were added to 1.6-mL centrifuge tubes, and they were gently flicked for several times. After the complete uniformity, instantaneous centrifugation was performed. After the covers of the tubes were removed, 4 μL ribonuclease A (RNase A) was added to each tube and the total vorticity was 15 s. After instantaneous centrifugation, the tubes were left standing for about 2 min at room temperature. 200 μL Buffer AL was added and the vorticity was 15 s. After incubation at 56°C for 10 min, the samples were removed from the metal bath and cooled to normal temperature for 2 min. 200 μL anhydrous ethyl alcohol was added and mixed with vorticity oscillation for 15 s. Then, the centrifugation was performed instantaneously. After centrifugation, the solvent was transferred to the purification column, avoiding the contact with the nozzle of the purification column as far as possible. After the centrifugation at 6,000 rpm for 1 min, the purification column was transferred into a new collection catheter, and another 500 μL Buffer AW1 was added to the purification column for elution. After the centrifugation at 6,000 rpm for 1 min, the waste liquid was discarded, and the purification column was transferred into a new collection catheter. 500 μL Buffer AW2 was added to the purification column for elution. After the centrifugation at 20,000 rpm for 3 min, the waste liquid was discarded. Then, the residual liquid at the end of the tube was wiped with absorbent paper, and it was centrifuged at 20,000 rpm for 1 min. A new 5-mL centrifuge tube was taken, into which the purification column was transferred. 200 μL Buffer AE was added. Subsequently, after the cover of the centrifuge tube was closed, it was placed at the indoor humidity for 2 min, and the centrifugation was performed at 6,000 rpm for 1 min. Finally, the liquid in the centrifuge tube was collected, which was the obtained gDNA.

2.4. The Technology of mPCR to Amplify BRCA1/2 Gene. 5 μL PCR amplification product was taken and mixed with 1 μL 6 × Loading buffer. The sample was added to 0.6% agarose gel pore by pipetting gun. The amount of foam was minimized and the liquid was made to overflow the pore. Additionally, 3 μL 1 K DNA Marker was added into the pore for the molecular number marker. According to the direction of the positive and negative electrodes, the groove cover was closed, and the power switch was turned on. The voltage was adjusted to 120 V for constant pressure electrophoresis, and the duration of agarose gel electrophoresis was set to about 30 min. After the electrophoresis was completed, the power switch was turned off immediately. The gel was carefully extracted from the electrophoresis tank, and it was put into the two-dimensional gel electrophoresis image analyzer for image scanning observation and

inspection. Bright and clear amplification bands were obtained by 1 K DNA Marker, which were labeled with the size of 1 Kb–10 Kb. The electrophoresis bands of PCR products were observed to check the location and brightness of target bands and compared with Marker. If the bands of the amplification products were concentrated in the region of 2 K–7 K, the amplification was successful.

2.5. Construction of the Resequencing Library for the Second-Generation Sequencing. 5 μL amplified material and 1 μL 6 × Loading Buffer were evenly stirred and added to the wells of 2% agarose gel, with 3 μL 100 bp DNA Marker as the molecular weight marker. Constant pressure electrophoresis was performed at 120 V for nearly 25 minutes. The size of the electrophoretic band of each PCR product was carefully observed. If the interruption band was found in the size of 202 bp–423 bp, the interruption was successful.

2.6. Designing Primers for Mutation Sites. According to the mutation sites obtained by mPCR-NGS sequencing, the first-generation sequencing experiments were used for nonsynonymous single nucleotide polymorphism (SNP) (serial numbers were 1, 2, and 3) with low mutation frequency. Primers were designed by the primer premier 5.0, and the product width of primer amplification generally ranged from 400 bp to 750 bp.

2.7. Segmentation of Lump Based on WT. Any signal is described as a sum of string functions according to Fourier theory. Hence, when a signal is completely described by the Fourier function, it has a spectral resolution rather than a temporal resolution, which means that the frequencies involved in the message can be identified, but the time at which it occurs cannot be identified. In the 1980s, some French physicists invented the wavelet analysis theory to inherit the advantages of the traditional Fourier analysis method and overcome the deficiencies [18, 19]. Through the development of recent decades, the foundation of modern wavelet analysis theory has been basically laid. Moreover, wavelet analysis has not only become a new field of applied mathematics, but also started to be used in many engineering fields.

The WT coefficient refers to a specific waveform, which is unique in that the wavelets have a finite width and an average value of 0. It borrows from Fourier analysis to decompose an electronic signal into a series of overlapping sinusoids of different frequencies. Wavelet analysis resolves the signal into a series of overlapping wavelet functions, which are derived from a set of identical mother wavelet functions through corresponding transformations. Wavelet analysis helps to obtain the features of the date signal when the signal occurs, which is usually achieved by using the translation of mother wavelet. Similarly, the frequency features of the time signal are obtained by scaling the WT.

WT is classified into the continuous WT and the discrete WT. For continuous WT, all *y(s)* in $L^2(R)$ space are extended under the wavelet basis, which is called *y(s)* continuous WT. In equation (1), m represents zoom factor and γ represents time translation.

$$U_{(m,\gamma)} = < y(s), \theta_{m,\gamma}(s) \geq \frac{1}{\sqrt{m}} \int Ry(s)\overline{\theta\left(s - \frac{\gamma}{m}\right)} d_s. \quad (1)$$

For the discrete WT, in any $L^2(R)$ space, the discrete WT of *y(s)* is expressed by the (2) and (3). It is important to note that this discrete transformation is about continuous large-scale parameters and translational parameters, not about s.

$$U_Z(e\ ,\ i) = \int_Y y(s)\overline{\complement_{hl}(s)} ds, \quad (2)$$

$$\complement_{hl}(s) = \frac{1}{\sqrt{2^h}} \complement\left(\frac{s}{2^h} - i\right). \quad (3)$$

WT is based on Fourier transform, which is further developed. It inherits the advantages of Fourier transform and overcomes the shortcomings. Generally speaking, WT has both frequency parsing and time parsing. Compared with Fourier transform, one of the biggest advantages is judging the timing of specific phenomena in the signal. Furthermore, it is possible to extract different resolution features from information because of the multiscale features of WT. Finally, in terms of the operating rate of the WT, it is much faster than Fourier transform.

2.8. Automatic Segmentation of Lump Based on Gray Histogram WT. The gray histogram of mammogram was classified by WT, and the segmentation threshold was determined regarding the limit value of WT modulus. First, before the lump area was extracted, the influence of non-cancerous parts of the breast on the determination of the lump area was eliminated, including the capillaries in the breast and the scars on the skin. The wavelet filtering method was selected as the main method to achieve image smoothness. Consequently, the effect of texture signals such as capillaries and skin in breast images on the tumor detection results was reduced. After the smoothing filtering, the histogram of the breast image needed to be recalculated. For most medical images, before the histogram transformation was completed, the image was normalized and the gray scale range was controlled into the gray space of 0–225, because the gray level area was large and the gray histogram distribution was inconsistent. In (4), P represented the prenormative image, and K represented the image processed by the specification.

$$K(z,u) = (P(z,u) - \min(P) \times \frac{255 - 0}{\max(P) - \min(P)}. \quad (4)$$

In the breast image, there was usually a jumping in gray value where the lump area was connected to the nearby normal tissue area. Hence, only by finding the gray value of the distortion correction point was the whole breast separated. However, in the actual breast images, since the volume of the lump department was relatively small compared with the whole breast department, there were some small bumps in the high point of the gray order in the gray histogram,

which were not observed directly. The wavelet for singularity test was obtained from smooth function based on the principle of wavelet transform modulus maximum test. If $\delta(z)$ was the smooth function with low pass features, the first stage of derivative was expressed as

$$\omega(z) = \frac{d\delta(z)}{ds}. \tag{5}$$

It was used as a wavelet, and the (6) was obtained.

$$Q_d^1(a,z) = f(z) * \theta_a^1(z) = s\frac{d}{dz}\left[f(z) * \varphi_a(z)\right]. \tag{6}$$

(7) represents the scaling transformation of $\delta(z)$ under scale S.

$$\varphi_a(z) = \frac{1}{a}\varphi\left(\frac{z}{a}\right). \tag{7}$$

In (6), the WT of the signal $f(z)$ was expressed as a derivative of $f(z)$ after smooth processing at the scale S. At this point, the extreme point of the WT was the turning point of $f(z) * \delta(z)$, which was the step point in the extreme case. $Q_d^1(a,z)$ was used to represent the sign signal time. The maximum point corresponded to the place where the information changed dramatically, while the minimum point corresponded to the place where the information changed slowly. This special property of WT was used to locate the singular points in the gray histogram of breast image to find the corresponding singular points with small gray order value. The minimum point that was nearest to the singular point was found, and the corresponding gray scale threshold point of the singular point was the threshold point of lump suspected region division after scale regression. The actual size and morphology of the lump were considered, and the addition of the area information of the block and the ratio of length to width of the minimum enclosing rectangle became another restriction requirement for image segmentation. According to the above requirements, the general situation of the suspected range of the lump was obtained, and the range of the lump in the mammography was generally detected.

The gray histogram of mammogram was examined by using wavelet mode maximum method. After the segmentation threshold was found and the lump location was located, the lump location was observed to have different properties. Some lumps stuck to the surrounding tissues, so it was difficult to precisely locate the boundaries of these lumps. There were some obvious differences between the other parts of the tumor and the surrounding tissue boundaries. For those lumps with obvious differences in gray features between the lump area and the surrounding tissue, simple region growth method was used to obtain the boundaries of these lumps. For those lumps with fuzzy edges, active contour model was used to extract their boundaries. (8) was the contour model equation of local activity.

$$G(B, j_1, j_2) = \omega R(B) + wT(B) + \sigma_1 \int_{\text{inside}(B)} |H - j_1|^2 epeq + \sigma_2 \int_{\text{inside}(B)} |H - j_2|^2 epeq. \tag{8}$$

The parameters used in the contour model of local activity changed with the change of each point, and it was not a long-term invariant gray scale mean value of the target region or background region. Therefore, when there was a certain regional difference between the homogeneity of the target area and the background area, the image was well separated through this mode. Hence, the contour mode of local activity helped to solve the deficiency of image separation in CV mode where the homogeneity of two edges was different. With the introduction of local information, the boundary of the target region was accurately found when the problem of image separation with uneven gray scale distribution was solved.

3. Results

3.1. Sequencing Quality and Coverage Quality. Table 1 shows the quality detection of the sequencing samples.

Figure 1 shows the tests of the coverage depth of the samples.

3.2. Analysis of Sequencing Results. Table 2 shows the results of sample sequencing.

According to the detection results, no pathogenicity mutation sites with a low occurrence rate (<1%) in Asian population were detected.

3.3. The Verification Results of the First-Generation Sequencing. The recent generation of gene detection was confirmed by the nonsynonymous SNP with low mutation frequency in three Asians and high occurrence frequency in the samples, which was consistent with the results of sequencing (Table 3).

Figures 2–4 showed the sequencing results of specific mutation sites.

3.4. Results of Automatic Segmentation of Lumps by WT. To obtain the best experimental efficiency, four WTs with different sizes were selected for comparative experiments. After the test results were analyzed, the wavelet parameters to obtain the best segmentation effect were determined. Figure 5 shows the comparative test results.

3.5. Results of Lump Edge Detection Based on Active Contour Model. Among the 70 breast images that included 68 lumps, 66 lump areas were obtained by rough examination, and 59

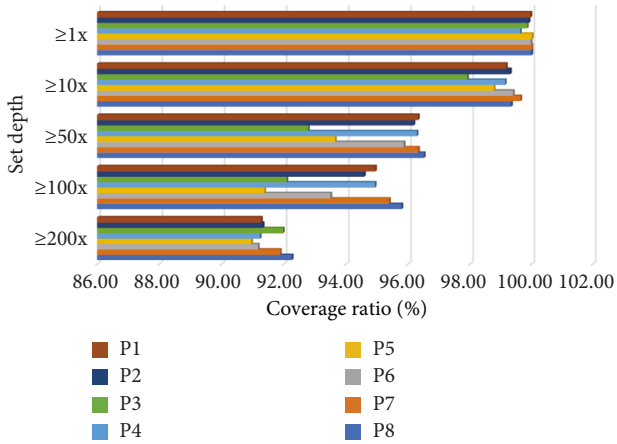


Figure 1: The tests of the coverage depth of experimental group and control group.

Table 1: Quality detection of samples.

Sample	*P1*	*P2*	*P3*	*P4*	*P5*	*P6*	*P7*	*P8*
Map bases	298237468	296811389	396791957	297232536	197714073	290648699	294785522	290241136
Clean bases	636832098	598907146	794252182	697338797	494135783	694778908	699269869	594654691
Clean reads	73558937	61226224	67547282	76173306	55717899	60804209	69899226	65224092

Table 2: Results of sample sequencing.

Sample name	*P1*		*P2*		*P3*	*P4*
Gene	BRCA2	BRCA2	BRCA1	BRCA1	BRCA1	BRCA2
Allele Freq Asn	23	120	36	36	36	100
Coordinate	39382638	30264927	419200372	43375502	41172649	32929387
HGVSc	c.H75A>C	c.6634T>C	c.4373A>G	c.2643A>G	c.237C>T	c.7397T>C
COSMIC ID	COSM103751		COSM134096		COSM166528	
PolyPhen	Benign (0.536)	Benign (0.231)	Benign (0.1)	Possibly damaging (0.732)	Benign (0.231)	Benign (0)
Geno type	horn	het	hom	het	hom	hom
Sample name	*P5*	*P6*	*P7*	*P8*	Control group	
Gene	BRCA2	BRCA1	BRCA2	BRCA1	BRCA1	BRCA2
Allele Freq Asn	100	33	100	26	100	100
Coordinate	33683462	43477289	32352679	32647899	30038853	31125674
HGVSc	c.72378>C	c.7532AXJ	c.2367T>C	c.7792A>C	c.2352T>C	c.1567T>C
COSMIC ID		COSM13783		COSM 14202		
PolyPhen	Benign (0.001)	Benign (0)	Benign (0.343)	Benign (0)	Benign (0.1)	Benign (0.116)
Geno type	hom	het	hom	het	hom	hom

Table 3: Site information of the first-generation sequencing.

Sample	*P2*	*P5*	*P6*
Gene	BRCA1	BRCA2	BRCA1
HGVSp	p.Serl 521Gly	p.Val2688AIa	p.Glul236Gly
HGVSc	c.1524A>G	c.8745T>C	c.1624A>G
Coordinate	42724678	39095326	42456093
	rs31566633	rs44291	rs773092

of them were completely identified and their boundaries were extracted.

4. Discussion

The mPCR-NGS was performed on seven clinically collected high-risk BRCA1/2 mutation carriers, and no pathogenicity mutation sites were found. Nevertheless, through the first-generation genetic sequencing, the experimental results were consistent with the results of sequencing proposed in the study. At the stage of rough segmentation of breast lump, the gray histogram of each image was firstly obtained, and the distribution of gray histogram of each image was also different with the difference of specific conditions in the

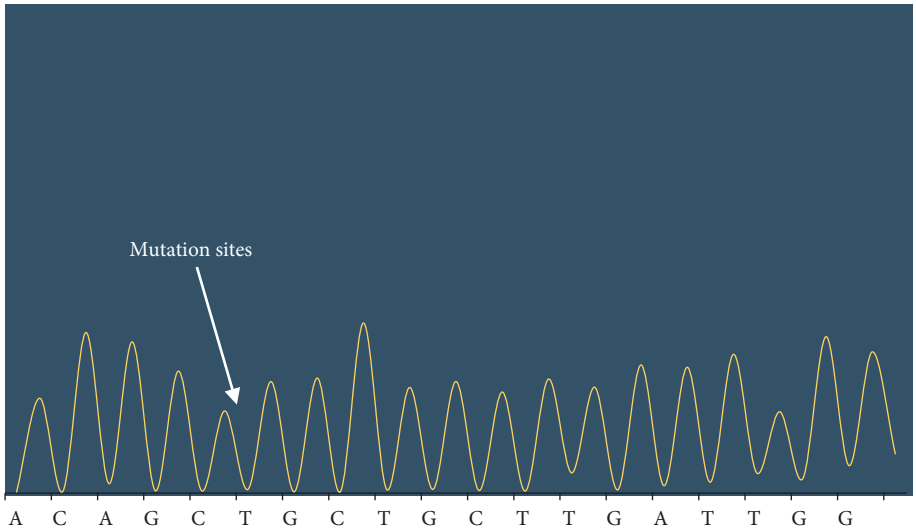


Figure 2: The first-generation sequencing was used to test sample $P2$ at site NO. 1, and the location of mutation sites was c.7397T>G. BRCA2 gene belonged to the positive chain, so the mutation of this site was A>G according to the principle of reverse complementation. The homozygosis was that one hundred percent of the mutation occurred.

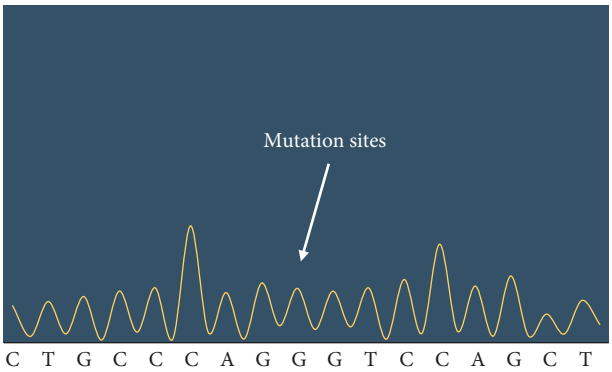


Figure 3: The first-generation sequencing was used to test sample $P5$ at site No. 2, and the location of mutation sites was c.4900A>G. BRCA1 gene was in negative chain, so the mutated gene was at this position A>G. The homozygosis was that one hundred percent of the mutation occurred.

imaging. Besides, it was necessary to implement the gray histogram standardized management.

The WT of Daubechies with the orders including $N = 5$, $N = 10$, $N = 15$, and $N = 20$ was selected to perform one-dimensional WT. The results showed that the WT of the order $N = 20$ obtained the best measurement effect, the detection rate was more than 92.90%, and the average number of false positive expression in each image was 0.86. According to the theoretical analysis, since the white noise was singular everywhere in the whole signal, it was pointed out that the amplitude and density of the maximum point of WT mode formed by white noise were inversely proportional to the large-scale parameters of wavelet. Nonetheless, the actual image information was different. The singularity of image information was mainly caused by the incoherence of image gray, and the maximum value of WT was little affected by the size change. With the influence of small-scale WT, it was difficult to distinguish the noise from the detail information of the signal because the dominant function of the maximum point formed in the detail signal of the image was similar. Moreover, it was why the detection accuracy was low and the false-positive rate was high when the small size was used. On the contrary, with the influence of the minimum wavelet function of large size, the mutation signal of the gray level of image had great effect on the maximum dominance of the whole signal. Therefore, the selection of large size wavelet parameters produced good singular value detection features, thus helping to achieve good efficacy in rough segmentation of lumps.

Additionally, although the selected WT with $N = 20$ still retained a relatively high false-positive rate, the possible reduction in the number of false-positive expressed lumps in subsequent tests was considered. The wavelet of $N = 20$ was chosen as the coefficient of WT and analyzed theoretically. The WT of order $N = 20$ had good processing efficiency for different wavebands, and it was helpful to cut the evenly distributed gray order histogram exactly. After the appropriate wavelet was selected, the odd features in the gray histogram were measured by using the basic principle of the modulus maximum of WT. Furthermore, this singularity corresponded to the gray distortion in the straight mammary gland image. After the appropriate singular loci was determined and the appropriate segmentation threshold was selected, the rough segmentation was performed automatically according to the segmentation threshold. In this experiment, 70 images from the MIAS database were examined. The results showed that the proposed method satisfactorily determined the size of hard lumps in mammogram, and the detection rate was 92.9%. As for undetectable lumps, there were generally two cases. The first case was that during the imaging process, the lump area and the pectoral muscle area were double merged and eliminated during the pretreatment process. The second case was that it was difficult to detect the lump in breast images because the gray data were very similar to those of the surrounding

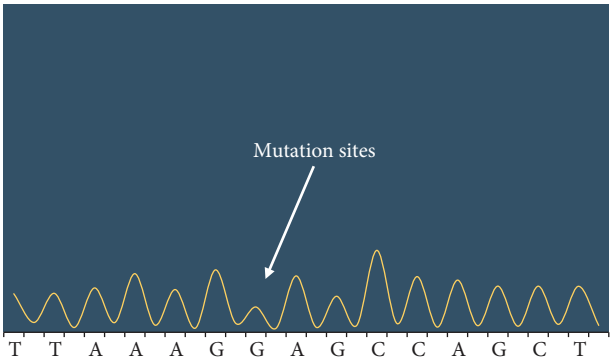


Figure 4: The first-generation sequencing was used to test sample *P6* at site NO. 3, and the location of mutation sites was c.3113A>G. BRCA1 gene was in negative chain, so this site was A>G after mutation. It was the heterozygosis that the target gene mutated into bimodal type.

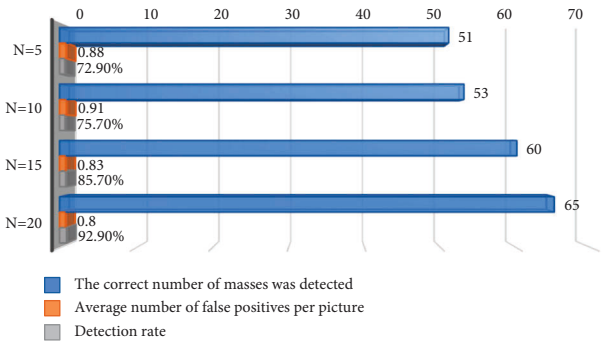


Figure 5: Comparative experiments of four WTs with different sizes.

normal tissues. However, as for the lumps detected by rough segmentation, the proposed method was helpful to well extract the boundary of lumps.

To sum up, this method basically achieved the automatic segmentation of digital mammography lumps by computer, and the segmentation effect was relatively ideal, which helped to promote the value of further study and adoption. In the modern society with the continuous development of high technology, people pay much attention to the diseases that they suffer from, and they have a comprehensive understanding of how to prevent and treat diseases [20]. In the process of collecting samples, both patients with breast cancer and specialist doctors had a strong interest in genomic screening and an emerging technical field, in which they voluntarily participated. It fully demonstrated that there was further development space in the field of genetics [21].

5. Conclusion

The techniques of breast tumor segmentation in mammography were systematically summarized and analyzed. After the study of various references, a new method for mammogram lump segmentation was provided. The principle of wavelet maximum was used to measure the singular points of high gray histogram blocks in breast image, and the rough segmentation threshold of hard blocks was determined by selecting appropriate singular points. Moreover, the specificity and sensitivity of BRCA1/2 genetic mutation sites were all detected by mPCR-NGS sequencing platform for clinically high-risk BRCA1/2 mutation carriers. Compared with traditional Sanger sequence testing and target sequence testing methods, mPCR-NGS sequencing platform had such advantages as easy execution, short time consuming, and low cost.

Any early diagnosis and prevention of breast cancer has always been the focus of biomedical research. Nevertheless, due to the limitations of research equipment and technology, the tumor isolation and genetic mutation detection methods on mammography images were investigated preliminarily. There are still many advanced algorithms that are not used, and there is still room for improvement and optimization in the experimental results. Additionally, there is still much work to be done in the future.

Data Availability

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethical Approval

This article does not contain any studies with human participants or animals performed by any of the authors.

Consent

Informed consent was obtained from all individual participants included in the study.

Conflicts of Interest

All authors declare that they have no conflicts of interest.

Acknowledgments

The authors acknowledge the help from the university colleagues. This work was supported by Scientific research project of higher science schools of the Department of education of Inner Mongolia (Grant no. NJZY21068) and Research on automatic detection and grading of diabetes retinopathy based on deep learning, Natural Science Foundation of Inner Mongolia (Grant no. 2021MS06010).

References

[1] A. Le Naour, A. Rossary, and M. P. Vasson, "EO771, is it a well-characterized cell line for mouse mammary cancer model? Limit and uncertainty," *Cancer Medicine*, vol. 9, no. 21, pp. 8074–8085, 2020 Nov.

[2] X. Li, H. Gao, M. Xu, Y. Wu, and D. Gao, "Breast papillary lesions diagnosed and treated using ultrasound-guided vacuum-assisted excision," *BMC Surgery*, vol. 20, no. 1, p. 204, 2020 Sep 15.

[3] T. A. Licari, "Mammography safety revisited," *Radiologic Technology*, vol. 91, no. 2, pp. 191–195, 2019 Nov, PMID: 31685598.

[4] Z.-M. Á, C. Castillo-Amature, J. M. Montiel-Company, and J. Mena-Álvarez, "Efficacy of computer-aided static navigation technique on the accuracy of endodontic microsurgery. A systematic review and meta-analysis," *Journal of Clinical Medicine*, vol. 10, no. 2, p. 313, 2021 Jan 15.

[5] K. Moustafa, "Publishers: save authors' time," *Science and Engineering Ethics*, vol. 24, no. 2, pp. 815-816, 2018.

[6] K. Ye, Q. Wei, Z. Gong et al., "Effect of norcantharidin on the proliferation, apoptosis, and cell cycle of human mesangial cells," *Renal Failure*, vol. 39, no. 1, pp. 458–464, 2017 Nov.

[7] M. Akram, M. Iqbal, M. Daniyal, and A. U. Khan, "Awareness and current knowledge of breast cancer," *Biological Research*, vol. 50, no. 1, p. 33, 2017 Oct 2.

[8] A. K. Gupta, S. Tulsyan, N. Thakur, V. Sharma, D. N. Sinha, and R. Mehrotra, "Chemistry, metabolism and pharmacology of carcinogenic alkaloids present in areca nut and factors affecting their concentration," *Regulatory Toxicology and Pharmacology*, vol. 110, Article ID 104548, 2020 Feb.

[9] Y. Zhang and Z. Liu, "Oncolytic virotherapy for malignant tumor: current clinical status," *Current Pharmaceutical Design*, vol. 25, no. 40, pp. 4251–4263, 2020.

[10] A. Anna and G. Monika, "Splicing mutations in human genetic disorders: examples, detection, and confirmation," *Journal of Applied Genetics*, vol. 59, no. 3, pp. 253–268, 2018 Aug.

[11] V. Zanardo and M. Parotto, "Ultrasonic Pocket Doppler, novel technology for fetal and neonatal heart rate assessment," *Resuscitation*, vol. 145, pp. 91-92, 2019 Dec.

[12] Y. Hu, Y. Zhang, and J. Cheng, "Diagnostic value of molybdenum target combined with DCE-MRI in different types of breast cancer," *Oncology Letters*, vol. 18, no. 4, pp. 4056–4063, 2019 Oct.

[13] A. O'Neil, J. D. Russell, K. Thompson, M. L. Martinson, and S. A. E. Peters, "The impact of socioeconomic position (SEP) on women's health over the lifetime," *Maturitas*, vol. 140, pp. 1–7, 2020 Oct.

[14] A. K. C. Leung, J. M. Lam, and K. F. Leong, "Childhood Langerhans cell histiocytosis: a disease with many faces," *World J Pediatr*, vol. 15, no. 6, pp. 536–545, 2019 Dec.

[15] A. Bitarafan, M. Nikdan, and M. S. Baghshah, "3D image segmentation with sparse annotation by self-training and internal registration," *IEEE J Biomed Health Inform*, vol. 25, no. 7, pp. 2665–2672, 2021 Jul.

[16] J. Joseph, D. Kalpana, S. Rao, and S. Raju Kurapati, "Digital dental photography," *Indian Journal of Dental Research*, vol. 29, no. 4, pp. 507–512, 2018 Jul-Aug.

[17] R. Struck, S. Cordoni, S. Aliotta, L. Pérez-Pachón, and F. Gröning, "Application of photogrammetry in biomedical science," *Advances in Experimental Medicine and Biology*, vol. 1120, pp. 121–130, 2019.

[18] A. Jaffe, C. Jiang, Z. Liu, Y. Ren, and J. Wu, "Quantum fourier analysis," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 117, no. 20, pp. 10715–10720, 2020 May 19.

[19] A. Kumar, H. Tomar, V. K. Mehla, R. Komaragiri, and M. Kumar, "Stationary wavelet transform based ECG signal denoising method," *ISA Transactions*, vol. 114, pp. 251–262, 2021 Aug.

[20] E. Grazioli, I. Dimauro, N. Mercatelli et al., "Physical activity in the prevention of human diseases: role of epigenetic modifications," *BMC Genomics*, vol. 18, no. Suppl 8, p. 802, 2017 Nov 14.

[21] U. Menon, C. Karpinskyj, and A. Gentry-Maharaj, "Ovarian cancer prevention and screening," *Obstetrics & Gynecology*, vol. 131, no. 5, pp. 909–927, 2018 May.

Hindawi
Computational Intelligence and Neuroscience
Volume 2022, Article ID 5993839, 11 pages
https://doi.org/10.1155/2022/5993839

Research Article

Construction and Analysis of Emotion Recognition and Psychotherapy System of College Students under Convolutional Neural Network and Interactive Technology

Minwei Chen,[1] **Xiaojun Liang**,[2] **and Yi Xu**[3]

[1]*College of Physical Education, Chongqing University, Chongqing 400044, China*
[2]*College of Humanities, Zhaoqing Medical College, Zhaoqing 526020, China*
[3]*Ministry of Basic Education, GuangdongEco-Engineering Polytechnic, Guangzhou 510520, China*

Correspondence should be addressed to Minwei Chen; cmw@cqu.edu.cn

Received 20 July 2022; Revised 24 August 2022; Accepted 6 September 2022; Published 17 September 2022

Academic Editor: Ning Cao

This study's aim is to effectively establish a psychological intervention and treatment system for college students and discover and correct their psychological problems encountered in a timely manner. From the perspectives of pedagogy and psychology, the college students majoring in physical education are selected as the research objects, and an interactive college student emotion recognition and psychological intervention system is established based on convolutional neural network (CNN). The system takes face recognition as the data source, adopts feature recognition algorithms to effectively classify the different students, and designs a psychological intervention platform based on interactive technology, and it is compared with existing systems and models to further verify its effectiveness. The results show that the deep learning CNN has better ability to recognize student emotions than backpropagation neural network (BPNN) and decision tree (DT) algorithm. The recognition accuracy (ACC) can be as high as 89.32%. Support vector machine (SVM) algorithm is adopted to classify the emotions, and the recognition ACC is increased by 20%. When the system's K value is 5 and d value is 8, the ACC of the model can reach 92.35%. The use of this system for psychotherapy has a significant effect, and 45% of the students are very satisfied with the human-computer interaction of the system. This study aims to guess the psychology of students through emotion recognition and reduce human participation based on the human-computer interaction, which can provide a new research idea for college psychotherapy. At present, the mental health problems of college students cannot be ignored; especially every year, there will be news reports of college students' extreme behaviors due to depression and other psychological problems. An interactive college student emotion recognition and psychological intervention system based on convolutional neural network (CNN) is established. This system uses face recognition as the basic support technology and uses feature recognition algorithms to effectively classify different students. An interaction technology-based psychological intervention platform is designed and compared with existing systems and models to further verify the effectiveness of the proposed system. The results show that deep learning has better student emotion recognition ability than backpropagation neural network (BPNN) and decision tree algorithm. The recognition accuracy is up to 89.32%. Support vector machine algorithm is employed to classify emotions, and the recognition acceptability rate increases by 20%. When K is 5 and d is 8, the acceptability rate of the model can reach 92.35%. The effect of this system in psychotherapy is remarkable, and 45% of students are very satisfied with the human-computer interaction of this system. This work aims to speculate students' psychology through emotion recognition, reduce people's participation via human-computer interaction, and provide a new research idea for university psychotherapy.

1. Introduction

With the rapid economic development, the society has undergone major changes, and various pressures such as life, employment, development, and learning have followed one another [1]. College is equivalent to half of society. As the main force of the next generation of society, college students are also faced with social, graduation, study, and life

problems. These pressures invisibly bring serious mental obstacles to the positive development of students [2]. Thus, the importance of mental education and mental consultation in colleges and universities has become increasingly prominent [3]. Studies have shown that 23% of students in Chinese colleges and universities have different levels of mental problems, which proves that the problem of mental health of college students in China is becoming more serious. Therefore, timely research on the mental health of college students is of great significance for talent training and social development [4]. United Nations (UN) experts have predicted that "from the 21st century, mental problems will be the chief culprit that plagues people's development and causes many other problems" [5]. The World Health Organization (WHO) has estimated that six of the top 20 major diseases will be mental diseases by 2020, accounting for 17.4% of all diseases [6]. As far as the level of mental health teaching in Chinese universities is concerned, some students with mental problems are unable to receive effective mental counseling and treatment in time due to the lack of mental counselors, which in turn leads to more students jumping off buildings and committing suicide [7]. Therefore, how to scientifically and effectively solve the shortcomings in mental counseling and treatment in colleges and universities has become the core issue discussed in the current mental education field.

Early warning of mental health is to timely discover and identify the factors affecting the mental health crisis through the identification and dynamic detection of the early warning object based on the analysis and processing of the object's behavior, emotions, and other data, so as to take timely preventive measures to reduce the more serious consequences caused by mental problems [8]. Therefore, it is widely used in various universities and enterprises. The early warning of mental health for students in colleges and universities is mainly realized through the relationship of student dormitory, class atmosphere, college system, and school pressure [9]. Wang et al. investigated on the mental health of some patients with diseases, revealing the protective factors related to the risk reduction of mental health problems in patients with diseases. This work revealed the high incidence of mental health problems among patients with diseases and the gap in mental health services, which also indicates the high degree of pain caused by the risk of disease environment and increased stress [10]. Hong et al. examined the mediating role of rumination and the moderating role of mindfulness in the association between social media exposure and COVID-19 information and psychological distress. A questionnaire survey was conducted among local college students to gain a better understanding of the development of psychological symptoms during the COVID-19 pandemic and provide insights into effective interventions for adverse mental health outcomes among college students [11]. Individuals and groups with abnormal psychology are screened, the student groups that may have mental problems are found and analyzed, the early warning messages are given based on their recent behaviors, and their behaviors are evaluated and analyzed by the mental processing centers of universities to achieve effective prevention of mental health crises [12]. Being different from students in other disciplines, physical education students suffer from greater employment restrictions, so they are more likely to have mental problems [13]. At present, early warning system of mental health in college often suffers from high cost, poor predictive performance, and more complicated system, which greatly hinder the development of college mental health education, making it difficult for early warning system of mental health in college to be applied in more colleges and universities [14].

In order to effectively provide early warning of mental health for physical education students and prevent college students from greater tragedies due to mental problems, a mental health prediction and treatment system for contemporary college students is constructed in this study starting with the existing problems of the existing college mental health early warning system and using the big data and artificial intelligence technology. Through data analysis, a database under different mental conditions is established, which can achieve effective human-computer interaction with the help of data. Therefore, this study can provide new research ideas and methods for mental health education in colleges. Firstly, the concept of deep learning-based emotion recognition algorithm is introduced, and the methods and measures of feature extraction of mental health data are explained. Secondly, the technical strategy of psychotherapy based on artificial intelligence is illustrated, and data training and parameter optimization are explained. Finally, the established model is employed to conduct performance evaluation and interaction analysis on the psychological parameters of college students to obtain the research results, and then the conclusion is drawn. This investigation provides some reference theory for college students' mental health diagnosis and treatment system. The innovation lies in the introduction of a psychological feature extraction method as the main basis of mental health judgment.

2. Literature Review

As places for talent training, colleges and universities have to understand the critical significance of big data to the mental health work [15]. Mental health educators in colleges and universities also need to develop the habit of emphasizing data investigation and forming data thinking and collect more relevant data through mental consultation and mental teaching of students [16]. There are many related studies on big data and mental health. Kulesza et al. used social media big data, natural language processing, and machine learning techniques to achieve mental health monitoring [17]. Huang et al. established a multiple regression model to examine the combined effects of fear and collectivism on the public's prevention intentions for COVID-19. The simple slope test was adopted to examine the interaction between fear and collectivism on preventive intention. It was found that fear and collectivism reduced the positive impact of each other on people's preventive intention [18]. Simon used big data to study the relationship between genome and mental health, which closely linked the genetic inheritance and mental health [19]. Thus, it can

be concluded that the use of big data analysis methods can help people deeply analyze the behavior and thinking of the human body, establish a mental health knowledge system, and provide a basis for establishing an early warning system of mental health.

In recent years, with the continuous in-depth research and rapid development as well as the application of expert systems in the field of medical diagnosis, a good foundation has been laid for the development of intelligent mental consultation [20]. The intelligent auxiliary diagnosis system can give full play to the role of domain experts in diagnosis through the effective acquisition, transmission, processing, and utilization of diagnosis information [21, 22]. There are many related research studies on artificial intelligence and mental health. D'Alfonso et al. enhanced the user participation with computing and artificial intelligence methods and further modified the system and improved the ability of young people to cope with the mental stress with novel mechanisms [23]. Graham et al. discussed in detail the various applications of artificial intelligence in mental health, further proving the potential of applying machine learning algorithms to solve mental health problems [24]. Kalmady et al. found that machine learning techniques can be well applied to mental health analysis research based on the combination of machine learning and functional magnetic resonance imaging system [25]. Graham et al. demonstrated the potential of using machine learning algorithms to address mental health problems, identify these disorders at an early or progenitor stage when intervention may be more effective, and personalize treatment regarding an individual's unique characteristics. However, caution is required to avoid over-interpreting preliminary results, and more work should be done to bridge the gap between mental health research and artificial intelligence in clinical care [24]. Laacke et al. introduced the (ethical) discussion of artificial intelligence in medicine, with reference to mental health. Two AIDD models using social media data and different usage scenarios were proposed, and the extended concept of health-related digital autonomy was proposed [26]. Adikari et al. proposed a comprehensive framework for emotion modeling and analysis under self-structured artificial intelligence. The framework systematically integrates modeling capabilities for unstructured and untagged social media data at a granular level. The validity and effectiveness of the application in real social media environment further proves the methodological novelty of this set of self-organizing artificial intelligence for deep emotions [27]. The above research studies show that artificial intelligence technology can play a very good auxiliary role in the field of mental health, effectively reduce the mental stress, and enhance the ability to cope with difficulties.

3. Methods and Design

3.1. Emotion Recognition Algorithm Based on Deep Learning. DCNN is featured with a convolutional neural network (CNN) structure with discriminative deep structure and is considered to be the first deep network learning algorithm [28]. CNN imitates the response of biological neurons to external stimuli to achieve the image recognition [29]. The convolutional layer is the feature extraction layer of CNN, and the features of the original input data are extracted through the convolution operation of this layer. The input image or the output of the pooling layer is adjusted by the bias function, and the nonlinear expression of the network is enhanced by the activation function, which can be expressed as

$$\mathrm{ConvLa}y_j^l = g\left(\sum_{i\in M_j} I_i^{l-1} \otimes W_{ij}^l + B_j^l\right), \tag{1}$$

where $g()$ represents the activation function, I_i^{l-1} is the input of the l-1th layer to the jth neuron, W_{ij}^l refers to the convolution kernel of the connected neurons between the l-1th layer and the lth layer, $\otimes$ refers to the convolution operation among connected neurons, and B_j^l represents the offset parameter amount. After the convolution operation in convolution layer, dimension of the feature is very large. When the learned features are directly used to train the classifier in the system, it can cause over-fitting of the model, so that generalization ability of the model is low [30]. Therefore, dimensionality of the feature is reduced through downsampling after the convolutional layer to inhibit the over-fitting.

At present, this algorithm has the following advantages:

(1) *Preweighting*. The high-frequency part of music signal generally has low energy, so it is necessary to filter out the low-frequency interference to obtain the high-frequency spectrum more easily.

(2) *Framing*. Music signal is a kind of nonstationary signal, but due to the related reasons of human vocal organs, music signal has the nature of being short-term stationary. Therefore, it needs to cut the audio into relatively smooth segments. The length of the cut should not only ensure the stability of the fragment but also be easy to handle. Music signals often appear to be some large peak signals, so the signal is not smooth enough to process. Hence, it is necessary to add windows to the frame signal to smooth the signal.

(3) *Mute Frame Detection*. Sometimes there are mute frames without signal parameters in music signal, which affect the result of recognition. A threshold based on the short time energy of the audio signal can be set in this section to detect silent frames. If the energy is below the set threshold, it will be removed.

(4) *Feature Extraction*. The quality of extracted feature is directly related to the output result. Features are one of the most used features in acoustic tasks. The human cochlea acts as a set of filters that distinguish between different sounds by filtering characteristics such as frequency bands and bandwidths. Traits mimic this ability.

At the same time, there are some drawbacks, such as the high dimension of the original time-domain features, as well

as a lot of redundancy and noise. Therefore, it is necessary to reduce the dimension of the input training data. Principal component analysis is adopted. Principal component analysis is a multivariate statistical method to investigate the correlation between multiple variables. It explores how to reveal the internal structure of multiple variables through a few principal components.

Moreover, when the number of samples is very small, the feature performance extracted by using features is not as good as general features. This also confirms that the deep learning network needs a certain number of samples for training in order to better fit the test samples. Due to the problem of parameter setting, its accuracy has not been greatly improved.

3.2. Feature Extraction of Mental Health Data. During analysis of the mental health of college students, the more the analysis items, the more they truly reflect the mental health of the student [31]. In order to reduce the amount of calculation, support vector machine algorithm is adopted to reduce the dimensions of many extracted mental health data of college students. Support vector machine is a learning model that analyzes data and recognizes patterns. It is based on the statistical learning related theories and is particularly good at processing small samples of nonlinear data [32]. The core of the support vector machine lies in the kernel function, and it can convert the samples in the low-dimensional space nonlinearly into the high-dimensional space by mapping the kernel function. The optimal classification function of the support vector machine is defined as follows:

$$f(x) = \operatorname{sgn}\left(\sum_{i=1}^{n} \alpha_i y_i K(x_i, x) + b\right). \tag{2}$$

The main steps for support vector machine to realize model classification are given as follows: input data of the support vector machine are defined, a set of training samples x_i and y_i are given (of which, y_i refers to the type of training sample x_i), maximum value of the function is calculated, and α_i is obtained. Such process has to meet the following constraints:

$$\sum_{i=1}^{n} \alpha_i y_i = 0, 0 \le \alpha_i \le C, i = 1, 2 \ldots n. \tag{3}$$

Function $Q(\alpha)$ can be calculated with the following equation:

$$Q(\alpha) = \sum_{i=1}^{n} \alpha_i - \frac{1}{2} \sum_{i,j=1}^{n} \alpha_i \alpha_j y_i y_j K(x_i, x_j). \tag{4}$$

Then, W and b can be calculated as follows:

$$\begin{aligned} W &= \sum_{i=1}^{n} \alpha_i y_i x_i, \\ b &= \frac{1}{y} - W \cdot x_s. \end{aligned} \tag{5}$$

In the above equations, x_s is a specific support vector, a certain kernel function $K(x,x_i)$ is selected to calculate the optimal classification function of the support vector machine for the test sample x, and the calculation result is +1 or −1, which can be taken as the basis for category judgment.

3.3. Mental Treatment Based on Artificial Intelligence. Based on the analysis of the mental health of physical education students, an artificial intelligence mental consultation platform is established. The main functional modules to be implemented include system maintenance, mental knowledge presentation, auxiliary decision-making diagnosis, artificial intelligence mental consultation, and consultation result analysis (Figure 1).

The functional structure is divided into two layers: system management and maintenance layer and system user layer. The system management and maintenance layer refers to the management and maintenance on the server side. It is mainly designed for knowledge engineers and system maintenance personnel, who can operate all module functions of the system. The system user layer is mainly for remote access of ordinary users, who can only realize the mental knowledge view, auxiliary decision-making diagnosis, artificial intelligence mental consultation, and consultation result analysis functions (Figure 2).

3.4. Data Training and Parameter Optimization

Data Source. The dataset applied in this study is the Kaggle facial emotion recognition dataset (Fer-2013), which was released by Pieere and Aaron at the ICML2013 seminar and consisted of 35,887 facial emotion pictures. There are 4,953, 547, 5,121, 8,989, 6,077, 4,002, and 6,198 pictures on angry, disgust, fear, happiness, sadness, surprise, and neutral emotions [33]. The dataset consists of three parts: the training set contains 28,709 pictures, the validating set contains 3,589 pictures, and the testing set contains 3,589 pictures. Part of the dataset is shown in Figure 3.

The Model training Caffe's framework is adopted in the experiment, the operating system is Ubuntu16.04, the memory is 8G, the central processing unit (CPU) is Intel Core i5-4590@3.30 GHz × 4, and the graphics processing unit (GPU) is GTX1050Ti. The training learning rate is 0.001, momentum is 0.9, Test_interval is 10,000, and the learning strategy is stochastic gradient descent (SGD). Facial expression recognition has related concepts in computer and psychology. In the field of computer science and technology, emotion recognition on account of emotion computing is an important part of the emotion computing program, and the individual's emotional and emotional state can be judged according to the research results of the above content. In the field of psychology, facial emotion recognition refers to the recognition of the specific facial emotional state of the observed object, thereby determining the psychological emotional state of the recognized object currently. For facial emotion recognition, the operational definition of facial emotion recognition is as follows. The recognition responses of individuals to face pictures given in the experiment

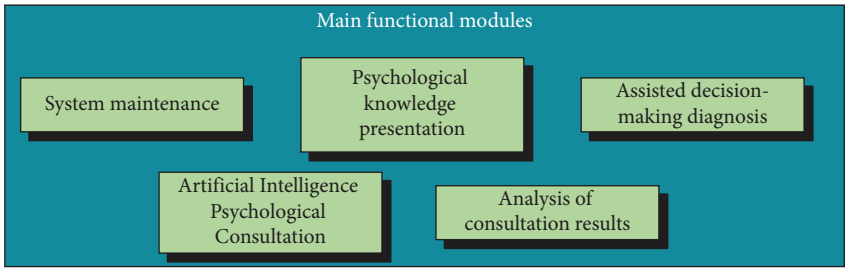


Figure 1: Main functional modules.

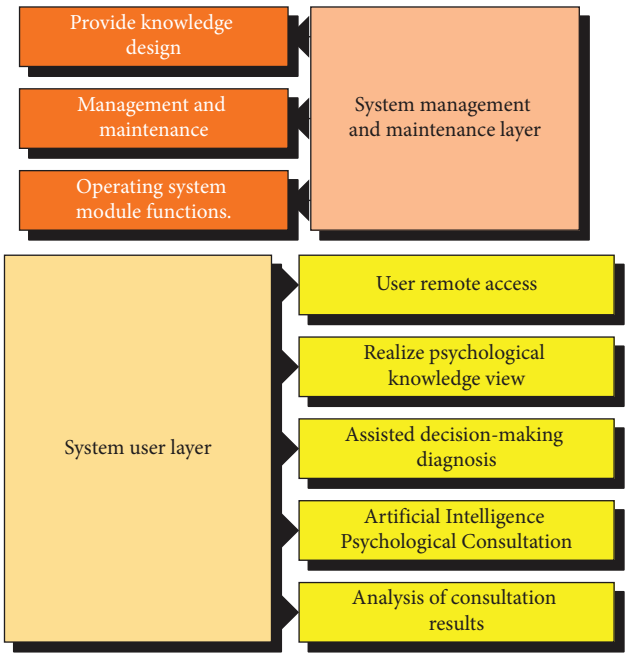


Figure 2: Hierarchy function diagram.

include response speed and response accuracy. The pictures with obvious emotional characteristics in the standard emotion library are displayed, including positive emotions (smile), negative emotions (sad and scared), and neutral emotions (blank expression), and then subjects are asked to identify the emotional pictures and press the corresponding response button at the same time. The difference of emotion recognition is measured by individual recognition response time and accuracy.

Due to the technological changes brought about by the emergence of computers, the research on facial emotion recognition technology began to emerge and prevail in recent decades. In the process of cognitive science research, due to the diversity of facial expressions and the complexity of the range involved, it has been difficult to identify emotions in the past, and it is extremely difficult to study emotions, especially facial emotions. Therefore, compared with other works involving biology and psychology, the development of emotion recognition is relatively slow. Its results not only are few but also involve only limited applications.

In the influencing factors of facial emotion recognition, the influence of facial muscle cues on emotion recognition has been mentioned. In this experiment, a body cognition-based experimental paradigm is utilized, and facial manipulation is used to influence the perception of emotion. It is found that when they held the pen in their mouth, they rated the cartoons funnier. Other works, such as using emotional understanding of visual images as a dependent variable, have used the same face manipulation paradigm. The results show that facial manipulation can significantly affect emotional understanding of images. After comparison of the response times, it is proved that the response time of the positive picture is shorter than that of the negative picture when biting the pen triggered the smile expression. There is no significant difference in response time between the two images when suppressing the smiling expression.

Parameter Optimization. The sample dataset is kept fixed in the experiment. The mental health data structure of each student is composed of a one-dimensional column vector of 11 elements, and the maximum dimension is 11, so the range of d can be determined as [2, 10], and the step size is 1; the value interval of the neighbor parameter K is [3, 10], and the step size is 1. In the experiment, SVM is undertaken as the classifier, and its kernel function is class support vector classification (C-SVC). The specific values of the parameters involved are degree = 3, gamma = 0.5, $\text{coef}_0 = 0$, and penalty factor $C = 1$.

3.5. Performance Evaluation and Interactive Analysis

Performance Evaluation. In mental health prediction, the most common algorithms are BPNN and DT algorithm [34], while locally linear embedding (LLE) algorithm is recently published in many articles with the ACC as high as 92% [35]. For this reason, different algorithms are applied for comparative analysis. In order to analyze the performance of the model more directly, ACC is adopted to represent the accuracy of the model in the testing set, and F1 is undertaken as the comprehensive performance evaluation index of the model, which is the harmonic mean of the ACC and the recall (Rec) [36]. The specific equation is as follows:

$$\text{ACC} = \frac{\sum_{i=1}^{s} Tp_i}{\sum_{i=1}^{s} (Tp_i + \text{FN}_i)}, \tag{6}$$

$$F_1 = \frac{2 * \text{Pre} * \text{Rec}}{\text{Pre} + \text{Rec}}. \tag{7}$$

The facial emotions in the testing set are divided into categories, Tp_i is used to represent the number of correctly

Figure 3: Some samples from Fer-2013.

Table 1: Questionnaire validity evaluation.

Content	Level		
	Very high (%)	High (%)	Low (%)
Content evaluation	45	52	3
Quantity evaluation	39	60	1
Structure evaluation	26	59	15

recognized emotions in a category, and FN_i represents the number of incorrectly recognized emotions in this category [37–39]. In (7), *Pre* refers to the precision, which means the proportion of all samples classified as positive, which are correctly classified, and *Rec* refers to the recall, which means the proportion of samples classified as positive in the actual positive samples.

Interactive Analysis. 120 students majoring in physical education in a university are selected and divided into two groups. One group is named as the experimental group (using an interactive system for teaching), and the other group is named as the control group (with normal mental counseling), which is accepted with a three-day mental counseling. When the experiment is over, questionnaires are distributed and collected to evaluate the two teaching methods [40–43]. The questionnaires are based on online questionnaires, covering three aspects: the effect, role, and acceptance of artificial intelligence mental counseling. 110 questionnaires are received, including 105 valid questionnaires. Then, the data are tested with the reliability and validity analysis, correlation analysis, and hypothesis test, so as to obtain the analysis results of the effect of human-computer interaction psychotherapy. The evaluation results of validity and reliability of the questionnaire are shown in Table 1.

Standardized questionnaires must be tested for reliability and validity before they are used. If small sample surveys are used to predict target groups, they have higher requirements to produce questionnaires or they are prone to statistical errors. According to the analysis in Table 1, the validity evaluation results of the questionnaire are good, so it can be used in the mental health investigation of college students.

4. Results and Analysis

4.1. Comparison of Facial Emotion Recognition Feature Algorithms. As shown in Figure 4, different algorithm models show greater differences in the training set and small differences in the testing set. The traditional back-propagation neural network (BPNN) and decision tree (DT) algorithm have the worst performance in the early warning of mental health for college students. After the locally linear embedding (LLE) algorithm with local linear embedding is adopted, performance of the model has been improved with an average improved adaptive cruise control (ACC) of 2%. When the SVM algorithm is added, the average ACC has increased by 3.47%. DCNN algorithm has the best performance with an average ACC of 81.86%. After the SVM classification is added, the performance of the algorithm has been greatly improved, so that the highest ACC can reach 94% with an average ACC of 87.28%. The above results show that the deep learning significantly improves the ACC of model recognition through data mining compared with traditional algorithms, and the SVM classification performance can complement the over-fitting of DCNN, thereby significantly improving the ACC of the model.

4.2. Determination of Optimal Parameters for Early Warning of Mental Health. As shown in Figure 5, if k is too small, it will affect the global performance of the model, and if k is too large, the original nonlinear characteristics will not exist after the data are reduced. The data in the status space may overlap if d is too small, and the noise will be introduced in the reduced dimensionality dataset so that the ACC is decreased if d is too large. It is found that the highest ACC of the model is 94.5% when k is 5 and d is 8, so they are applied for model training and early warning in subsequent experiments so as to accurately assess the mental health of physical education students.

4.3. Psychotherapy Effect of Human-Computer Interaction System. Figure 6 suggests that most of the students are quite satisfied with the artificial intelligence mental counseling, and the proportion of "very satisfied" is the highest (45%). The satisfaction to the artificial intelligence method is

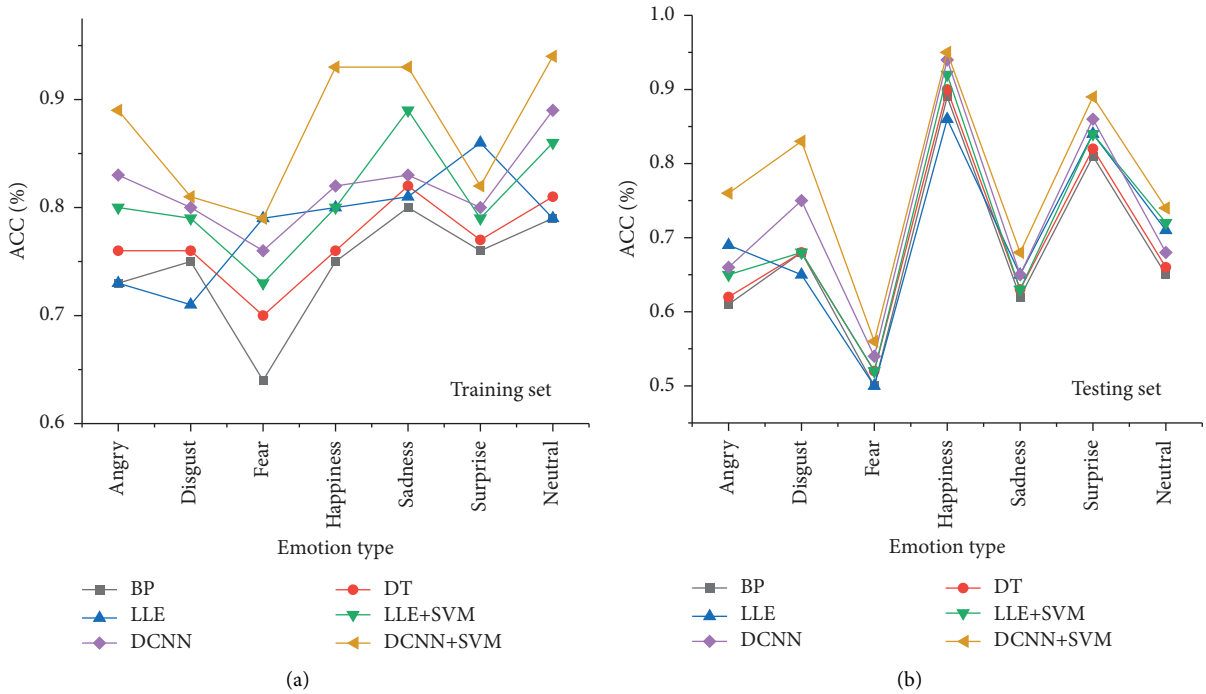
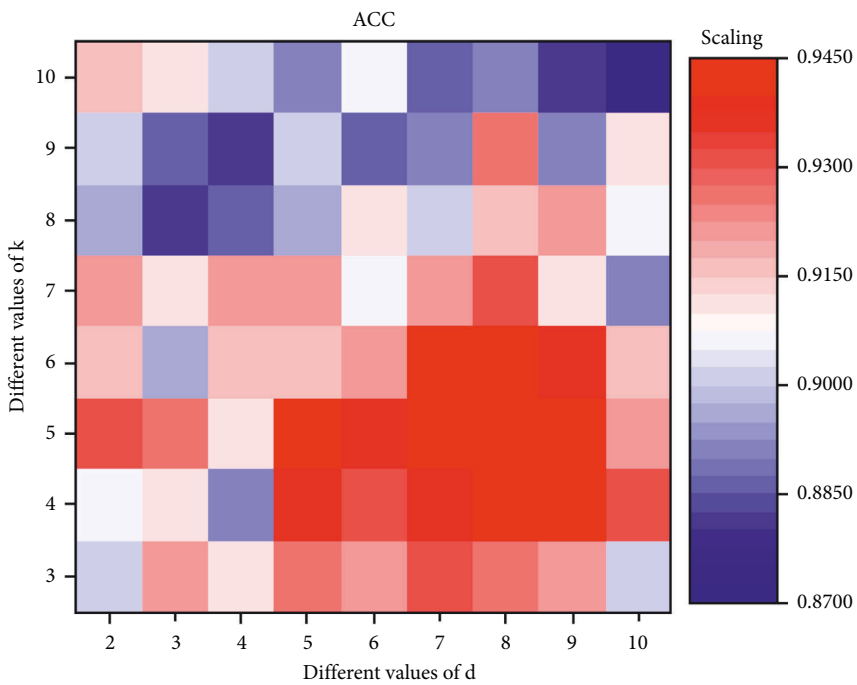
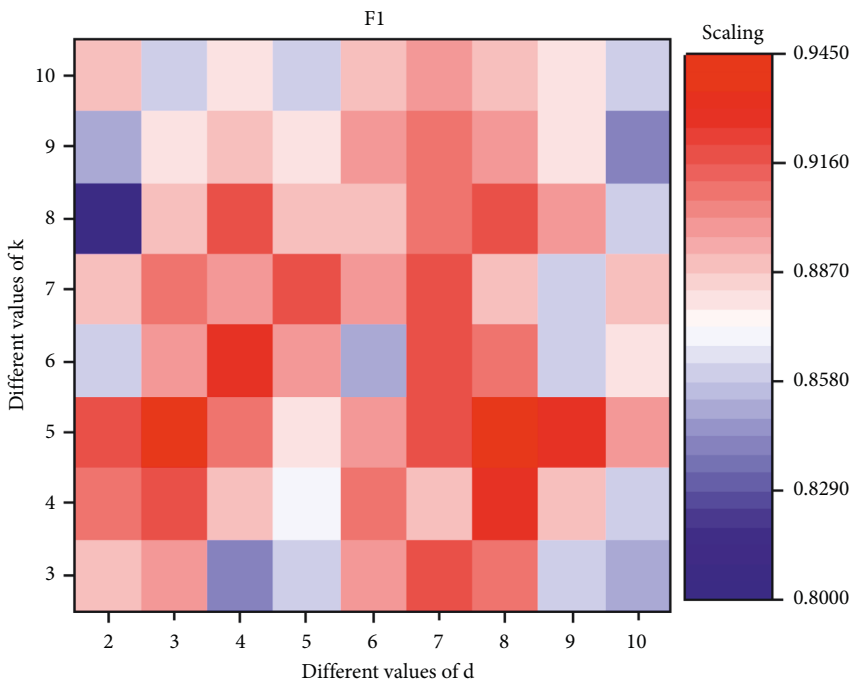


Figure 4: Comparison results of facial emotion recognition feature algorithms. (a) The numerical results of training. (b) The numerical results of testing.



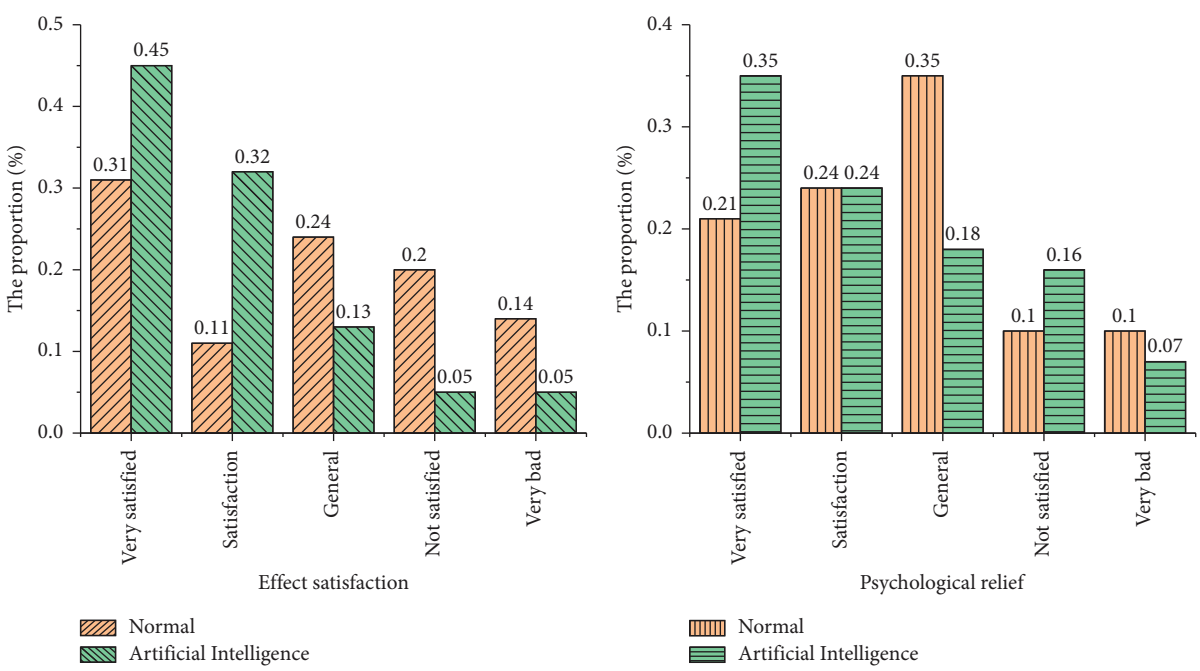
(a)

Figure 5: Continued.



(b)

Figure 5: Determination of optimal parameters of early warning model of mental health. (a) The numerical results of training. (b) The numerical results of testing.



(a) (b)

Figure 6: Continued.

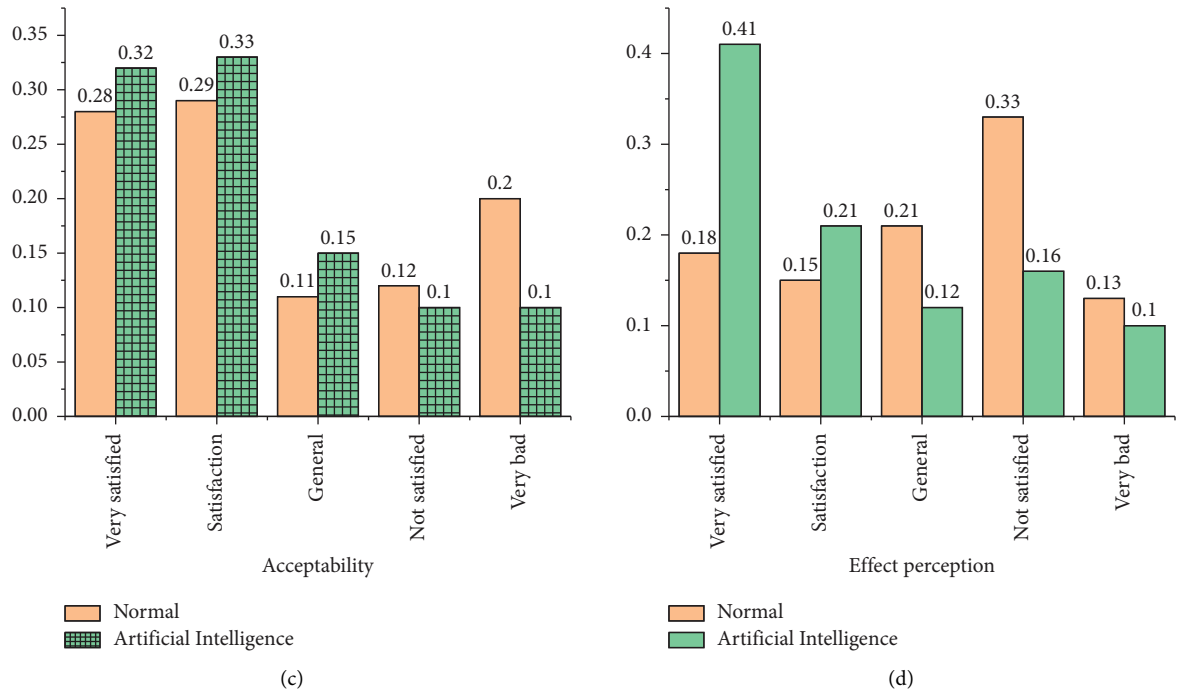


Figure 6: Psychotherapy effect analysis of human-computer interaction system. (a) The result of effect satisfaction. (b) The result of psychological relief. (c) The result of acceptability. (d) The result of effect perception.

obviously higher than that of professional teachers. General mental questions are related to privacy, and the application of human-computer interaction can solve it effectively. On the other hand, human-computer interaction is more targeted and can effectively analyze the mental problems of physical education students. Compared with the traditional way, 30% of students think that the human-computer interaction is acceptable. After mental counseling with artificial intelligence, 41% of students think that they are more comfortable mentally, while only 18% of students think the traditional way of mental counseling is more comfortable. Based on the above results, it is found that the use of artificial intelligence for psychotherapy for physical education students is more effective, and this method is also widely accepted by students. This system plays a good auxiliary role in daily learning and making friends of physical education students.

5. Discussion

By using different algorithms, it is found that the performance of the model can be significantly affected. SVM and DCNN algorithms can achieve the highest performance of the model. It is also found that different values of k and d will affect the data in space, resulting in data overlap. Hence, the data noise is too large, affecting ACC value. After the questionnaire is utilized to investigate the psychological status of the students, it is found that the satisfaction of the artificial intelligence method is significantly higher than that of professional teachers. General psychological problems are related to privacy, and the application of human-computer interaction can be effectively solved. Moreover, human-computer interaction is more targeted and can effectively analyze the psychological problems of sports students.

6. Conclusion

Based on the analysis of the current research status of mental health and the construction of mental early warning systems for college students, an emotional recognition system with higher recognition ACC is established for physical education students through deep learning of the face database. Compared with other algorithm models, the established system has the ability of learning and analysis. Through the optimal analysis of parameters, its mental health prediction performance has reached a better state. With the help of artificial intelligence technology, the automatic early warning treatment system of mental health teaching has achieved better practical results. An automatic early warning platform for mental health education is constructed for physical education students, but there are still many shortcomings in the study. Firstly, it considers the effect of activation function and model structure on the performance and optimizes the main parameters, but the other parts of the current neural network can significantly improve the prediction performance of model. Secondly, it can consider using data mining technology for a more in-depth analysis of the acquired face data to improve the consulting effect of human-computer interaction and establish different communication methods for different emotions to avoid more mechanized communication. Mechanized communication is a type of queuing system utilized to describe the communication recovery process of mechanized repair units. The working algorithm of the simulation model of the

recovery process of the communication means of the mechanized braiding unit is given. Therefore, how to better improve the performance and calculation accuracy of the model will be the main research objective in the subsequent research work. In the follow-up, further analysis and development will focus on these two aspects, so as to provide more effective diagnosis and treatment systems for college students of different majors.

Data Availability

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Consent

Informed consent was obtained from all individual participants included in the study.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

The authors acknowledge the help from the university colleagues.

References

[1] D. Gallie, "Research on work values in a changing economic and social context," *The Annals of the American Academy of Political and Social Science*, vol. 682, no. 1, pp. 26–42, 2019.

[2] J. C. Vela, W. D. Smith, J. F. Whittenberg, R. Guardiola, and M. Savage, "Positive psychology factors as predictors of Latina/o college students' psychological grit," *Journal of Multicultural Counseling and Development*, vol. 46, no. 1, pp. 2–19, 2018.

[3] G. . K. Kassymova Tokar, O. V. Tashcheva, A. I. Slepukhina et al., "Impact of stress on creative human resources and psychological counseling in crises," *J. International journal of education and information technologies*, vol. 13, pp. 26–32, 2019.

[4] Y. Song, "Gratifications for social media use in entrepreneurship courses: learners' perspective," *Frontiers in Psychology*, vol. 10, pp. 1270–1276, 2019.

[5] C. Oduaran and N. Okorie, "Community radio, family and psychological support for sexual harassment issues: a study of Yoruba usage," *J. Media Watch.*vol. 9, pp. 291–301, 2018.

[6] S. I. Ahmad, B. L. Leventhal, B. N. Nielsen, and S. P. Hinshaw, "Reducing mental-illness stigma via high school clubs: a matched-pair, cluster-randomized trial," *Stigma and Health*, vol. 5, no. 2, pp. 230–239, 2020.

[7] R. Yang, X. You, Y. Zhang, L. Lian, and W. Feng, "Teachers' mental health becoming worse: the case of China," *International Journal of Educational Development*, vol. 70, pp. 102077–102082, 2019.

[8] M. Kalkbrenner and T. J. Hernández, "Community college students' awareness of risk factors for mental health problems and referrals to facilitative and debilitative resources," *Community College Journal of Research and Practice*, vol. 41, no. 1, pp. 56–64, 2017.

[9] M. Chen, "The impact of expatriates' cross-cultural adjustment on work stress and job involvement in the high-tech industry," *Frontiers in Psychology*, vol. 10, pp. 2228–2234, 2019.

[10] Y. Wang, Z. Duan, Z. Ma et al., "Epidemiology of mental health problems among patients with cancer during COVID-19 pandemic," *Translational Psychiatry*, vol. 10, pp. 263–310, 2020.

[11] W. Hong, R. D. Liu, Y. Ding, X. Fu, R. Zhen, and X. Sheng, "Social media exposure and college students' mental health during the outbreak of CoViD-19: the mediating role of rumination and the moderating role of mindfulness," *Cyberpsychology, Behavior, and Social Networking*, vol. 24, no. 4, pp. 282–287, 2021.

[12] J. Wang, Z. Zhang, H. Luo, Y. Liu, W. Chen, and G. Wei, "Research on early warning model of college students' psychological crisis based on genetic BP neural network," *American Journal of Applied Psychology*, vol. 8, pp. 120–128, 2019.

[13] M. Huhtiniemi, A. Sääkslahti, A. Watt, and T. Jaakkola, "Associations among basic psychological needs, motivation and enjoyment within Finnish physical education students," *Journal of sports science & medicine*, vol. 18, pp. 239–246, 2019.

[14] F. Liu, N. Zhou, H. Cao et al., "Chinese college freshmen's mental health problems and their subsequent help-seeking behaviors: a cohort design (2005-2011)," *PLoS One*, vol. 12, no. 10, pp. 0185531–0185539, 2017.

[15] A. Kaplan and M. Haenlein, "In my hand: who's the fairest in the land? On the interpretations, illustrations, and implications of artificial intelligence," *Business Horizons*, vol. 62, no. 1, pp. 15–25, 2019.

[16] N. Shah, Z. Irani, and A. M. Sharif, "Big data in an HR context: exploring organizational change readiness, employee attitudes and behaviors," *Journal of Business Research*, vol. 70, pp. 366–378, 2017.

[17] M. Kulesza, A. Odrowaz-Coates, and A. Perkowska-Klejman, "Suicide rates amongst adolescents A mental health practitioner's perspective in Poland and a global, Big Data context," *Resocjalizacja Polska*, vol. 21, pp. 461–486, 2021.

[18] F. Huang, H. Ding, Z. Liu et al., "How fear and collectivism influence public's preventive intention towards COVID-19 infection: a study based on big data from the social media," *BMC Public Health*, vol. 20, pp. 1707–1709, 2020.

[19] G. E. Simon, "Big data from health records in mental health care: hardly clairvoyant but already useful," *JAMA Psychiatry*, vol. 76, no. 4, pp. 349-350, 2019.

[20] S. H. Park and K. Han, "Methodologic guide for evaluating clinical performance and effect of artificial intelligence technology for medical diagnosis and prediction," *Radiology*, vol. 286, no. 3, pp. 800–809, 2018.

[21] S. H. Park and H. Y. Kressel, "Connecting technological innovation in artificial intelligence to real-world medical practice through rigorous clinical validation: what peer-reviewed medical journals could do," *Journal of Korean Medical Science*, vol. 33, no. 22, pp. e152–e254, 2018.

[22] W. Wu, H. Wang, C. Zheng, and Y. J. Wu, "Effect of narcissism, psychopathy, and machiavellianism on entrepreneurial intention—the mediating of entrepreneurial self-efficacy," *Frontiers in Psychology*, vol. 10, pp. 360–368, 2019.

[23] S. D'Alfonso, O. Santesteban-Echarri, S. Rice et al., "Artificial intelligence-assisted online social therapy for youth mental health," *Frontiers in Psychology*, vol. 8, pp. 796–806, 2017.

[24] S. Graham, C. Depp, E. E. Lee et al., "Artificial intelligence for mental health and mental illnesses: an overview," *Current Psychiatry Reports*, vol. 21, no. 11, pp. 116–121, 2019.

[25] S. V. Kalmady, R. Greiner, R. Agrawal et al., "Towards artificial intelligence in mental health by improving schizophrenia prediction with multiple brain parcellation ensemble-learning," *Npj Schizophrenia*, vol. 5, pp. 2–11, 2019.

[26] S. Laacke, R. Mueller, G. Schomerus, and S. Salloch, "Artificial intelligence, social media and depression. A new concept of health-related digital autonomy," *The American Journal of Bioethics*, vol. 21, no. 7, pp. 4–20, 2021.

[27] A. Adikari, G. Gamage, D. de Silva, N. Mills, S. M. J. Wong, and D. Alahakoon, "A self structuring artificial intelligence framework for deep emotions modeling and analysis on the social web," *Future Generation Computer Systems*, vol. 116, pp. 302–315, 2021.

[28] K. H. Jin, M. T. McCann, E. Froustey, and M. Unser, "Deep convolutional neural network for inverse problems in imaging," *IEEE Transactions on Image Processing*, vol. 26, no. 9, pp. 4509–4522, 2017.

[29] H. Liang and Q. Li, "Hyperspectral imagery classification using sparse representations of convolutional neural network features," *Remote Sensing*, vol. 8, no. 2, pp. 99–106, 2016.

[30] Q. Xu, M. Zhang, Z. Gu, and G. Pan, "Overfitting remedy by sparsifying regularization on fully-connected layers of CNNs," *Neurocomputing*, vol. 328, pp. 69–74, 2019.

[31] R. Rogoza, M. Żemojtel-Piotrowska, M. M. Kwiatkowska, and K. Kwiatkowska, "The bright, the dark, and the blue face of narcissism: the Spectrum of narcissism in its relations to the metatraits of personality, self-esteem, and the nomological network of shyness, loneliness, and empathy," *Frontiers in Psychology*, vol. 9, pp. 343–356, 2018.

[32] S. Huang, N. Cai, P. P. Pacheco, S. Narrandes, Y. Wang, and W. Xu, "Applications of support vector machine (SVM) learning in cancer genomics," *CANCER GENOMICS and PROTEOMICS*, vol. 15, no. 1, pp. 41–51, 2018.

[33] K. S. Gautam and S. K. Thangavel, "Video analytics-based facial emotion recognition system for smart buildings," *International Journal of Computers and Applications*, vol. 43, no. 9, pp. 858–867, 2019.

[34] X. L. Feng and J. Gao, "Attendance data analysis with decision regression tree on spark," *Journal of Inner Mongolia University of Technology (Natural Science Edition)*, pp. 7–14, 2018.

[35] S. A. Vella, C. Swann, M. Batterham et al., "Ahead of the game protocol: a multi-component, community sport-based program targeting prevention, promotion and early intervention for mental health among adolescent males," *BMC Public Health*, vol. 18, pp. 390–412, 2018.

[36] D. Chicco and G. Jurman, "The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation," *BMC Genomics*, vol. 21, no. 1, pp. 6–13, 2020.

[37] J. J. Chen, T. N. Jiang, and M. F. Liu, "Family socioeconomic status and learning engagement in Chinese adolescents: the multiple mediating roles of resilience and future orientation," *Frontiers in Psychology*, vol. 12, Article ID 714346, 2021.

[38] K. Liu, F. Ke, X. Huang et al., "DeepBAN: a temporal convolution-based communication framework for dynamic WBANs," *IEEE Transactions on Communications*, vol. 69, no. 10, pp. 6675–6690, 2021.

[39] F. Meng, X. Xiao, and J. Wang, "Rating the crisis of online public opinion using a multi-level index system," *The International Arab Journal of Information Technology*, vol. 19, no. 4, pp. 597–608, 2022.

[40] Z. Zhang, J. Tian, W. Huang, L. Yin, W. Zheng, and S. Liu, "A haze prediction method based on one-dimensional convolutional neural network," *Atmosphere*, vol. 12, no. 10, p. 1327, 2021.

[41] K. Shang, Z. Chen, Z. Liu et al., "Haze prediction model using deep recurrent neural network," *Atmosphere*, vol. 12, no. 12, p. 1625, 2021.

[42] Y. Liu, J. Tian, R. Hu et al., "Improved feature point pair purification algorithm based on SIFT during endoscope image stitching," *Frontiers in Neurorobotics*, vol. 16, Article ID 840594, 2022.

[43] Y. Choi, J. Wang, Y. Zhu, and W. F. Lai, "Students' perception and expectation towards pharmacy education: a qualitative study of pharmacy students in a developing country," *Indian Journal of Pharmaceutical Education and Research*, vol. 55, no. 1, pp. 63–69, 2021.

Hindawi
Computational Intelligence and Neuroscience
Volume 2022, Article ID 8431982, 11 pages
https://doi.org/10.1155/2022/8431982

Research Article

Construction and Application of a Piano Playing Pitch Recognition Model Based on Neural Network

Guobin Wu and Wei Chen

Changchun Humanities and Sciences College, Changchun, Jilin 130117, China

Correspondence should be addressed to Wei Chen; chenwei@ccrw.edu.cn

Received 14 July 2022; Revised 16 August 2022; Accepted 20 August 2022; Published 17 September 2022

Academic Editor: Ning Cao

The intonation recognition of piano scores is an important problem in the field of music information retrieval. Based on the neural network theory, this study constructs a piano playing intonation recognition model and uses the optimized result as the feature of piano music to realize the prediction of the music recognition of the intonation preference. The model combines the behavioral preference relationship between intonation and musical notation to measure the similarity between intonations, which is used to calculate the similarity between intonation preference and music, and solves the quantification problem of intonation recognition. In the simulation process, the pitch preference feature of piano playing is used as the identification basis, and the effectiveness of the algorithm is verified through four sets of experiments. The experimental results show that the average symbol error rate of the improved network model is reduced to 0.3234%, and the model training time is about 33.3% of the traditional convolutional recurrent neural network, which is optimized in terms of recognition accuracy and training time in single-class pitch feature. In the recommended method of multi-category evaluation of pitch features, the recognition accuracy of multi-category pitch features is 42.89%, which effectively improves the musical tone recognition rate.

1. Introduction

With the rapid growth of the number of digital music, the piano playing pitch recognition algorithm performs personalized identification by analyzing the historical behavior of pitch-on-demand music [1]. As a hotspot of development in the new century, the neural network has gained more and more attention and application depending on its advantages in nonlinearity, self-learning, robustness, and self-adaptation [2–4]. As one of the most successful models of neural networks, the BP network has been able to simulate complex nonlinear models by virtue of its powerful nonlinear mapping capability, parallel distributed processing capability, adaptive capability, fault tolerance capability, and generalization capability [5]. The error is minimized; therefore, we adopt the BP network as the reference model for musical tone recognition [6].

Traditional machine learning techniques are limited in processing natural data in its raw form [7]. For a long time, the construction of recognition models is very difficult and requires considerable time and effort to complete [8]. In addition, professional knowledge in related fields is also essential, because feature extraction classifier needs assistance [9–11] to perform effective analysis and appropriate feature representation on the original data (such as image pixels and audio signals) [12] for the classification subsystem to analyze the input raw data, and the complexity of the original data requires a lot of time and effort to analyze, and the analysis results are difficult to apply to other data scenarios [13].

In this study, a piano playing intonation recognition model is constructed, and the tool is used to generate the spectrum of piano music. Four music categories are set: blues, classical, jazz, and pop. For the training of CNN (convolutional neural network), two activation functions, ELU (exponential linear unit) and ReLU (rectified linear unit), and two gradient descent methods were explored. Experiments show that ELU is more stable than ReLU, and

Adam has faster gradient descent than RMSProp. Comparing the classification methods based on the spectrum, the latter has an improvement of 1.5% on the basis of the former's classification accuracy of 96%. In addition, aiming at the problem of octave interference in the existing autocorrelation identification method, an improved note identification method is designed using the frameshift method. Taking the comprehensive features as the classification basis, the automatic recognition method of notes is explored as an explicit classification feature.

2. Related Work

The piano tone signal is composed of the fundamental tone and the overtone [14]. Due to the significant differences between Western music and Chinese music in musical styles and modal systems, each has its own independent system in music theory: therefore, the modal tonality analysis module also treats Western music and folk music separately. The output of the module is a clear textual representation indicating the modal key of the music entered into the module for analysis [15].

Román et al. [16] proposed a research method on music score images, including digitization, recognition, and restoration, and summarized the software and hardware required by the pitch recognition system, explored potential application scenarios of pitch recognition, and promoted pitch recognition identification research. Nakamura et al. [17] used a neural network model for optical score recognition and gradually adopted an end-to-end piano score recognition method to simplify the research method based on a general framework. Segmentation is achieved using the region growing method and improved Hough transform detection for the fringes, and the rest of the notes are extracted again through template matching. Watts [18] used a multi-track data set to train and test the model, using a convolution-recurrent neural network that effectively learned both spectral features and audio contextual information. The creative achievement of Qian et al. [19] is mainly to achieve as many as 18 instrument categories at the same time, which also includes the identification of male singers (tenor and bass) and female singers (soprano and alto). The recognition of pure musical instruments also achieves high accuracy.

Although the staff information is the auxiliary information of the note, that is, the position of the note in the staff determines its pitch and other information, to a certain extent, the existence of the staff will interfere with the identification of the note, so the accuracy of staff detection and deletion will directly affect the note [20]. However, the cross-overlapping characteristics of staff and notes make it difficult to delete staff. When the staves are not deleted enough, the remaining staves will affect the recognition of notes. Combining collaborative filtering with intonation context information, the original intonation data scoring model of collaborative filtering is extended to form a three-dimensional data model of intonation-data-context, which expands the correlation dimension and achieves more accurate personalized recognition [21]. In this indirect identification scheme, although there is no correlation between users and data, a third party can be used [22]. Because of the strong correlation between notes, the reconstruction of notes needs to incorporate contextual information. It is not only necessary to combine the relevant information of the staff but also inseparable from the time signature, key signature, clef, and other information in the music score, and grammar-based reconstruction can effectively solve this problem.

3. Construction of a Neural Network-Based Piano Playing Pitch Recognition Model

3.1. Hierarchical Distribution of Neural Network. The neural network only needs to build a suitable learning network, and multiple hidden layers in the network will automatically learn various features, and there is a progressive relationship between the hidden layers, and the latter layer is a more abstract description of the previous layer. The neural network generally consists of three parts: first, the input layer x (c, s); second, the hidden layer v (x); the neural network obtains the features of each part of the original data samples and further compresses and abstracts the features. Third, the output layer g (x–x') compares and classifies the feature results, compares the expected results during training, back propagates the error of the results, and calculates the probability y–s of each classification during prediction.

$$\begin{aligned} &\frac{1-x(c,s)}{N} = \sum \ln(c-1) - r(y-s), \\ &\partial\beta(x)v(x) - y(1-\beta)\left(g\left(x-x'\right)\right) - 1 = 0. \end{aligned} \tag{1}$$

Each node of the hidden layer represents a neuron, each arrow represents a feature transfer function r (c, s) of the data, and the weight carried by each function represents the neuron's "perception" of the data k, and finally, the weight sim (c–1) is adjusted by the error fed back by the output layer, making the neuron's "perception" of the data more accurate. To control each training error sim (c–s), the gradient descent method is generally used to ensure that each iteration training can reduce the error.

$$\begin{aligned} &\overline{r(c,s)} - r(c,s) = \begin{cases} k\sum sim(c-1) \\ 1 - sim(c-s) \end{cases}, \\ &\frac{1}{2}k = \frac{1}{\sum sim(x)} - di\,am(x) - |sim(x)|. \end{aligned} \tag{2}$$

However, activation functions sim (x) such as rectified linear units that introduce sparsity by reducing the coupling ability in certain ranges are equivalent to pre-training in unsupervised learning diam (x). The starting point of the horizontal line on the lower side of the center line of the long ruler was located to the left of the starting point of the horizontal line on the upper side of the center line of the time length ruler, indicating that the note was sung fast.

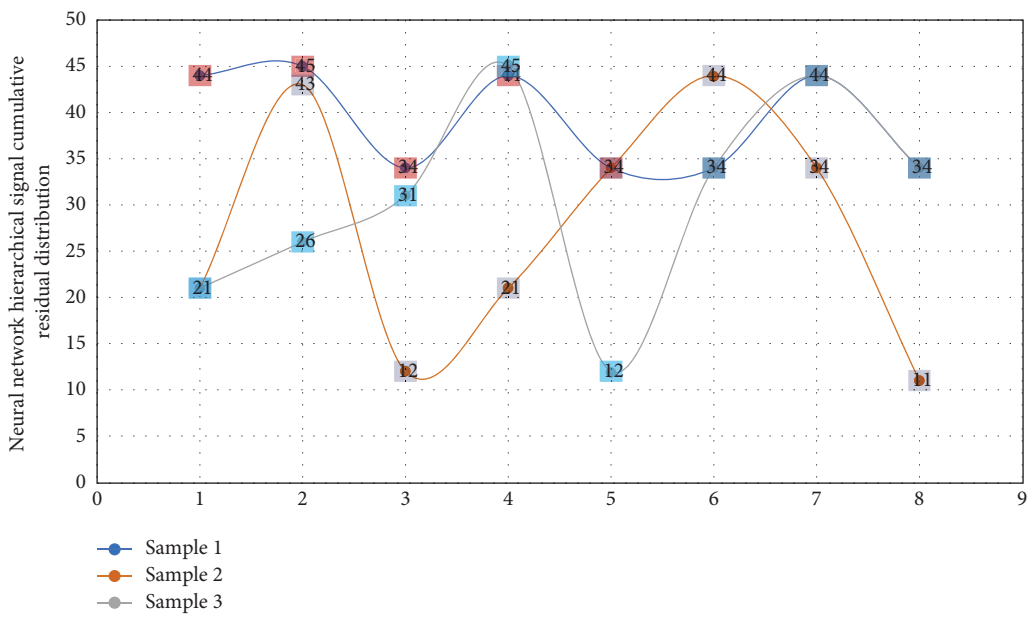


Figure 1: Distribution of cumulative residuals of neural network hierarchical signals.

Obviously, the amount of data after the pooling operation s (x) c (x) is reduced to a quarter of the original. The feature map c (x) obtained after the pooling operation is reduced in dimension compared with the original, which realizes the network dimension reduction.

$$y = \begin{cases} \sum c(x) - |c(x)|, \\ 1 - |s(x)c(x)|, \end{cases} \quad (3)$$
$$C = 1 - |c(1), c(2), c(3), c(4), \ldots, c(x)|.$$

Since the statistical characteristics of one part of the image are the same as other parts, the features learned from one part can also be used in another part, so a convolution kernel can share the same set of parameters in the process of convolution of the input lim (x). To further compare the features, a pooling operation sim $(x\ (i))$ is performed after the convolution. That is, the actual output of each output unit in the neural network is basically consistent with the expected output.

$$\lim_{x \longrightarrow \infty} \left[\frac{1}{n} \sum \operatorname{sim}(x) - \operatorname{sim}(x-1) \right] - \sqrt{\operatorname{sim}(x(i) - x)} = 0, \quad (4)$$
$$F(x, a) \in f(a, n) - 2^{1 - n \sin \theta}.$$

The purpose of learning and training the neural network $F(x, a)$ is to construct the inner connection $f(a, n)$ and law of the musical sound signal through the training sample library, that is, to construct the basic model of recognition. The output layer number of nodes is the same as the number of hidden layer nodes, that is, the number of categories of piano monophony. For the identification of musical tones, we can divide it into t-ends and use the corresponding model. Figure 1 achieves the purpose of music segmentation optimization by calculating its small cumulative prediction residual.

The objects analyzed by the mode key analysis module cover most of the music—Western tonal music and national modes. The input of the module is the output (result) of the feature extraction module of the pitch and duration of the monomelody music, and of course, it can also be a two-dimensional array containing pitch and duration information. The difference in harmonic structure is mainly reflected in the first three-order harmonics, and the energy distribution of the latter harmonics with higher orders is similar, so when paying attention to the first three-order harmonics, the identification score is the highest, and the number of harmonics concerned is the highest.

3.2. Frequency Domain Classification of Piano Playing. First, we need to construct a baseline model for musical instrument recognition as a basis for subsequent model improvements and experimental comparisons. We adapt the proposed neural network model, which has been shown in the literature to be effective for the problem of automatic music labeling. Furthermore, we use training data with frame-level precision labels as supervision signals for training the network, while they train the model in a weakly supervised manner. The ratio s (i) of the sampling frequency to the pitch period is the pitch frequency, and there is a one-to-one correspondence between the pitch frequency and the note.

$$S(i, j) = \begin{cases} s(i) | i = 1, 2, 3 \ldots, n; \\ s(j) | j = 1, 2, 3 \ldots, n, \end{cases} \quad (5)$$
$$\frac{1}{3} (h(\theta, x) - h(x)/i > 0) - \overline{h(x, y) - y(i)^2} = 1.$$

It can be seen that if you want to obtain the frequency value $h\ (x)/i$ of a certain time period i, you need to count the number of all frequency values in this time period and select the largest number as the frequency value of this time period $h\ (x, y)$. The shorter the time period, the more accurate the frequency value of the time period. Doing this every other time period, sim (r) produces a graph of the frequency distribution $c\ (i, j)$ over time. If the amplitude feature $a\ (x)$ is added to the spectrum, a multidimensional spectrum can be obtained.

$$\begin{cases} 0 < \operatorname{sim}(r, c(i-1, j-1)), \\ 0 < \sum \operatorname{sim}(r, c(i, j)) * r(c-s), \end{cases} \tag{6}$$

$$\theta(i, x) - a(x)\frac{\partial a/i}{\partial i} - \overline{i(x, y) - j(i)} = 1.$$

Note pitch frequencies are compressed vertically to 128 levels, and grayscale levels are also 256 levels, where $p\ (r)$ is the expectation factor. The first condition for the above formula to be established in the system is that the process is stable. For the specific training method, we can choose the optimal gradient descent method. The instantaneous value of $b\ (g{-}r)$ at the time gate is used to replace the sum $x\ (r{-}b)$ of squares of errors to solve the defects of the statistical characteristics of the objective function $a + r$ of the whole system at one time.

$$\begin{cases} y(p(r) + 1 - b(g-r)) \geq g, \\ x(r-b) \geq g, \end{cases} \tag{7}$$

$$\Delta\theta = 1 - \frac{\alpha - r}{\partial a + r} - x(r-b) - g.$$

Because of its segmented and derivable advantages, it is applied to practical problems. The nonlinear characteristics enhance the fitting ability $x\ (r{-}b)$ of the function to the data, and the linear characteristics in the range of positive real numbers make the gradient calculation of the model simple and fast during the training process.

The input layer and the hidden layer are fully connected, but the weight is fixed at 1. There is no need to pre-determine the weight between the hidden layer and the output layer, and it can be calculated directly in the next step. According to the scheme, in order to facilitate setting the threshold value and enhance the signal's adaptability to the signal, we first normalize the signal, then divide the data into frames, and obtain its short-term energy and short-term zero-crossing rate frame by frame, and we set two threshold thresholds.

Figure 2 "forces" the convolution by constructing a harmonic sequence matrix to focus on the corresponding number of harmonics in the process of learning intermediate features. The nodes are discarded directly, and only the parameters between the solid line nodes are learned; that is, the state transition matrix of this model is sparse, which reduces the computational complexity of model training. In each round of training, nodes are randomly selected in a certain proportion, and the remaining node information is discarded. Since the nodes removed in each round are not fixed, it is equivalent to training different models with the same data, and the results will be more accurate when integrated, which effectively alleviates the huge amount of parameter calculation introduced by the full connection. However, due to the complexity of the performance situation, the situation of playing wrong is also intricate, for example: playing a few bars of notes, and the subsequent bars are all correct; this situation will be considered by the system: starting from the first bar where the error occurs.

3.3. Pitch Recognition Feature Extraction. In the process of recognizing multiple pieces of piano music, it was found that when the rhythm of the piano music was slow, that is, when the number of notes per second was less than 2, the recognition accuracy of both was very high; when the number of notes per second was in the range of 2~4, the recognition accuracy rate of the traditional method begins to decline, but the improved algorithm can still ensure high accuracy; when the number of notes per second exceeds 5, it means that the interval between each two notes is less than 0.2 seconds, which will lead to harmonics of the previous note. The wave has not decayed enough to interfere with the harmonics of the latter note, so neither method has a high rate of recognition. Under normal circumstances, the number of notes per second in piano music is in the range of 2 to 3, and the number of notes per second over 5 usually only occupies a small section, so the improved algorithm can basically meet the classification requirements.

This model is developed on the basis of the left-to-right model without spanning, which allows the state to jump at intervals, and describes the phenomenon that some pronunciation units in speech are absorbed or deleted in actual speech. For a more complex model, there are two parallel branches, and jumps are also allowed between the branches. Based on the aforementioned research, we choose an extraction algorithm with better performance, MFCC (Meier cepstral coefficient) feature extraction algorithm to process the signal. The specific extraction results are shown in Figure 3.

Through analysis, we select LPCC (linear prediction cepstrum coefficient) and feature parameters extracted by the MFCC algorithm as the input vector of BP network. After that, we normalize the input data, so that some abnormal input can fall within the expected range, and strengthen the adaptability of the system to musical sound data.

According to the different calculation methods of intonation preference features, this study proposes a recognition method for comprehensive evaluation of user features and a multi-category evaluation method for intonation features and makes an experimental comparison between the two methods. The experimental results show that the recognition method of comprehensive evaluation of intonation features is suitable for the recognition of single-class intonation, while the recognition method of multi-class evaluation of intonation features is suitable for the recognition of multi-class intonation. On the whole, the recognition result of comprehensive evaluation of intonation features is better than that of intonation.

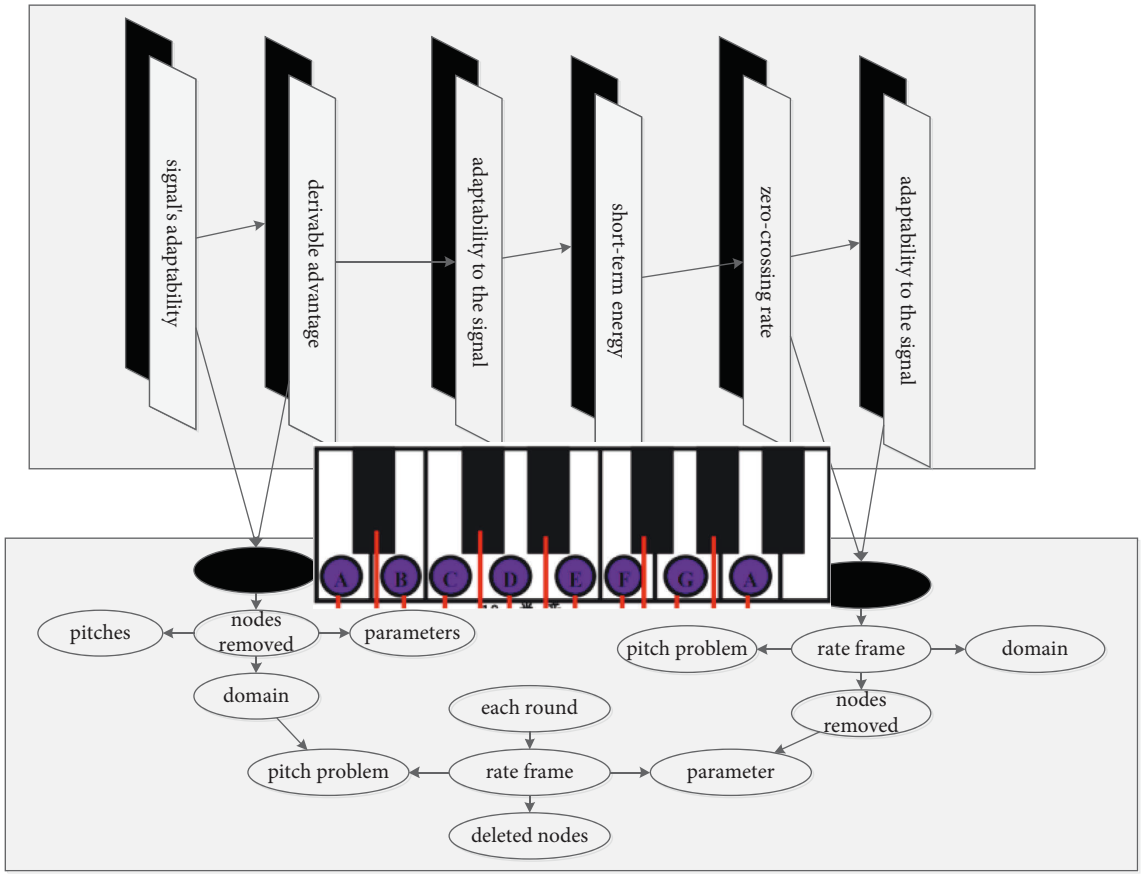


Figure 2: Frequency domain structure of piano playing.

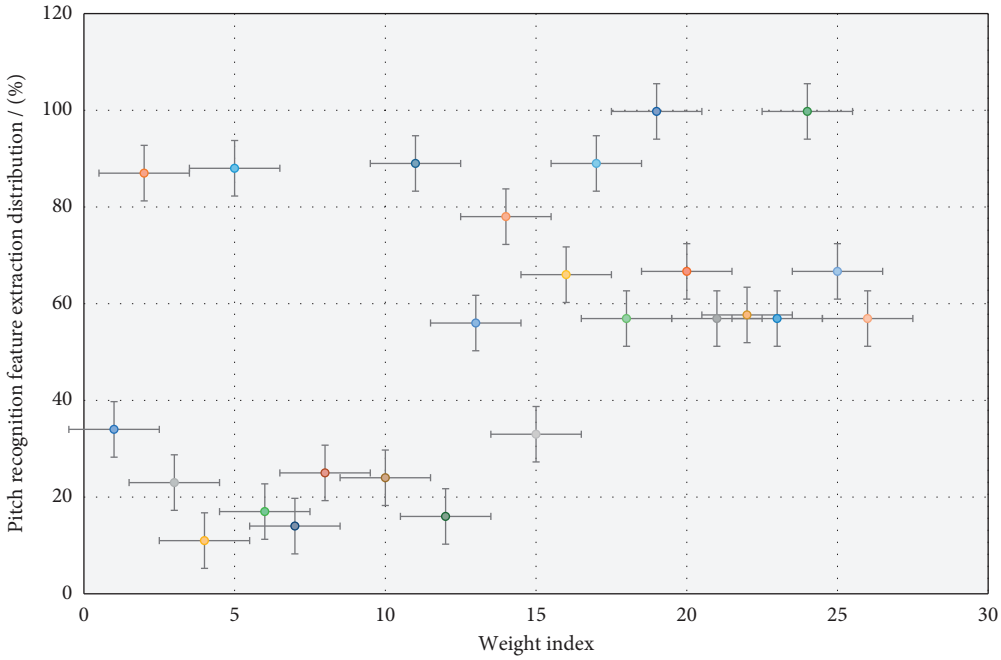


Figure 3: Pitch recognition feature extraction distribution.

3.4. Neural Network Data Convolution. The signal was acquired and saved with the file name "elise. wav." Then, we use the program to read it into MATLAB. In the above-mentioned recognition system using the 88 key tones of the piano as the sample training library, we constructed the signal recognition model through signal preprocessing and feature extraction. On the basis of again, we conducted a comparative experiment on the extraction methods of hidden layer neurons and feature parameters and selected the optimal number of hidden layer neurons and a feature parameter extraction scheme with better performance.

One of the most useful properties of this feature vector is as follows: it can encode the chords contained in a given song. Therefore, two audio frames with similar harmonic content will have the same feature vector. There is a positive correlation between the pitch estimation effect and the instrument recognition effect. It can be seen that after the completion of 200 epochs of AlexNet and its improved network, the spectral recognition accuracy of the verification set music tends to be stable, and the verification set spectral recognition accuracy of AlexNet-improved-v2 has reached the highest. The spectral recognition accuracy of the validation set is 1.08% higher than that of AlexNet-improved-v1 and 1.90% higher than that of AlexNet. Based on the experiments in Figure 4, AlexNet-improved-v2 is more suitable for the music genre recognition model.

Like single-note recognition preprocessing, continuous tones also need to undergo signal de-noising, endpoint detection, and single-note segmentation, but the difference from the single note is that the endpoint detection and signal segmentation of the single note are mainly to further simplify the signal, increase the proportion of useful information, and improve the recognition accuracy of the system; the recognition of continuous musical tones is to extract the continuous signal from the continuous signal. In addition, since timbre and dynamics are relative quantities, it is not convenient to compare, only the corresponding information is extracted, and this factor is not considered in the comparison and results. This function can be improved in future research and expansion. With this feature vector, we can calculate the correlation of the feature vectors of the two audio frames and measure the similarity of the two audio frames accordingly.

3.5. Similarity Error Analysis. The experimental environment is the same as that of the pitch feature extraction experiment. We divided the Bach10 data set, the Mixing Secrets data set, the MedleyDB data set, and the self-built data set, four data sets with annotation labels of frame-level instrument activity levels, into training and test sets in a ratio of 9 : 1, respectively. The audio data for the four categories were segmented to obtain approximately 8000 fragment image samples each. 40% of the image samples are used as training samples, 30% are used as verification samples, and 30% are used as test samples.

For the network optimization of the first-level classification model, Figure 5 uses the algorithm with a momentum of 0.9, a mini-batch size of 100, an initial learning rate of 0.05, and a weight decay factor of 4. For each residual network in the second-level classification model, the network was optimized using the algorithm with a momentum of 0.9, a mini-batch size of 64, and an initial learning rate of 0.1, and it divides the learning rate by 10 every 30 epochs and set the max training epochs to 100 and weight decay to 4. The scene text classification module uses the convolutional neural network technology to identify the business scene of the input text sequence, and the named entity recognition module receives the scene and text sequence from the scene text classification module and uses the NER model based on the composite framework to identify the named entity.

For the needs and convenience of research, many scholars do not consider diminished chords. For each scale, this article defines major chords, minor chords, diminished chords, and three categories of 36 chords according to their names. Experiments found that this set of chord types is appropriate, and it is between over-learning and under-learning. This shows that when the number of windows is the same, the entropy sequence output of the spectrum obtained is basically the same, but when the difference in the number of windows is large, the obtained entropy sequences are slightly different. After experimental verification, when the difference in the number of windows is greater than 30, the obtained entropy sequence has a large difference. Thus, it was concluded that when the number reached 20, the number of neurons increased, although it takes more system training time, the recognition performance cannot be improved again.

4. Application and Analysis of Piano Playing Pitch Recognition Model Based on Neural Network

4.1. Neural Network Data Preprocessing. 100 experimental samples were tested, 82 samples with recognition rate above 9 s%, 1 sample with recognition rate lower than 50%, and 6 samples with recognition rate lower than 70%. Observing the samples in which the R chord is less than 70% of the experimental results, it is found that such experimental samples belong to absolute consonant chords in music, and the two-tone interval is more than 13° apart: that is to say, this module has absolute consonant interval resolution for polyphonic intervals. The presentation layer is mainly the interface for the direct interaction between users and the system. The user can input text sequences through this module and then perform operations on viewing named entities; the logic layer is the business processing part of the entire system, which implements all the logical functions of the system, mainly including scene text.

It can be seen that the musical tone recognition system of Figure 6 has a high musical tone recognition efficiency, which basically meets the requirements for the main frequency of musical tones. However, we also see that due to the longer length of the last note of the elise continuous tone, the proportion of its overtone in the tone signal increases, which has a certain impact on the identification of the fundamental tone, and a certain error occurs in the system.

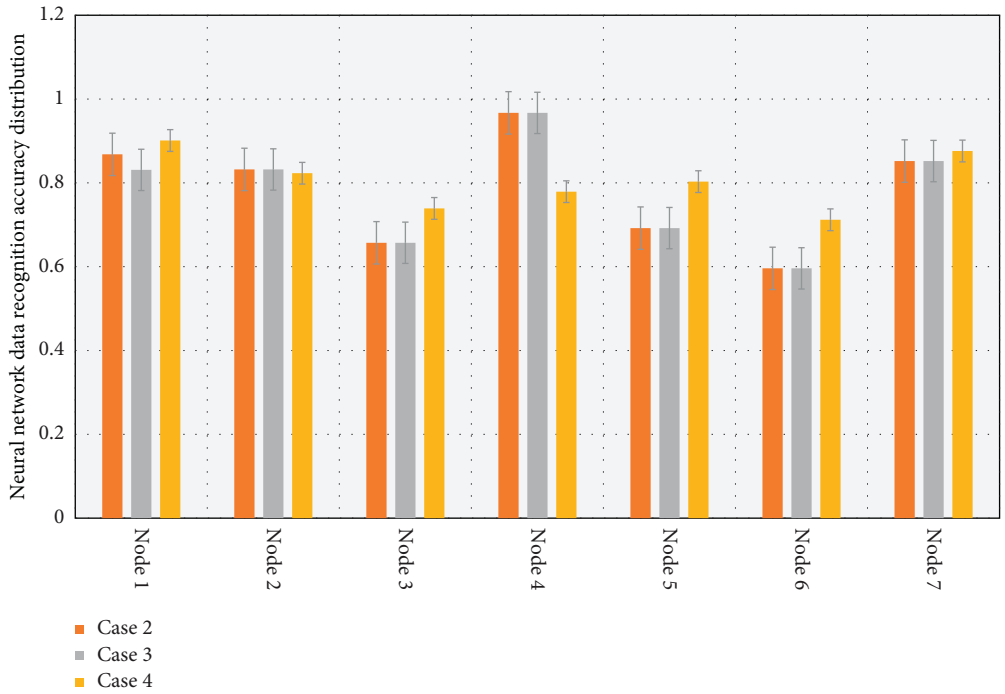


Figure 4: Accuracy of neural network data recognition.

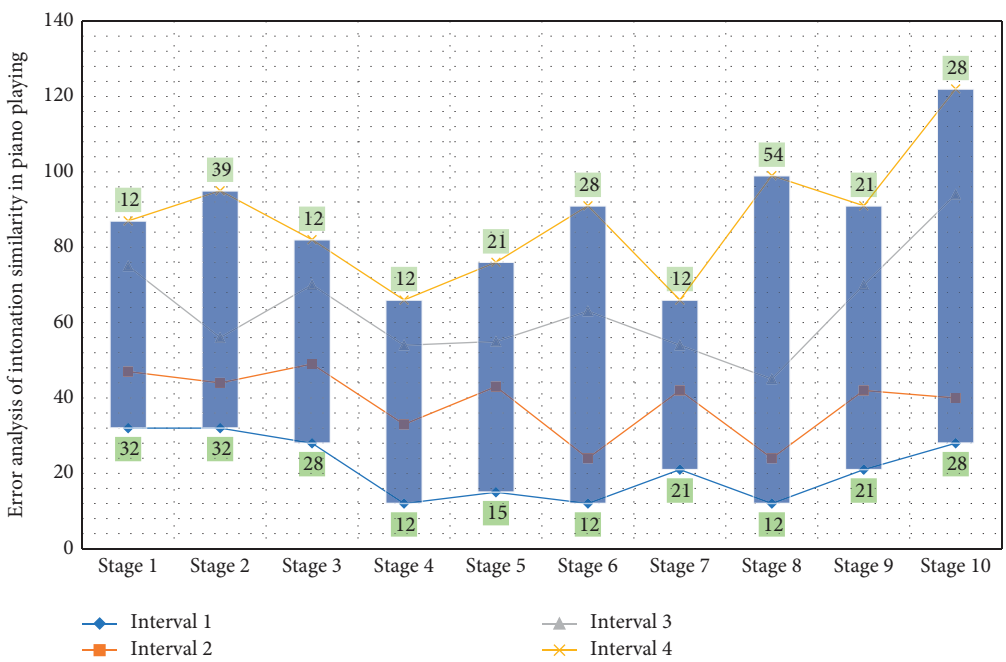


Figure 5: Analysis of intonation similarity error of piano playing.

When the minimum Euclidean distance is greater than 0.45, most of the speech entropy sequences in the template library with the smallest Euclidean distance from the test template entropy sequence as the recognition result are wrong. Therefore, in this experimental system, the threshold for judging the recognition result is set to 0.45. When the minimum Euclidean distance is greater than 0.45, the recognition result cannot be obtained; when the minimum Euclidean distance is less than 0.45, the speech signal represented by the entropy sequence of the minimum Euclidean distance in the template library is returned. From the prediction results, the system fully meets the identification

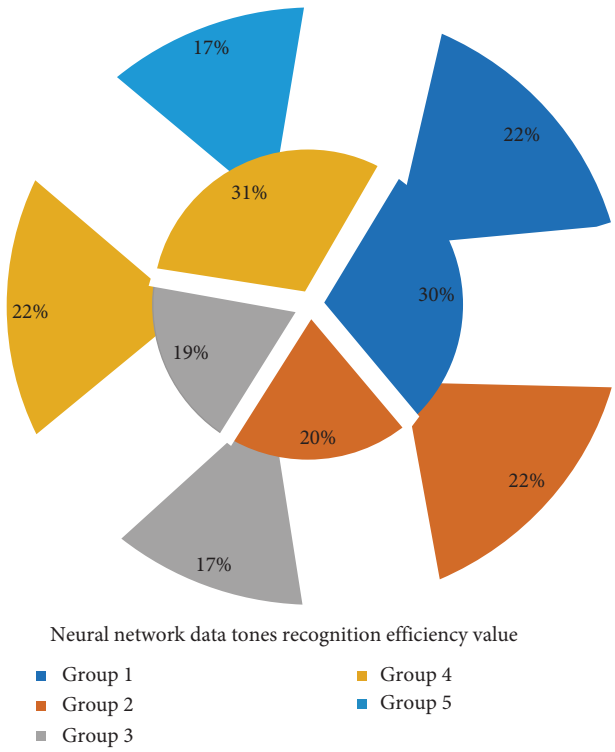


Figure 6: Music tone recognition efficiency of neural network data.

accuracy requirements. After we get the main frequency of the musical note, we can complete the simple score identification of musical notes through the one-to-one correspondence between the musical note name and the main frequency.

4.2. Simulation of Piano Playing Intonation Recognition. The experimental environment of this study is as follows: Ubuntu 16.04 operating system, Intel Core i7-8700 CPU, 16G running memory, Nvidia GTX 1080Ti GPU, and TensorFlow deep learning framework. The model is optimized by the Adam adaptive learning rate algorithm during the training process, which combines the momentum-based algorithm and the adaptive-based algorithm and sets the initial learning rate to 1e-3 and the batch size to 16. In the next iteration, the algorithm evaluates the symbol error rate on the validation set to verify the accuracy of the model.

The system distinguishes and recognizes chords by distinguishing different states. A model is a stochastic finite state machine in which each state produces an observation. The music crawler system can download music on the Internet in batches to build a self-built data set based on the Internet music library to improve the robustness of the music genre identification system.

The network output in Table 1 is converted to the conditional probability distribution on the label sequence. The two-tone pitch duration extraction module accurately reflects the chord duration and pitch information. Since the analysis window is deterministic, if the input is a fast-paced segment, it means that the window will span more notes, and a frame of data is likely to contain more than one type of chord, thus confusing the system. Through this module, the pitch and vibration time of two-tone chords can be analyzed. Through the work of this study, the work of comparing musical scores and audio files has been transformed into a relatively mature string search and comparison work with the help of music feature extraction. It can be expected that if algorithms such as fuzzy search are used for this matching and comparison. More precise conclusions will be made for more performance situations.

First, the C-BiLSTM is trained in the data set in Figure 7, and its effectiveness is verified by comparing the recognition error rates before and after image enhancement; then, the CNN is improved to a residual CNN and its effect on the generalization ability of the model is verified. Secondly, a multi-scale fusion algorithm is added to the residual CNN, and its influence on the feature extraction ability of the model is proved by comparing the feature maps of each convolutional layer of the model and the symbol error rate.

The half-part structure is the top-down feature fusion part, and the pixel-level fusion of the deeper feature map C5 containing semantic information and the upper-level feature map C4 is performed, because the fusion requires the two feature maps to maintain the size and dimension, so the size of C5 is kept consistent with the size of the feature map C4 through 2 times upsampling, and the feature map C4 is passed through a 1×1 convolution kernel to ensure that the feature dimension after upsampling is the same as that of C5, and the feature map F5 is obtained after fusion. The same is done for feature maps F5 and C3, and finally, feature map F4 is obtained. As a result, features containing different levels of information are obtained, to achieve multi-scale feature fusion, so that the feature vector used in subsequent note recognition contains more detailed and comprehensive information.

4.3. Example Application and Analysis. The CNN in the C-BiLSTM network is improved to a residual CNN to form a residual convolutional cyclic neural network, namely RC-BiLSTM. In the same experimental environment, the performance of the model before and after the improved CNN was compared, and the C-BiLSTM network and RC-BiLSTM were trained separately and their symbol error rates were compared. The change in the loss function value of the two networks with the number of iterations during the model training process is shown in the figure, which shows that the loss value of each iteration of the RC-BiLSTM network is lower than that of the C-BiLSTM network, and the loss value of the RC-BiLSTM network has been reduced to 5 and stabilized, while the C-BiLSTM network has only dropped to around 10 with large fluctuations all the time. A collection of neurons is called a layer. On this basis, the outermost layer used for input in the network structure is called the input layer, and the layer used for the final output is called the output layer.

First, we perform feature parameter calculation based on the LPCC (linear prediction cepstral coefficient) algorithm. Using the specific execution steps of the introduced algorithm, considering the accuracy of musical tone recognition

Table 1: Piano pitch recognition algorithm.

Piano pitch network output	Recognition algorithm codes
Pbestmax = popmax;	Obtain the final target
Pbestmin = popmin;	In the next iteration $p(r)$
[newpopmax, newvmax] = updatepop (popmax, vmax);	To verify $\Delta\theta$
[newpopmin, newvmin] = updatepop (popmin, vmin);	The accuracy $sim(x-1)$
V [v < rangespeed [0]] = rangespeed [0]	On the validation set
V [v > rangespeed [1]] = rangespeed [1]	Most likely label $x+y$
For j = 1: popsize	The algorithm evaluates $g(x-x')$
If newvalue_max (j) > value_max (j)	Case of the model
Pbestmax(:, j) = newpopmax (:, j);	By selecting the $\alpha - r$
Gbestvaluemax = newgbestvaluemax;	The symbol error rate
Gbestmax = newgbestmax;	$b(g-r)$
End	For the given input sequence

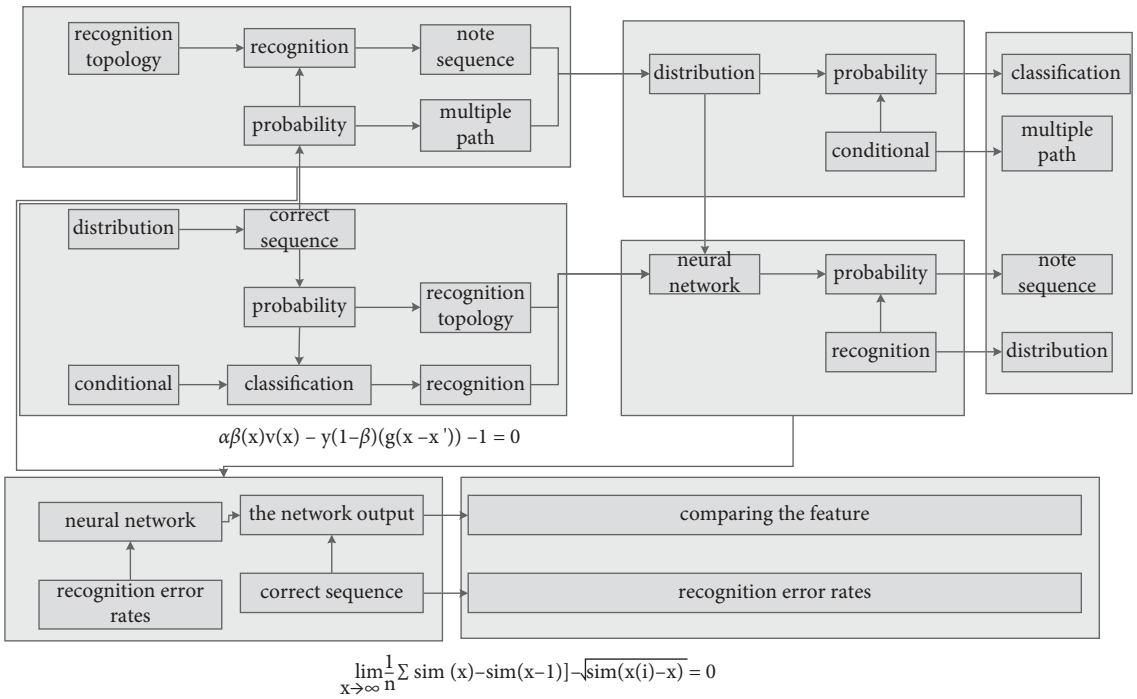


Figure 7: Pitch recognition topology based on neural network.

and the unity of opposites between the amount of calculation, we choose to use the 12th-order LPCC extraction algorithm, the audio signal FFT (fast Fourier transform algorithm) transformation length is set to 256, the sampling frequency is 20500 Hz, and the length of each frame is 256 points. Through initialization and determination of algorithm prediction coefficients, the feature parameter extraction based on the LPCC algorithm is completed, and 6-frame feature parameters are taken in Table 2.

At the same time, the symbol error rates in the two algorithms are compared in the validation set after every 1000 iterations, and the results are shown in the following figure. During the entire training process, the symbol error rate of the C-BiLSTM network dropped to a minimum of about 4%, but at about 5.2×10^4 iterations, the symbol error rate increased significantly, showing that the model did not converge, while the symbol error rate of the network can be stably reduced to less than 2%, and the fluctuation is small. It can be seen that the accuracy of RC-BiLSTM network note recognition has been significantly improved.

This shows that residual CNN can not only improve the accuracy of the model but also solve the problem of model degradation and enhance the generalization ability of the model. For the MFCC feature parameter extraction method, to facilitate the comparative analysis, our parameter settings are the same as the LPCC algorithm and then the DCT (discrete cosine transformation) parameter solution, pre-emphasis filter filtering, and musical sound signal are go troughed. The frames are further divided, and finally, the MFCC parameters of the musical sound signal are calculated. The data set provides the original quartet chorus pieces, and we can generate more recordings with a different polyphony by exploring the different combinations of voices in each piece. The performance dynamics of these new

Table 2: Calculation of musical notation characteristic parameters.

Musical notation index	Feature number	Feature percentage	Root mean square	Error rate
Parameter 1	42	86.46	0.52	0.03
Parameter 2	83	42.91	0.07	0.01
Parameter 3	57	48.95	0.29	0.01
Parameter 4	56	96.70	0.06	0.07
Parameter 5	99	62.58	0.35	0.02
Parameter 6	66	65.67	0.24	0.01

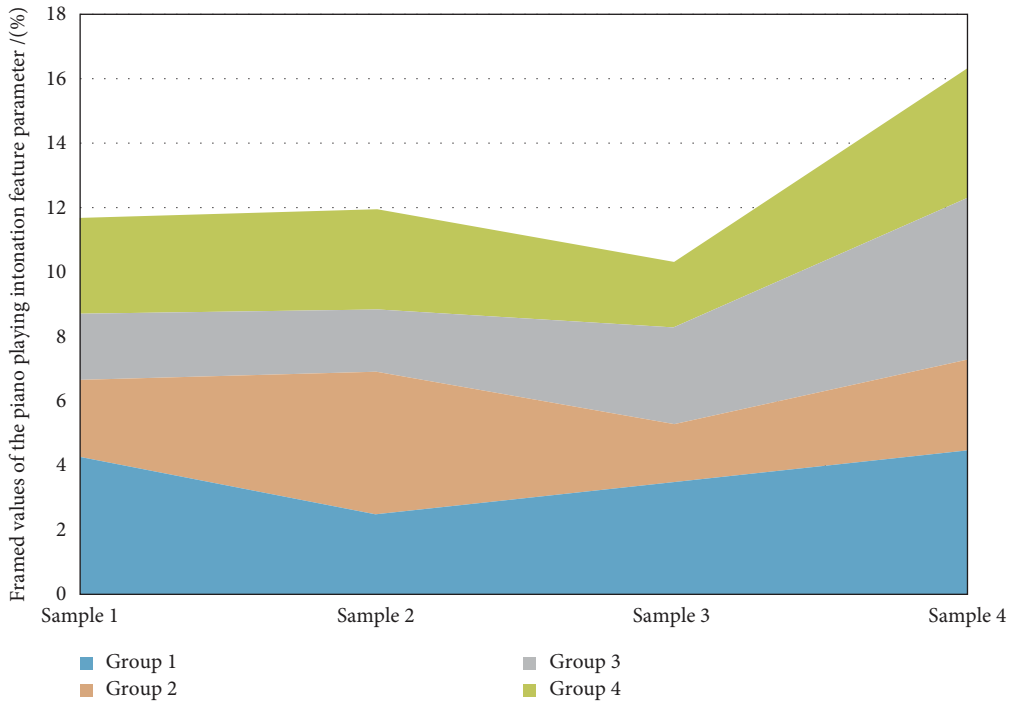


Figure 8: Framing of the intonation characteristic parameters of piano playing.

recordings are the same as the original recordings, providing instrument recognition algorithms with samples of multi-instrumental performance test data at a different polyphony.

The comparison of the recognition results of the two recognition methods for different pitch types is shown in Figure 8. The string matching method is to use strings to represent the melody and then uses string retrieval, quick matching, and other methods to match, by retrieving from the database the characteristic strings of the humming melody. It can be seen that the experimental results and the theoretical values have a good approximation. The experimental data are still derived from the performance of piano players. However, the data used in the module test are the basic two-tone chords in music, and they will not change after the vibration is triggered; that is, they are stable on two tones. To test the performance of this module, 100 two-tone chords with different ranges, different spans, and different degrees of consonance were selected. If the time value feature is further normalized and the length of the sound is specified with a strict musical time value concept, it can be obtained. Here, the normalization method is to take the duration of each note played as the smallest duration unit (such as a quarter beat) as a metric and take its proportional value. Compared with capella-scan, the recognition effect of tuplets and ties is poor. When the algorithm in this study is large, there will be serious errors in simple note recognition, and when the note occupies more pixels, the recognition rate is significantly improved.

5. Conclusion

Based on the neural network, this study constructs a piano playing intonation recognition model and uses the optimized classification result as the classification feature of piano music. The interaction between the user and the system includes two use cases of inputting text sequences and viewing named entities. The viewing named entity use cases include scene text classification and named entity recognition. These four use cases provide their own functions for obtaining named entities of text sequences, so that users can accurately obtain the required data information. First, image enhancement is performed on some score examples in the data set to expand the score image data and improve the robustness of the training model. By analyzing the LPCC feature parameters and MFCC feature parameters, through the analysis and comparison of their principles, the

superiority of the MFCC feature parameters in musical tone recognition is determined. Named entity recognition receives business scenarios from scene text classification and selects the corresponding NER model. The NER model is trained based on the composite framework of transformer network and recurrent neural network. The NER model is used to identify named entities in text sequences and feedback to users. Using the heuristic estimation algorithm when estimating pitch offset, this dynamic programming-based method can meet the requirements of people with general humming level, but the dynamic programming method takes a long time to match large amount of calculation. The results show that the accuracy of the recognition method in this study is more than 50%, which is close to the performance level of similar international methods. The identification method in this study still has a certain gap and needs to be improved.

Data Availability

The data used to support the findings of this study can be obtained from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no known conflicts of financial interests or personal relationships that could have appeared to influence the work reported in this study.

Acknowledgments

This work was supported by a Key Project of Jilin Province Higher Education in 2021: Innovation Research on the Online and Offline Blended Teaching of Piano Course in Colleges and Universities (no. JGJX2021C89).

References

[1] M. Taenzer, S. I. Mimilakis, and J. Abeßer, "Informing piano multi-pitch estimation with inferred local polyphony based on convolutional neural networks," *Electronics*, vol. 10, no. 7, p. 851, 2021.

[2] L. Liu, V. Morfi, and E. Benetos, "Joint multi-pitch detection and score transcription for polyphonic piano music[J]," *ICASSP Acoustics, Speech and Signal Processing (ICASSP). IEEE*, pp. 281–285, 2021.

[3] S. Liu, L. Guo, and G. A. Wiggins, "A Parallel Fusion Approach to Piano Music Transcription Based on Convolutional Neural network[J]," *Acoustics, Speech and Signal Processing (ICASSP). IEEE*, pp. 391–395, 2018.

[4] X. Wang, L. Liu, and Q. Shi, "Harmonic Structure-Based Neural Network Model for Music Pitch detection[J]," *Machine Learning and Applications (ICMLA). IEEE*, pp. 87–92, 2020.

[5] Y. Zhang, "Design of the piano score recommendation image analysis system based on the big data and convolutional neural network[J]," *Computational Intelligence and Neuroscience*, vol. 23, 2021.

[6] A. Ycart and E. Benetos, "Polyphonic Music Sequence Transduction with Meter-Constrained LSTM networks[J]," *Acoustics, Speech and Signal Processing (ICASSP). IEEE*, pp. 386–390, 2018.

[7] R. Kelz, S. Böck, and G. Widmer, "Deep Polyphonic Adsr Piano Note transcription[J]," *ICASSP Acoustics, Speech and Signal Processing (ICASSP). IEEE*, pp. 246–250, 2019.

[8] X. Wang, L. Liu, and Q. Shi, "Exploiting Stereo Sound Channels to Boost Performance of Neural Network-Based Music transcription[J]," *Machine Learning and Applications (ICMLA). IEEE*, pp. 1353–1358, 2019.

[9] A. Ycart, D. Stoller, and E. Benetos, "A Comparative Study of Neural Models for Polyphonic Music Sequence transduction [J]," *Acoustics, Speech and Signal Processing (ICASSP)*, 2019.

[10] T. Wang, "Neural Network-Based Dynamic Segmentation and Weighted Integrated Matching of Cross-Media Piano Performance Audio Recognition and Retrieval Algorithm[J]," *Computational Intelligence and Neuroscience*, 2022.

[11] K. Deng, G. Liu, and Y. Huang, "An efficient approach combined with harmonic and shift invariance for piano music multi-pitch detection[J]," *Pattern Recognition. SPIEL*, vol. 11198, pp. 133–138, 2019.

[12] A. Wiggins and Y. E. Kim, "Guitar tablature estimation with a convolutional neural network," *ISMIR.*, pp. 284–291, 2019.

[13] J. Wang, "The application of intelligent speech recognition technology in the tone correction of college piano teaching [J]," *Journal of Physics: Conference Series*, vol. 1852, no. 2, Article ID 022086, 2021.

[14] P. Steiner, A. Jalalvand, and P. Birkholz, "Improved Acoustic Modeling for Automatic Piano Music Transcription Using Echo State Networks[J]," *Artificial Neural Networks*, pp. 143–154, 2021.

[15] Z. Meng and W. Chen, "Automatic music transcription based on convolutional neural network, constant Q transform and MFCC," *Journal of Physics: Conference Series*, vol. 1651, no. 1, Article ID 012192, 2020.

[16] M. A. Román, A. Pertusa, and J. Calvo-Zaragoza, "An end-to-end framework for audio-to-score music transcription on monophonic excerpts[J]," *ISMIR*, pp. 34–41, 2018.

[17] E. Nakamura, E. Benetos, and K. Yoshii, "Towards Complete Polyphonic Music Transcription: Integrating Multi-Pitch Detection and Rhythm quantization[J]," *Acoustics, Speech and Signal Processing (ICASSP). IEEE*, pp. 101–105, 2018.

[18] L. Watts, "DeepPitch: Wide-Range Monophonic Pitch Estimation Using Deep Convolutional Neural networks[J]," *Interspeech*, 2018.

[19] H. Qian, P. Gu, and R. Yan, "Robust Multipitch Estimation of Piano Sounds Using Deep Spiking Neural Networks[J]," *Symposium Series on Computational Intelligence (SSCI). IEEE*, pp. 2335–2341, 2019.

[20] X. Wang, "Research on the improved method of fundamental frequency extraction for music automatic recognition of piano music," *Journal of Intelligent and Fuzzy Systems*, vol. 35, no. 3, pp. 2777–2783, 2018.

[21] S. Gururani, C. Summers, and A. Lerch, "Instrument activity detection in polyphonic music using deep neural networks," *ISMIR.*, pp. 569–576, 2018.

[22] Q. Kong, B. Li, X. Song, Y. Wan, and Y Wang, "High-resolution piano transcription with pedals by regressing onset and offset times," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 29, pp. 3707–3717, 2021.

Hindawi
Computational Intelligence and Neuroscience
Volume 2022, Article ID 1823762, 12 pages
https://doi.org/10.1155/2022/1823762

Research Article

Intelligent Logistics System Design and Supply Chain Management under Edge Computing and Internet of Things

Tianxia Wang,[1] Hong Chen,[2] Rui Dai,[3] and Delong Zhu[4]

[1]*School of Economics and Trade Management, Yibin Vocational and Technical College, Yibin 644003, China*
[2]*School of Management, Guangzhou Xinhua University, Dongguan 523133, China*
[3]*College of Finance and Statistics, Hunan University, Changsha, China*
[4]*School of Economics and Management, Hubei University of Automotive Technology, Shiyan 442002, China*

Correspondence should be addressed to Delong Zhu; 201431050134@chd.edu.cn

Received 11 July 2022; Revised 18 August 2022; Accepted 2 September 2022; Published 16 September 2022

Academic Editor: Ning Cao

In order to carry out practical innovation of the intelligent logistics system and promote the practicality of the intelligent logistics system and supply chain management process, this study aims to optimize the design of the intelligent logistics system and supply chain management under edge computing (EC) and the Internet of Things (IoT). The flower pollination algorithm performs the positioning function in the intelligent logistics system and supply chain management. Based on the research on the design of the intelligent logistics system and supply chain management under the EC and IoT, this thesis analyzes the positioning of intelligent logistics systems and supply chain management through the flower pollination algorithm. The eXtreme Gradient Boosting (XGBoost) model is used to predict user information in the system of supply chain management information. Finally, the operation of intelligent logistics and supply chain management systems, the prediction model of supply chain management under XGBoost, and the change of supply chain management and material flow are analyzed. The results show that with the increase in the number of iterations, the optimized algorithm improves the comparison distance error by 53.57%, which has high accuracy and can meet the requirements of positioning and tracking of the intelligent logistics system and logistics status query in supply chain management. The waiting time of the intelligent logistics system is shorter than that of the supply chain management system, and the average waiting time of the system increases by 121.252 ms. The XGBoost model can well predict user information under supply chain management. After discussing the changes of the intelligent logistics system from 2018 to 2020, it is found that the operation efficiency of the supply management system is higher with the increase of the system operation days. The intelligent logistics system has a significant impact on the development of the logistics industry. This research gives a reference for establishing the intelligent logistics system and supply chain management system.

1. Introduction

Logistics is the basis of the development of the Internet of Things (IoT). As an economic activity, it appears with the emergence of commodities and develops with the development of commodity production [1]. The development of the IoT is dependent on the support of logistics. It can be said that logistics is one of the earliest industries in the IoT. Many logistics systems involve some advanced technologies, such as informatization, digitization, networking, integration, intelligence, flexibility, agility, visualization, and automation. Originally, the IoT is called the sensor network and used in the logistics industry. For example, radio frequency identification (RFID) is one of the most basic applications of the IoT [2–4]. In traditional storage, manual scanning is needed to obtain data and the efficiency is low. The storage location is not defined clearly, which causes the phenomenon of the random stack of goods. After the IoT is applied to the traditional warehousing management system, an intelligent warehousing management system is obtained. It can improve the efficiency of scoring and delivering goods, expand the storage capacity, reduce the labor intensity and cost, monitor the goods delivery process, and complete the data query [5–8]. Through the systematic management of

logistics vehicles, the freight cars and goods can be tracked and monitored in real time, and the temperature and humidity of goods can also be detected. In the process of transporting goods, the vehicle speed, fuel consumption, tire temperature, and other driving behaviors are monitored in real time. In this way, the transportation efficiency is improved, and the transportation cost and loss are reduced [9–11].

With the combination of Internet technology and the logistics industry, the demand for intelligent logistics markets is increasing. China's modern logistics and supply chain management are still in their infancy. The suppliers only provide goods and have no value-added services [12], and the income of logistic companies and outsourcing service providers mainly comes from basic services (transportation management and warehouse management) and value-added services (supply chain integration services, supply chain financial services, and supply chain platform construction) [13, 14]. Google Cloud IoT is a platform designed for intelligent services, likefully hosted integrated services, which can be used to easily and safely connect to a large number of devices distributed around the world, manage and extract the data on these devices, visualize these data in real time, improve business agility and speed up decision-making, and make business changes [15, 16]. Edge computing (EC) can transform the pattern of the manufacturing industry from the traditional centralized control mode to the mode of real-time collection and processing of network edge data [17]. Since 2017, Google has announced two related new products, namely hardware Chip TPU and software Cloud IoT Edge, which helps promote the development of edge networking devices. Google said, "cloud IoT edge can extend the powerful data processing and machine learning capabilities of Google cloud to billions of edge devices, such as robot arms, wind turbines, and oil rigs, so that they can operate the data from their sensors in real time and predict the results locally [18–20]." The standardization of EC gradually attracts the attention of major organizations and relevant working groups have been established to do research on it [21]. Regarding the importance of studying the intelligent logistics system, Çakmak et al. pointed out in the optimization study of the express logistics system that the current logistics system played a great role in improving the efficiency of transportation logistics symbolized by express delivery [22]. Yan and Li illustrated the application of Radio Frequency Identification (RFID) technology in logistics information systems, and demonstrated the application of technology in logistics information systems [23]. Da Silva et al. proposed that logistics system and supply chain management interact and relate each other, which is the technical basis of supply chain management process. It has transparency and security of information in rapid response and model construction [24]. Regarding the technological development of intelligent logistics system, Feng and Ye realized intelligent transportation and warehousing operation process by optimizing the algorithm architecture of logistics information system in intelligent logistics system [25]. Li et al. proposed a control method and system under an intelligent logistics system [26].

Based on the research on the design of the intelligent logistics system and supply chain management under the EC and IoT, this thesis analyzes the positioning of intelligent logistics systems and supply chain management through the flower pollination algorithm. The eXtreme Gradient Boosting (XGBoost) model is used to predict user information in the system of supply chain management information. Finally, the operation of intelligent logistics and supply chain management systems, the prediction model of supply chain management under XGBoost, and the change of supply chain management and material flow are analyzed. This thesis has certain reference significance for the establishment of intelligent logistics systems and supply chain management. The innovation of this study lies in using the flower pollination algorithm to carry out the positioning work in the process of intelligent logistics management and using the IoT and EC to design the intelligent logistics system and supply chain management process.

This research is divided into four sections. Section 1 describes the development of the modern intelligent logistics system, explains the research of logistics in China and foreign countries, and explains the research method and research frame. Section 2 is the theory and research method. Firstly, the edge computing theory is introduced. Secondly, the intelligent logistics system and supply chain management process are designed combined with the IoT and EC. Finally, the positioning in the process of intelligent logistics management is carried out through the flower pollination algorithm. Section 3 mainly analyzes the operation of intelligent logistics system and supply chain management system, and the changes of supply chain management and logistics flow. Section 4 mainly summarizes the results and puts forward the specific prospect of future research and development.

2. Method

2.1. Theory of Edge Computing. EC is a distributed architecture that moves the operations of applications, data, and services from the central node of the network to the edge [27]. It decomposes the large-scale services initially processed by the central node into smaller parts and distributes them to the edge nodes for processing. The edge node is closer to the user terminal device, which can speed up the data processing and transmission speed. Under this framework, the data can be analyzed closer to the data source, so the framework can process a large number of data. EC is a decentralized computing architecture. It can store the data where they are needed and make most of the calculations done in distributed nodes. "Edge" refers to any calculations and network resources between the data source and the cloud data center. For example, smart phones are the "edge" between individuals and the cloud, and gateways in smart homes are the "edge" between home devices and the cloud. The basic principle of EC is to calculate near the data source. It is an open platform that integrates the core capabilities of networks, calculations, storage, and application and provides edge intelligent services nearby the edge of the

network near the object or data sources. The goal is to make any application or function closer to the execution device.

EC is a supplement and optimization of cloud computing. It is also an important supporting technology of the 5th generation mobile networks (5G) and IoT [28, 29]. Cloud computing is to calculate big data in the cloud computing center (Data Center), which can realize on-demand access. However, it cannot meet the needs of real-time data processing and the explosive growth of networking equipment in the era of 5G and IoT. As a supplement to cloud computing, EC can be used in local decision-making and other scenarios. Table 1 shows the collaboration points between EC and cloud computing.

Table 1 shows the collaborative points of EC and cloud computing. The edge collects, stores, and preprocesses real-time data. After most redundant and unimportant data are excluded, only the cleaned data are sent to the cloud, reducing the network's pressure. It can store the data when the network signal is poor, and upload the stored data to the cloud until the network is restored. The cloud only needs to process and learn the massive data collected by each edge, and push the updated prediction model to the edge, which greatly relieves the workload of the cloud. Compared with cloud computing, EC is arranged nearby and can realize the unprecedented connectivity, centralization, and intelligence of IoT, which can meet the needs of agile connection, real-time business, data optimization, application intelligence, security, and privacy protection. It is an important support for realizing distributed autonomy and industrial control automation. Figure 1 shows the process of EC under IoT.

Figure 1 shows the information integration solution provided by the EC under the IoT. It includes the servers, algorithms, applications, and equipment access capabilities needed to carry out edge services. The IoT edge computing provides cloud management and local management for all-in-one machines. The EC management platform under the IoT is a management console running in the hardware of an all-in-one machine. It can manage the offline all-in-one machine's network configuration, algorithm tasks, and terminal equipment. The one-click deployment function of the EC console can synchronize the resources and configurations between the cloud and the edge and view the deployment progress and trace the deployment log. The EC preloads the official application software in the all-in-one machine under the IoT and focuses on developing business logic without consuming energy, such as program startup, message flow, log query, and process maintenance. The EC under the IOT also supports the development of container applications or light applications to expand businesses. It also provides a multilanguage software development kit (SDK) to support common IoT devices and systems, including sensor devices, NVR (network video recorder), business systems, and other services connected to the cloud.

2.2. Intelligent Logistics System Based on IoT. The application of IoT in logistics system can improve automation, optimize transportation management process, improve the transmission of traffic information, and improve traffic efficiency.

Table 1: Collaborative points of EC and cloud computing.

Coordination points	EC	Cloud computing
Networks	Data aggregation	Data analysis
Businesses	Agent	Business arrangement
Application	Microapplication	Lifecycle management
Intelligence	Distributed reasoning	Centralized training

It can also reduce transport risk, improve resource utilization, reduce transport costs, and realize the automation, visualization, controllability, intelligence, and networking of logistics transportation. It enables all parties in transportation to share traffic information, and improves the integration ability of transportation enterprises and the ability to perceive and respond to market changes. It can timely and accurately provide customers with relevant product transportation information, and formulate optimal transportation plans for customers. It can also provide customers with the most satisfactory services and improve the overall logistics and transportation service level and service quality. The application of the IoT in supply chain management can enable supply chain system managers to accurately track and locate any item in any link of the supply chain, and achieve transparent supply chain management. Applying the IoT in supply chain management has achieved a high degree of integration of all aspects of the supply chain and improved the overall management efficiency of the supply chain.

Intelligent logistics depends on the application of the Internet and IoT. It uses advanced technologies to collect, process, manage, circulate, and analyze data. Figure 2 shows the design of the intelligent logistics system.

Figure 2 shows the process of intelligently completing the packaging, transportation, distribution, loading, unloading, and warehousing. In this way, the flow status of goods is monitored in real time, so that the goods can be delivered to the demander efficiently and quickly, reducing the cost for the supplier. It can also significantly reduce the consumption of social and natural resources. Figure 3 shows the overall framework of the intelligent logistics system.

In Figure 3, based on the business process of the intelligent logistics system, the overall framework of the informatization of the intelligent logistics system is clarified, and on this basis, the informatization function requirements of each link are proposed. It mainly includes inbound logistics, warehouse management, material distribution, finished product logistics, container management, emergency logistics, and other links. Meanwhile, it also clarifies the input, information drive, information collection, output, and other information of each process activity node. Table 2 shows the functions of the intelligent logistics system.

According to the functions of the intelligent logistics system in Table 2, the design of the intelligent logistics system should meet the personalized and diversified needs of users. Figure 4 shows the framework of the intelligent logistics system.

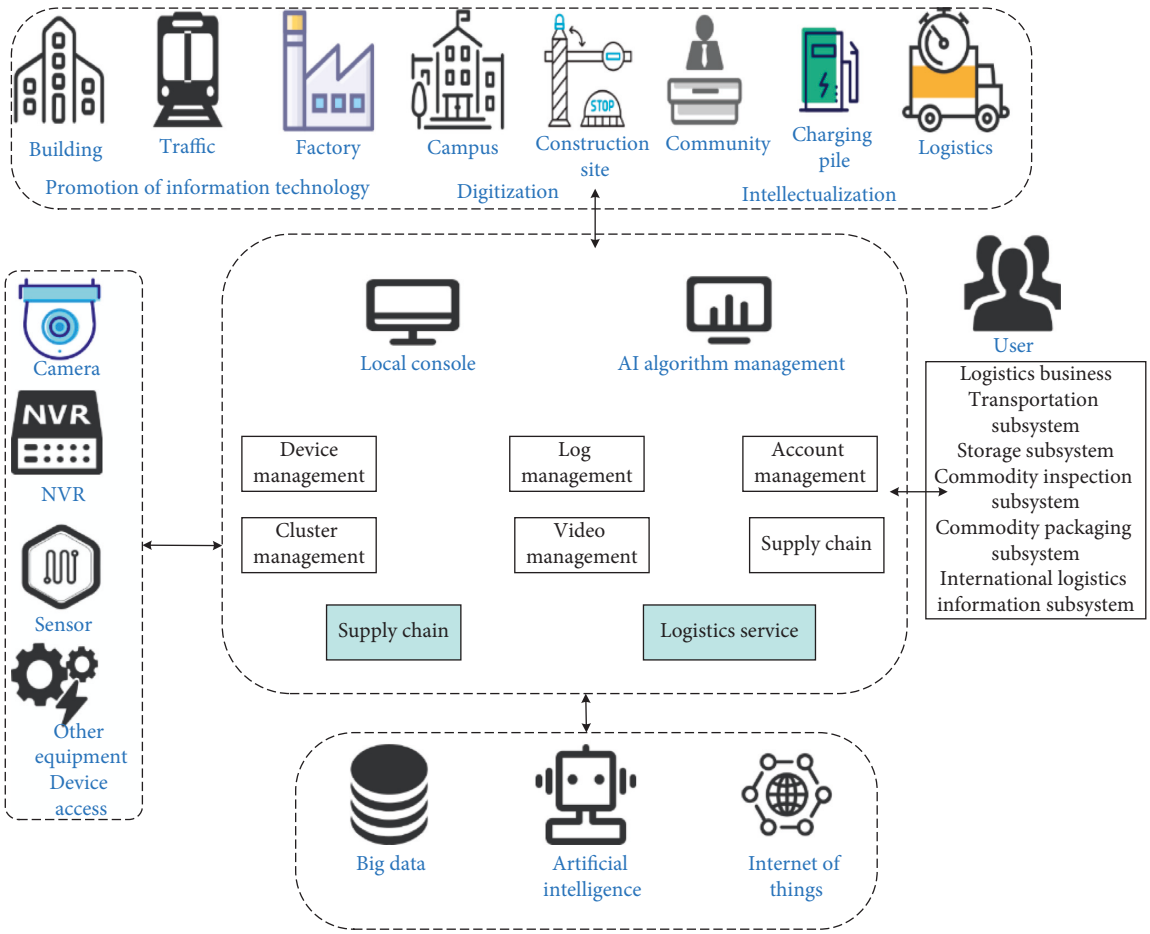


Figure 1: EC under IoT.

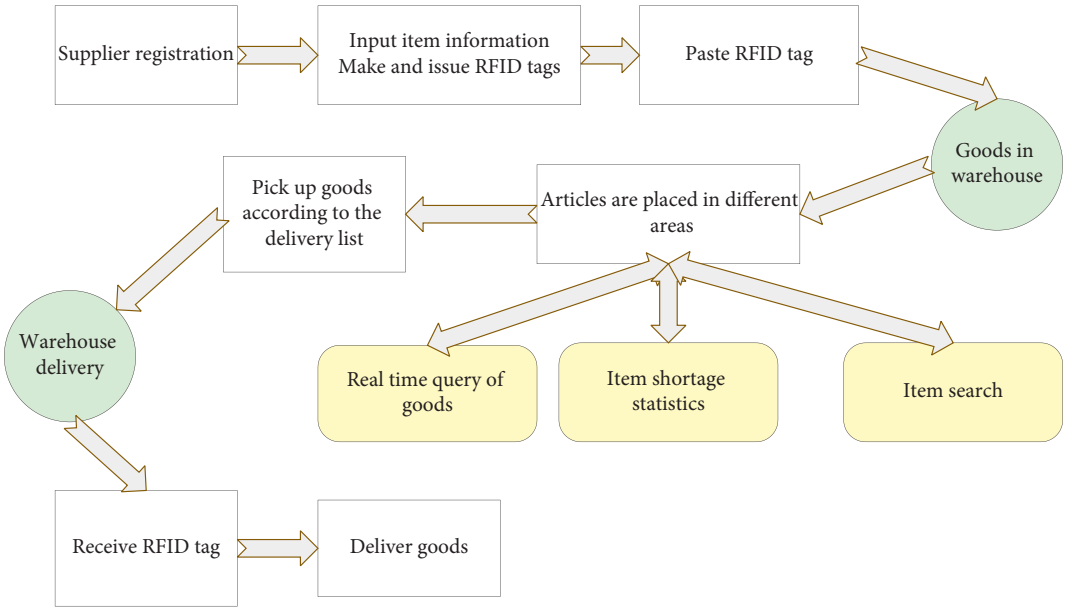


Figure 2: Design of the intelligent logistics system.

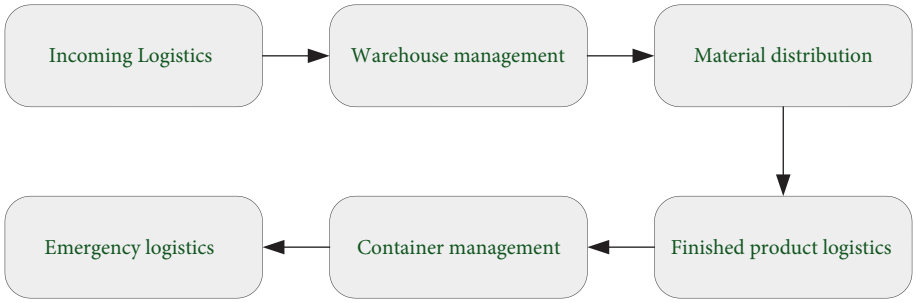


Figure 3: Overall framework of the intelligent logistics system.

Table 2: Functions of the intelligent logistics system.

Realization elements	Content	Functions
Logistics function integration	Integrating logistics services and value-added services	Meeting the personalized needs of users
Real-time tracking of users	Understanding the changes in the user's needs	Improving service levels
Information integration	Other logistics information	Providing integrated and networked services
Financial integration	Financial supervision and management of logistics businesses	Achieving financial objectives

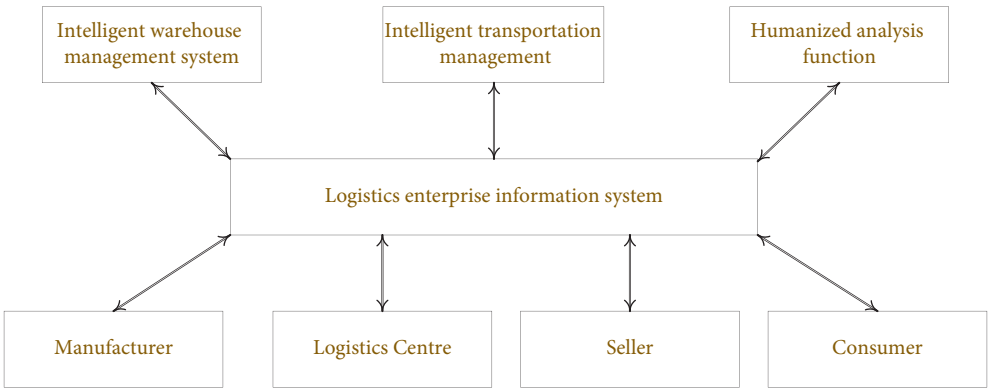


Figure 4: Framework of the intelligent logistics system.

Figure 4 shows that the intelligent logistics system includes the advanced traffic information service system, vehicle control system, traffic management system, operating truck management system, and electronic toll collection system. ITS can realize logistics distribution management and centralized dynamic control of vehicles, including providing road traffic information, route guidance information to optimize the decision-making transportation scheme, accurate arrival time of vehicles through real-time tracking of vehicles, the warehouse inventory strategy, and distribution plan of the logistics center. This can realize the information sharing between each node of the logistics network and the headquarters and between network nodes, and improve the efficiency of the whole logistics transportation system. The IoT is a huge network that combines the global positioning system (GPS), infrared sensors, radio frequency identification (RFID) devices, laser scanners, and other devices with the Internet. Through this network, any object can be connected with the Internet for information exchange and communication, realizing intelligent identification, positioning, tracking, monitoring, and management.

2.3. Design of the Intelligent Logistics Supply Chain. Supply chain logistics management is the logistics management system centered on the core products or core businesses of the supply chain. The former focuses on the logistics management of the supply chain organized with the manufacturing, distribution, and raw material supply of core products as the system, such as the logistics management of automobile manufacturing, distribution, and raw material supply chain, which is the logistics management system centered on automobile products. The latter is the logistics management system of the supply chain organized with the core logistics business as the system, such as the logistics management of the third-party logistics, distribution, storage, or transportation supply chain. The logistics management of these two types of the supply chain has both similarities and differences. The core of supply chain

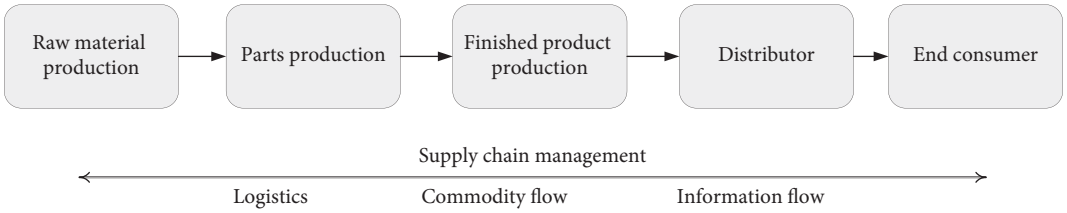


Figure 5: Process of logistics supply chain management.

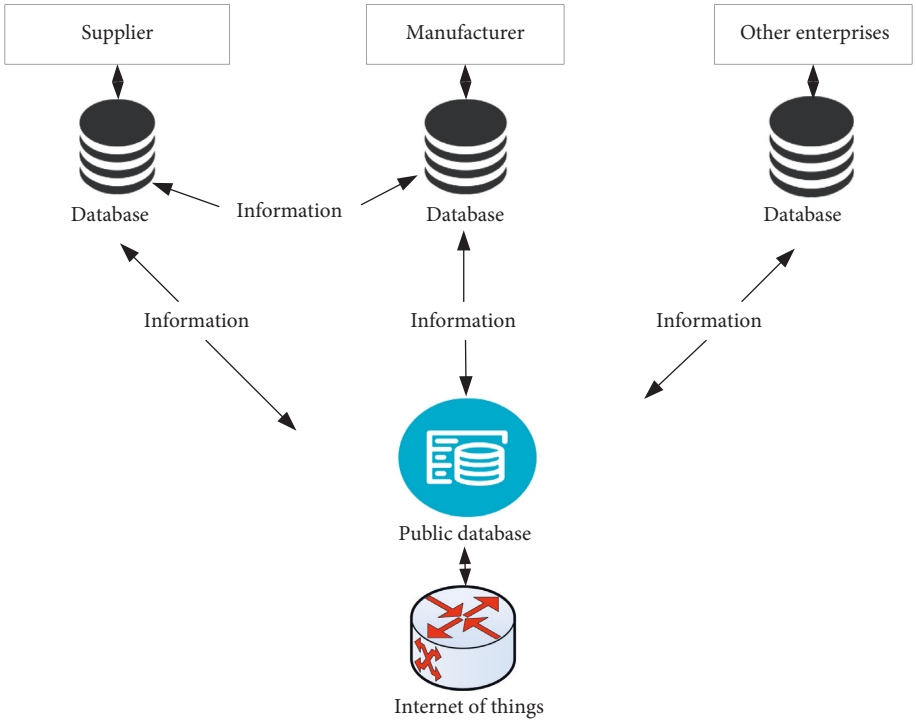


Figure 6: Information interaction mode of supply chain logistics management.

management is the logistics management of the supply chain. Capital flows serve to create conditions to ensure the smooth progress of logistics. Figure 5 shows the process of logistics supply chain management.

Figure 5 presents the logistics supply chain management business process, which is different from traditional logistics focusing on warehousing and distribution. The supply chain management covers the entire logistics link from raw material procurement to final distribution to end consumers and realizes the unification of logistics, information flow, and capital flow. The product manufacturing process provides a framework for how to develop new products and markets with suppliers. This process not only enables management to coordinate the flow of new products in the supply chain but also improves operations such as manufacturing, logistics, and marketing. Part production management process is a supply chain process, including all the necessary activities in the supply chain to obtain, implement, and manage the manufacturing flexibility, as well as the activities needed to move the product out of the factory. Distribution management processes based on customer value breakdown increase customer trust through customized product and service agreements. The supplier relationship management process formulates the healthy relationship between enterprises and suppliers. The process provides an important link in the supply chain among enterprises. Figure 6 shows the information interaction mode of supply chain logistics management.

Figure 6 shows that with the development of concurrent engineering, there are higher requirements for information management. The technology should support the collaborative work of multidisciplinary expert groups and realize the organic combination of interactive information, so that the correct information can be transmitted in real time. Therefore, in the actual process of information interaction based on SCLM, enterprises do not use a single mode for information interaction. They can select from the above three modes according to their needs and the confidentiality of the interactive information. On the basis of the first three interaction modes, the comprehensive information interaction mode integrates the interactive information to realize information integration and improve the efficiency of information interaction. Figure 7 shows the operation process of the intelligent logistics management system in the supply chain.

Figure 7 shows that the intelligent logistics management system on the supply chain takes enterprise resource

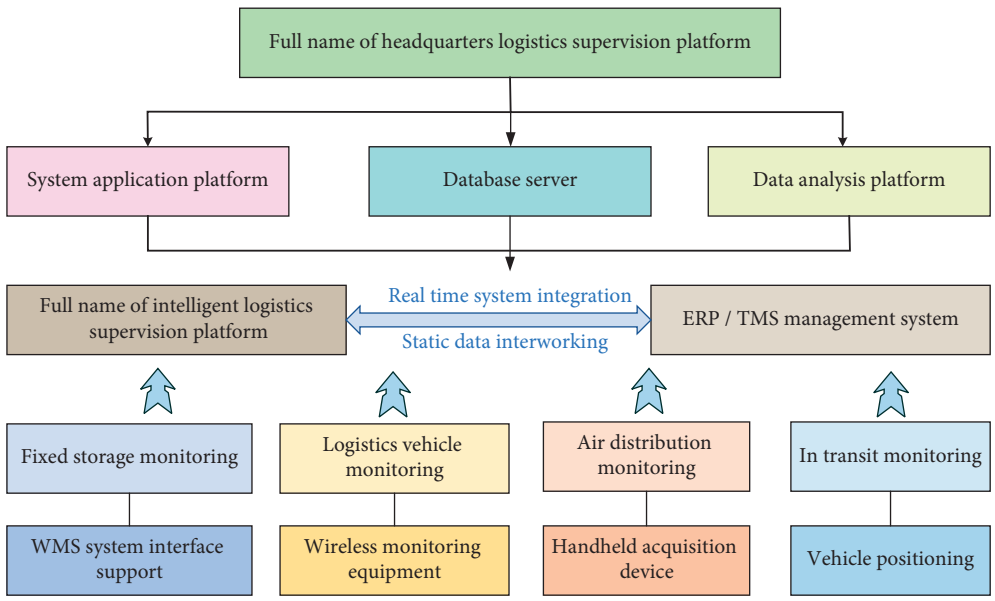


Figure 7: Operation process of the intelligent logistics management system in the supply chain.

planning (ERP) and transportation management system (TMS) as the technical and theoretical support and traces the flow of goods on the whole supervision platform of the logistics head office. In addition, the platform reserves a data interface to facilitate docking with the customer system and has the functions of data analysis and report exports. The platform integrates human, material, information, and other resources, optimizes business processes, improves operation efficiency, connects the warehouse management system (WMS), and big data analysis platform of logistics enterprises, and realizes data and information sharing. The vehicle, equipment, and personnel information used in each link, as well as the temperature of the intelligent equipment in picking up, warehousing, transferring, and signing in, can also be displayed. The flower pollination algorithm is used to locate the process of intelligent logistics management. The relevant parameters are initialized, including flower population number n and conversion probability p [30, 31]. The pollen position is updated by the following equation:

$$X_i^{t+1} = X_i^t + L\left(g^* - X_i^t\right). \tag{1}$$

In equation (1), X_i^{t+1} and X_i^t are the solutions of $t+1$ and t, respectively, g^* is the optimal global solution, and L is the step length. L is calculated by the following equation:

$$L \sim \frac{\lambda\Gamma(\lambda)\sin(\pi\lambda/2)}{\pi s^{1+\lambda}}, \left(s \gg s_0 > 0\right), \tag{2}$$

$$\Gamma(\lambda) = \int_0^{+\infty} x^{\lambda-1} e^{-x} \mathrm{d}x, \left(\lambda > 0\right). \tag{3}$$

In equation (2), $\Gamma(\lambda)$ is a standard gamma function. When $\lambda = 3/2$, the pollen is renewed and the boundary is crossed by

$$X_i^{t+1} = X_i^t + \epsilon\left(X_j^t - X_k^t\right). \tag{4}$$

In (4), ϵ is a random number between 0 and 1, X_j^t and X_k^t are the pollen of different flowers of the same plant. Conversion probability p is a constant, and the adjustment of P is made by

$$p = 0.8 + 0.2\mathrm{rand}. \tag{5}$$

In (5), rand is a random number between 0 and 1, and step factor γ of L is calculated by

$$\gamma = \frac{N-t}{N}. \tag{6}$$

In (6), N is the maximum times of iterations, and t is the current times of iterations. Global search is conducted by (7). The random number of rand is modified during the local search to improve the flower pollination algorithm. The specific changes are made using

$$X_i^{t+1} = X_i^t + \gamma L\left(g^* - X_i^t\right), \tag{7}$$

$$\varepsilon_{t+1} = \begin{cases} \varepsilon_t + \mathrm{ran\,d} * \varepsilon_u, & \mathrm{rand} < 0.1, \\ \varepsilon_t, & \mathrm{rand} > 0.1. \end{cases} \tag{8}$$

2.4. Prediction Model of Logistics Supply Chain Based on XGBoost. More and more enterprises use data analysis to deal with supply chain interruption and strengthen supply chain management. For example, the global COVID-19 pandemic highlights the strategic importance of an integrated supply chain as a response to disruption. The supply chain must be resilient and able to withstand shocks and accidents. Therefore, in the design of an intelligent logistics system, it is necessary to predict the important supply chain demand. Table 3 is a machine learning algorithm for predicting the supply chain.

According to the XGBoost model in Figure 3, the prediction model in supply chain management is built.

Table 3: Machine learning algorithm for forecasting supply chain.

Machine learning algorithm	Application	Algorithm advantages
Decision tree	Simple regression and classification	Low time complexity and high efficiency
XGBoost	Practical problems, data competition	Fast algorithm implementation speed and high operation processing speed
Random forest	Regression and classification tasks	Strong fitting ability and fast operation speed
Support vector machine	Data analysis and pattern recognition tasks	High fitting degree and high recognition rate

XGBoost is obtained based on the improvement of gradient boosting decision tree (GBDT). The XGBoost is also an additive model. If the given dataset has n samples and m features, the dataset can be expressed as $D = \{(x_i, y_i)\}$. The predicted values of K functions are added to fit the model. (9) is the specific predicted value:

$$\hat{y} = \sum_{k=1}^{K} f_k(x_i), f_k \in F. \tag{9}$$

In (9), $F = \{f(x) = w_{q(x)}\}(q\colon R^m \longrightarrow T, w \in R^T)$ is the function space constructed by all regression tree models. In terms of specific parameters, $q(x)$ is a function that maps the eigenvector x to the leaf node index of the decision tree. T is the number of leaf nodes corresponding to a decision tree. w is the weight corresponding to the leaf node, so each tree f_k corresponds to a tree structure feature vector q and the weight vector w corresponding to the leaf node.

First, the regularization term is added to the objective function. The objective function equations in the algorithm are

$$L(\varphi) = \sum_i l(\hat{y}, y_i) + \sum_i \Omega(f_k), \tag{10}$$

$$\Omega(f) = \gamma T + 0.5\lambda w^2. \tag{11}$$

$w = (w_1, w_2, \cdots, w_k)$. $l(\hat{y}, y_i)$ is the difference between the predicted value and the real value. Ω is a regularization term, which uses the L2 norm of the number of leaf nodes and the corresponding value vector of leaf nodes to control the complexity of the trained tree model, and λ is the corresponding coefficient. The system inventory, diversion route, user, product, and user demand of the supply chain management's logistics system are used as input variables. (12) is the construction of characteristic variables:

$$v(t) = \frac{\sum_{j=1}^{n} m_{t-1}}{n}. \tag{12}$$

In (12), $v(t)$ is the created second manual input variable, n is the statistics of the number of times the user purchases a certain product, and j is the same user purchased such products for the j time. m_{t-1} refers to the last purchase of such products by the user. (13) is the use frequency of characteristic variables:

$$w(t) = \sum_{o=1}^{p} ir. \tag{13}$$

In (13), $w(t)$ is the use frequency of characteristic variable user information in week t. p is the total usage. i is the usage of characteristic variable user information in week t. r is the usage of all user information. The initial value of the maximum depth of the decision tree in the model's use is 10. Besides, the mean square error is used as the evaluation standard of the model. Then, the number of decision trees in the random forest is 200. When splitting, the minimum sample initialization value of the splitting node is 2, and the minimum sample initialization number of the cotyledon node is 1. The minimum initialization value of the sum of the weights of all the cotyledon nodes is 0. Moreover, the feature quantity in the splitting is automatic splitting. The maximum characteristic number index reads

$$m = \log_2 N. \tag{14}$$

m is the maximum number of features, and N is the total number of features of the input variable. The mean absolute error (MAE), root mean squared error (RMSE), and average absolute percentage error are used to evaluate the prediction performance of this model in logistics supply chain management. (15) displays the MAE:

$$\text{MAE} = \frac{1}{m}\sum_{i=1}^{m}\left|h(x_i) - y_i\right|. \tag{15}$$

$h(x_i)$ is the predicted demand value of the model for the x_i-th sample point. y_i is the real demand data of the x_i-th sample point. m is the number of predicted samples. (16) displays the RMSE:

$$\text{RMAE} = \sqrt{\frac{1}{m}\sum_{i=1}^{m}\left(\left|h(x_i) - y_i\right|\right)^2}. \tag{16}$$

(17) is Spearman's rank correlation coefficient.

$$r_s = \frac{\sum_{i=1}^{m}\left[\left(h(x_i) - \overline{y}\right)\left(h_1(x_i) - \overline{y}_1\right)\right]}{\sqrt{\sum_{i=1}^{m}\left(h(x_i) - \overline{y}\right)^2}\sqrt{\sum_{i=1}^{m}\left(h_1(x_i) - \overline{y}_1\right)^2}}. \tag{17}$$

In (17), $-1 \le r_s \le 1$. The larger the r_s is, the greater the correlation between the two feature information sequences IS. When $h(x_i)$ and $h_1(x_i)$ are the values of the two feature information sequences at the x_i-th sample point, respectively, $\overline{y}$ and $\overline{y}_1$ are the average values of the two feature information sequences.

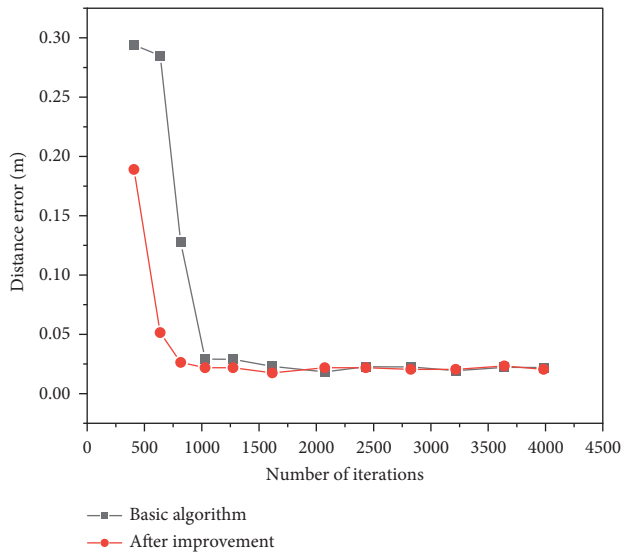


Figure 8: Distance errors of the flower pollination algorithms.

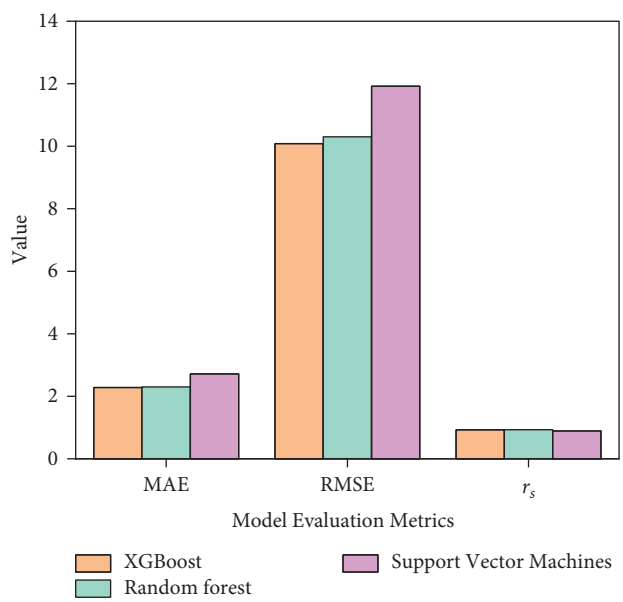


Figure 9: Error of the prediction model in supply chain management.

3. Results and Discussion

3.1. Positioning and Tracking of the Intelligent Logistics System and the Logistics Status in Supply Chain. The intelligent logistics state is perceived according to the flower pollination algorithm. The distance errors of the products in the logistics system are compared before and after the optimization of the algorithm. The comparison results of the distance errors of the flower pollination algorithms are shown in Figure 8.

Figure 8 shows that when the number of iterations is 500, the distance error of the traditional flower pollination algorithm is 0.28 M, and that of the optimized algorithm is 0.13 M. This shows that the more the times of iterations are, the more stable the distance errors will be before and after optimization. With the increase of iteration times before and after optimization, the optimized algorithm has high accuracy in the distance error, which realizes the positioning and tracking of the intelligent logistics system and the logistics state query in the supply chain management.

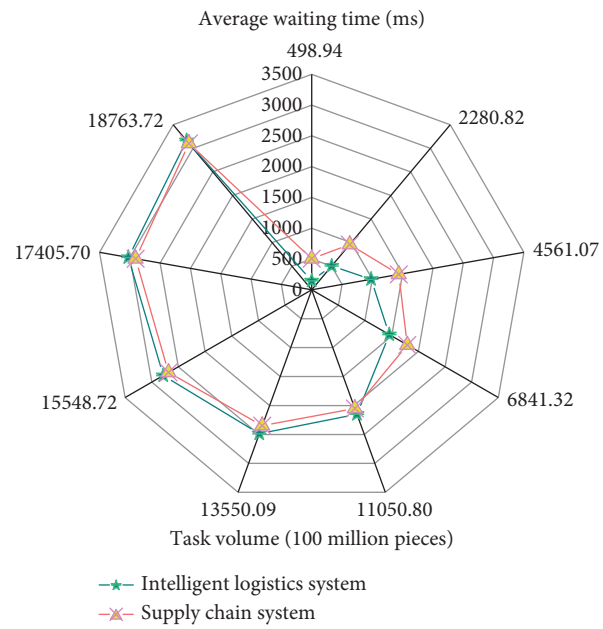


Figure 10: Average waiting time of the intelligent logistics system and supply chain management system in operation.

3.2. Analysis of a Model of Supply Chain Management Prediction and Operation Efficiency in the Supply Chain Management System. XGBoost is compared with random forest and support vector machine algorithms. Figure 9 shows the error of the prediction model in supply chain management.

Figure 9 shows that there is little difference between the MAE value of XGBoost and that of random forest, indicating that the average fitting ability of XGBoost and random forest model in the prediction model of supply chain management is similar. The MAE value of the support vector machine model is the largest, indicating that the support vector machine model is not suitable for evaluating the prediction model. The RMSE of XGBoost is the smallest, which indicates that the predicted value of XGBoost t is the closest to the real value in the supply chain prediction model, and the numerical prediction ability of XGBoost is higher than that of the other two models. In addition, the fitting ability of the three models is good. The intelligent logistics monitoring system analyzes the data uploaded by each terminal when storing data in the cloud. The actual route and node time of goods transportation are displayed, so that the transportation status of goods is known. If the goods transportation route is set in advance, the alarm can be triggered when the goods deviate from the set route. The average waiting time of the intelligent logistics system and supply chain management system in operation is analyzed, as shown in Figure 10.

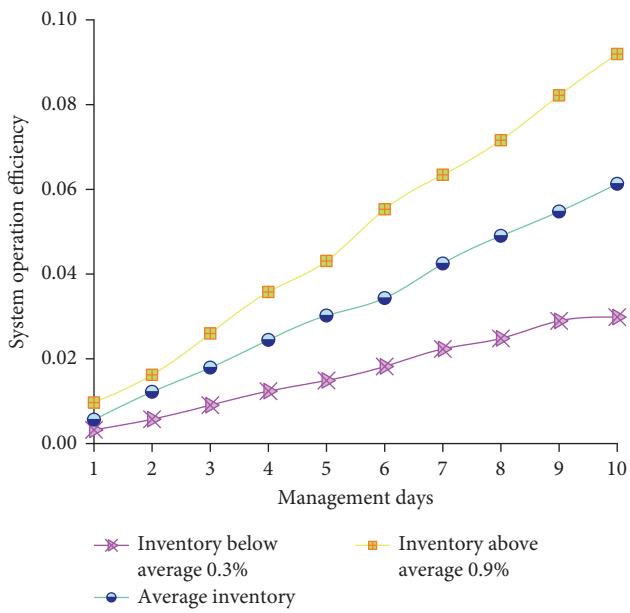


Figure 11: Operation efficiency of the supply chain management system.

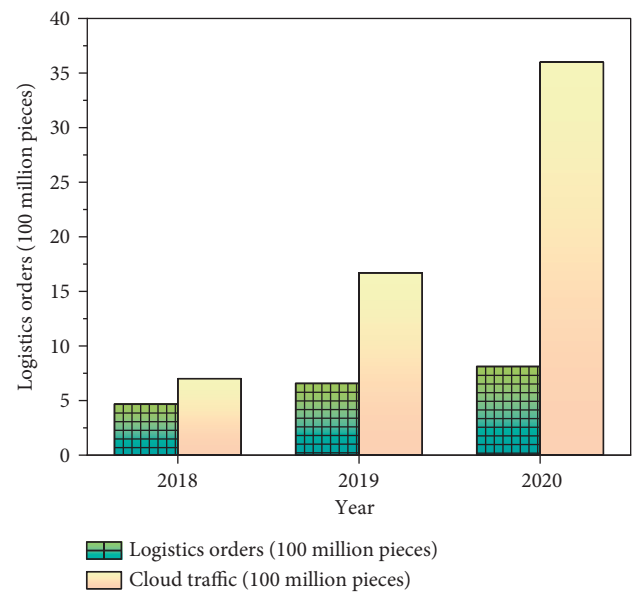


Figure 12: Changes of the material flow of the intelligent logistics system.

Figure 10 shows that the average waiting time of the intelligent logistics system in operation is 1855.379 milliseconds, and that of the supply chain system is 1976.631. This shows that the larger the amount of the task is, the longer the waiting time for system operation will be. In different periods, the waiting time of intelligent logistics system is shorter than that of supply chain management system.

3.3. Operation Efficiency of the Supply Chain Management System. The operation efficiency of the intelligent system and the supply chain management within 10 days is analyzed. After an in-depth analysis of the product transportation, route, and quantities, an exemplary management system is established based on the supply chain to realize the planning and management of the intelligent logistics supply chain. Figure 11 shows the relationship between the supply chain management system and its inventory.

Figure 11 shows that the operation efficiency of the supply chain management system is compared and analyzed when its logistics inventory is at the average, 0.3% less than the average, and 0.9% less than the average inventory. It is found that the operation efficiency of the supply chain management system with an average inventory is lower than that of 0.3% less than the average inventory and higher than that of 0.9% less than the average inventory. With the increase of system operation days, the operation efficiency of the supply management system is more elevated.

3.4. Material Flow Changes of the Intelligent Logistics System. Based on the analysis of the changes of material flows of the intelligent logistics system from 2018 to 2020, intelligent cloud logistics iterate rapidly under AI, and the "intelligent revolution" changes the pattern of logistics markets. This shows that symbiosis and sharing are the developing trends, the service economy and experience economy are deepened, new ways of division and cooperation are widely used, and a win-win logistics ecosystem is established. It is estimated that by 2025, the revenue of the intelligent logistics market will exceed 1 trillion. Figure 12 shows the changes in the material flow of the intelligent logistics system.

Figure 12 shows that intelligent cloud logistics develops rapidly from 2018 to 2020. Compared with Intelligent Cloud logistics, logistics orders are much fewer than traditional logistics systems. The difference in the revenue between the two is 233 million in 2018 and 2.788 billion in 2020. Intelligent logistics system has a significant impact on the development of the logistics industry.

4. Conclusion

The intelligent logistics system and the supply chain management system under the IoT are discussed based on EC, and the flower pollination algorithm is used to determine the positions in the system. When the times of iterations are 500, the distance error of the traditional algorithm is 0.28 M and that of the optimized algorithm is 0.13 M. The more the times of iterations are, the distance error tends to be stable before and after optimization. The average waiting time of the intelligent logistics system is 1855.379 milliseconds and that of the supply chain system is 1976.631 milliseconds. This shows that the larger the amount of the task is, the longer the waiting time for system operation is. XGBoost and random forest model have similar average fitting ability in the prediction model of supply chain management, while support vector machine model is not suitable for evaluating the prediction model. In this supply chain prediction model, the predicted value of XGBoost model is the closest to the real value, and the numerical prediction ability of XGBoost is higher than that of the other two models. The changes in the

material flow of the intelligent logistics system from 2018 to 2020 show that with the increase of system operation days, the operation efficiency of the supply management system is higher. The intelligent logistics system has a great impact on the development of the logistics industry. This research gives a reference for establishing the intelligent logistics system and supply chain management system. However, with the rapid development of the IoT, the operating system in the logistics industry should be updated to adapt to the changes in the new era.

Data Availability

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Consent

Informed consent was obtained from all individual participants included in the study.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

[1] K. Leng, L. Jin, W. Shi, and I. Van Nieuwenhuyse, "Research on agricultural products supply chain inspection system based on internet of things," *Cluster Computing*, vol. 22, no. S4, pp. 8919–8927, 2019.

[2] M. Humayun, N. Z. Jhanjhi, B. Hamid, and G. Ahmed, "Emerging smart logistics and transportation using IoT and blockchain," *IEEE Internet of Things Magazine*, vol. 3, no. 2, pp. 58–62, 2020.

[3] M. Ben-Daya, E. Hassini, and Z. Bahroun, "Internet of things and supply chain management: a literature review," *International Journal of Production Research*, vol. 57, no. 15-16, pp. 4719–4742, 2019.

[4] Y. P. Tsang, C. H. Wu, H. Y. Lam, K. L. Choy, and G. T. S. Ho, "Integrating Internet of Things and multi-temperature delivery planning for perishable food E-commerce logistics: a model and application," *International Journal of Production Research*, vol. 59, no. 5, pp. 1534–1556, 2021.

[5] C. Liu, Y. Feng, D. Lin, L. Wu, and M. Guo, "Iot based laundry services: an application of big data analytics, intelligent logistics management, and machine learning techniques," *International Journal of Production Research*, vol. 58, no. 17, pp. 5113–5131, 2020.

[6] S. Winkelhaus and E. H. Grosse, "Logistics 4.0: a systematic review towards a new logistics system," *International Journal of Production Research*, vol. 58, no. 1, pp. 18–43, 2020.

[7] T. Bosona, "Urban freight last mile logistics—challenges and opportunities to improve sustainability: a literature review," *Sustainability*, vol. 12, no. 21, p. 8769, 2020.

[8] K. T. Park, Y. H. Son, and S. D. Noh, "The architectural framework of a cyber physical logistics system for digital-twin-based supply chain control," *International Journal of Production Research*, vol. 59, no. 19, pp. 5721–5742, 2021.

[9] Y. Ding, M. Jin, S. Li, and D. Feng, "Smart logistics based on the internet of things technology: an overview," *International Journal of Logistics Research and Applications*, vol. 24, no. 4, pp. 323–345, 2021.

[10] V. W. B. Martins, R. Anholon, O. L. G. Quelhas, and W. Filho, "Sustainable practices in logistics systems: an overview of companies in Brazil," *Sustainability*, vol. 11, no. 15, p. 4140, 2019.

[11] M. Viu-Roig and E. J. Alvarez-Palau, "The impact of E-Commerce-related last-mile logistics on cities: a systematic literature review," *Sustainability*, vol. 12, no. 16, p. 6492, 2020.

[12] M. K. Anser, M. A. Khan, U. Awan et al., "The role of technological innovation in a dynamic model of the environmental supply chain curve: evidence from a panel of 102 countries," *Processes*, vol. 8, no. 9, p. 1033, 2020.

[13] S. B. Keller, P. M. Ralston, and S. A. LeMay, "Quality output, workplace environment, and employee retention: the positive influence of emotionally intelligent supply chain managers," *Journal of Business Logistics*, vol. 41, no. 4, pp. 337–355, 2020.

[14] Y. Wang, "Negotiation and scheduling decisions of marine intelligent supply chain based on stochastic distribution," *Journal of Coastal Research*, vol. 98, no. sp1, pp. 183–186, 2019.

[15] A. Darwish, A. E. Hassanien, M. Elhoseny, A. K. Sangaiah, and K. Muhammad, "The impact of the hybrid platform of internet of things and cloud computing on healthcare systems: opportunities, challenges, and open problems," *Journal of Ambient Intelligence and Humanized Computing*, vol. 10, no. 10, pp. 4151–4166, 2019.

[16] P. Sotres, J. Lanza, L. Sánchez, J. R. Santana, C. Lopez, and L. Munoz, "Breaking vendors and city locks through a semantic-enabled global interoperable internet-of-things system: a smart parking case," *Sensors*, vol. 19, no. 2, p. 229, 2019.

[17] X. Wang, Y. Han, V. C. M. Leung, D. Niyato, X. Yan, and X. Chen, "Convergence of edge computing and deep learning: a comprehensive survey," *IEEE Communications Surveys & Tutorials*, vol. 22, no. 2, pp. 869–904, 2020.

[18] H. Fuketa and K. Uchiyama, "Edge artificial intelligence chips for the cyberphysical systems era," *Computer*, vol. 54, no. 1, pp. 84–88, 2021.

[19] C. Resende, D. Folgado, J. Oliveira et al., "TIP4. 0: industrial internet of things platform for predictive maintenance," *Sensors*, vol. 21, no. 14, p. 4676, 2021.

[20] G. Furano, G. Meoni, A. Dunne et al., "Towards the use of artificial intelligence on the edge in space systems: challenges and opportunities," *IEEE Aerospace and Electronic Systems Magazine*, vol. 35, no. 12, pp. 44–56, 2020.

[21] S. A. A. Hakeem, A. A. Hady, and H. W Kim, "5G.-V. 2X.: Standardization, architecture, use cases, network-slicing, and edge-computing," *Wireless Networks*, vol. 26, no. 8, pp. 6015–6041, 2020.

[22] E. Çakmak, İ Önden, A. Z. Acar, and F. Eldemir, "Analyzing the location of city logistics centers in istanbul by integrating geographic information systems with binary particle swarm optimization algorithm," *Case Studies on Transport Policy*, vol. 9, no. 1, pp. 59–67, 2021.

[23] X. Yan and J. Li, "Animal intelligent logistics management based on RFID technology," *Revista Científica de la Facultad de Ciencias Veterinarias*, vol. 29, no. 6, pp. 1772–1781, 2019.

[24] V. L. Da Silva, J. L. Kovaleski, and R. N. Pagani, "Technology transfer in the supply chain oriented to industry 4.0: a literature review," *Technology Analysis & Strategic Management*, vol. 31, no. 5, pp. 546–562, 2019.

[25] B. Feng and Q. Ye, "Operations management of smart logistics: a literature review and future research," *Frontiers of Engineering Management*, vol. 8, no. 3, pp. 344–355, 2021.

[26] S. Li, Q. Sun, and W. Wu, "Benefit distribution method of coastal port intelligent logistics supply chain under cloud

computing," *Journal of Coastal Research*, vol. 93, no. sp1, pp. 1041–1046, 2019.

[27] S. Mao, S. Leng, S. Maharjan, and Y. Zhang, "Energy efficiency and delay tradeoff for wireless powered mobile-edge computing systems with multi-access schemes," *IEEE Transactions on Wireless Communications*, vol. 19, no. 3, pp. 1855–1867, 2020.

[28] B. Lin, F. Zhu, J. Zhang et al., "A time-driven data placement strategy for a scientific workflow combining edge computing and cloud computing," *IEEE Transactions on Industrial Informatics*, vol. 15, no. 7, pp. 4254–4265, 2019.

[29] M. Ke, Z. Gao, Y. Wu, X. Gao, and K. K. Wong, "Massive access in cell-free massive MIMO-based Internet of Things: cloud computing and edge computing paradigms," *IEEE Journal on Selected Areas in Communications*, vol. 39, no. 3, pp. 756–772, 2021.

[30] W. Hu, "An improved flower pollination algorithm for optimization of intelligent logistics distribution center," *Advances in Production Engineering & Management*, vol. 14, no. 2, pp. 177–188, 2019.

[31] G. S. Shehu and N. Çetinkaya, "Flower pollination–feedforward neural network for load flow forecasting in smart distribution grid," *Neural Computing & Applications*, vol. 31, no. 10, pp. 6001–6012, 2019.

Hindawi
Computational Intelligence and Neuroscience
Volume 2022, Article ID 7131143, 7 pages
https://doi.org/10.1155/2022/7131143

Research Article

Application of Neural Network with Autocorrelation in Long-Term Forecasting of Systemic Financial Risk

Junzhi Zhang and Lei Chen

School of Management, China University of Mining and Technology-Beijing, Beijing 100083, China

Correspondence should be addressed to Lei Chen; lchen@student.cumtb.edu.cn

Received 11 August 2022; Revised 4 September 2022; Accepted 10 September 2022; Published 16 September 2022

Academic Editor: Ning Cao

Carrying out early warning of systemic financial risk is a prerequisite for timely adjustment of monetary policy and macroprudential policy to effectively prevent and resolve systemic financial risks. This paper constructs a systemic financial risk monitoring and early warning system for China's banking industry based on isolated forest anomaly detection and neural network with autocorrelation mechanism and uses low-frequency data with high credibility to effectively identify the ten factors that have the greatest impact on systemic financial risk in China's banking industry, improving the prospective and accuracy of risk early warning. The conclusions can help regulators to adjust their policies prospectively to curb the rise of systemic financial risks.

1. Introduction

The subprime crisis in 2007 made governments and international academics realize that microprudential regulation for individual financial institutions insufficiently reflects the actual risk accumulation in the financial system and that macroprudential regulation for financial system needs to be strengthened to maintain financial stability. In the following decade or so, governments have devoted themselves to the establishment and improvement of macroprudential policy frameworks. Among them, early warning (EW) of systemic financial risk (SFR) is crucial in macroprudential supervision. On December 31, 2021, People's Bank of China (PBC) formulated and published Macroprudential Policy Guidelines (Trial Implementation), which emphasize the need to establish a sound framework for SFR's monitoring, assessment, and early warning and actively explore the use of big data technology in SFS's monitoring and early warning. In China's Monetary Policy Implementation Report (2022 Q1), PBC pointed out that financial work should further improve the macroprudential policy framework, improve SFR's monitoring, assessment, and early warning capabilities, and enrich macroprudential policy toolbox. Thus, it is clear that exploring the construction of SFR's EW system will be a long-term topic in financial research.

Regarding the EW models of SFR, the commonly used models include the logit model [1–3], probit model [4], KLR signal model [5], STV cross-sectional regression model [6], and so on. In addition, many scholars have also combined the analytical frameworks such as binary classification tree (BCT) [7], GARCH model [8], quantile projection [9], and jump undetermined equity analysis [10] to further improve the prediction models of systemic financial risk. Further, with the continuous research in this field, the latest research has made good progress in effectively predicting systemic financial risk using cutting-edge machine learning models such as random forest [11–13], neural network [12, 14–16], and support vector machine (SVM) [12, 17, 18].

At present, domestic research studies on systemic financial risk early warning are also quite abundant [19–23], but they also face the following difficulties: (1) the use of high-frequency data, such as stock data, can improve the timeliness of risk early warning. However, relevant studies also point out that the development of China's financial market is relatively short and immature, and the effectiveness of public market data needs to be further tested [24, 25];

(2) the risk early warning system using asset and liability data can better reflect the risk changes in related domestic fields, but the low frequency of data limits the foresight of early warning; and (3) risk forecasting techniques based on econometric models, such as GARCH-like models, are to fit in each sequence separately, which cannot take advantage of the similarity patterns existing in different sequences.

The development of deep learning provides a useful exploration to solve the above difficulties. The autoformer model is improved on the basis of transformer based on deep decomposition architecture and the autocorrelation mechanism [26]. The model has improved the efficiency of long-term prediction through progressive decomposition and sequence-level connection and also makes the model itself capable of making long-term time series prediction based on less data.

Motivated by the above analysis, we construct a SFR's EW system for China's banking industry around the unique characteristics of the autoformer model. The reason for taking the banking industry as the early warning research object is that the banking industry in China accounts for more than 90% of the total assets of the financial industry for a long time, and the stability of banks is crucial to financial stability.

This paper is displayed as follows: firstly, we review the influencing factors of SFS and organize the set of covariate indicators for EW of bank systemic financial risk (BSFR) indexes in a phase dependent manner; secondly, we adopt the variance threshold method for initial screening of indicators and propose an indicator selection method based on the isolated forest method for further screening of indicators, the purpose of which is to downscale indicators and screen out indicators with high correlation with BSFR, and lay the foundation for the monitoring and EW of BSFR; thirdly, the early warning model is constructed around autoformer to predict the changes of bank systemic risk. Finally, this research is summarized, and conclusions and recommendations are drawn.

The contributions are as follows: (1) The indicator screening method based on isolated forest anomaly detection is completely data-driven, which not only identifies several factors that have the greatest impact on SFR of China's banking industry but also alleviates the interference of human subjective factors to a certain extent and (2) constructing an EW system for SFR based on autoformer not only exploits the high credibility of low-frequency data but also achieves the forward-looking nature of early warning.

2. Materials and Methods

2.1. Covariate Set Selection

2.1.1. Theoretical Analysis of SFR's Influencing Factors. Systemic financial risk refers to the risk caused by internal factors such as financial vulnerability or external factors such as policy adjustments and macroeconomic fluctuations, which can spread through the interinstitutional correlation network and cause the dysfunction of financial institutions, thus leading to the dysfunction of financial services, the spread of market panic, and ultimately leading to serious damage to the macro economy.

Based on the above definition of SFR and actuality of China's financial system, we grouped the factors into five dimensions:

Macroeconomic risk: Macroeconomic and financial systems have a clear "pro-cyclicality." When the macro economy continues to improve, credit in the banking sector rises rapidly and credit standards are relaxed; when the economy is in a downward phase, the downward pressure is transmitted to the banking sector, and the banks' asset quality comes under pressure, and credit contracts and the NPL rate rises. In the process of economic downturn and bank credit contraction, the default of the real economy induces liquidity crisis and credit crisis in the banking sector.

Credit risk: Against the backdrop of a continuous decline in macroeconomic growth, some enterprises' operating efficiency and debt-servicing ability have declined significantly, highlighting credit risk in the banking industry and deteriorating credit asset quality. As credit risk intensifies, it will lead to large-scale interest defaults, deterioration of bank-enterprise relationship, and mutual trust in the near term; in the long run, it will weaken the profitability of banks, and capital will be continuously eroded by nonperforming loans, and banks will be forced to contract credit when it is more difficult to replenish capital from external sources, resulting in a vicious circle of "credit contraction—contraction of the real economy—credit contraction."

Market risk: Declines in interest rates, equity, and foreign exchange and commodity prices can cause loss in banks' balance sheet operations. Currently, the capital finance lenders of equity financing and bond financing in China's capital market are mainly commercial banks, and the fall in the value of stocks and bonds erodes the quality of banks' assets and affects their normal operations. In addition, commercial banks' credit assets include mostly real estate collateral, and a decline in housing prices will lead to significant losses in bank assets.

Liquidity risk: Liquidity risk is the possibility that a bank will not have sufficient liquidity to settle liabilities as they fall due and meet customer withdrawal needs without increasing costs or losing asset value, resulting in a loss to the bank. The current liquidity risk of banks in China mainly comes from the business model of "borrowing short and lending long."

Correlation risk: Interbank correlation amplifies the contagion capacity of various risks. Under the modern credit-money system, monetary funds intended to maintain systemic stability and lower market interest rates may intensify commercial banks' over-reliance on wholesale funds and make profits through arbitrage by amplifying leverage and increasing maturity mismatch, thus generally forming a nested and intertwined interbank credit network, where problems in one institution will quickly spread to the periphery, increasing the demand for liquidity settlement and exacerbating systemic risk.

2.1.2. Selection of Factors Influencing Systemic Financial Risk. Based on the above theoretical analysis and previous researches [13, 18, 27–31], this paper selects covariate indicators for constructing the systemic financial risk early

warning in the banking sector system. The indicators of macroeconomic risk include GDP growth rate (GDP), M2 growth rate (M2), consumer confidence index (CCI), growth rate of total retail sales of consumer goods (TRSCG), consumer price index (CPI), product inventory growth rate of industrial enterprises (PIIE), fixed assets investment growth rate (FAI), industrial entrepreneur confidence index (IECI), industrial enterprise prosperity index (IEPI), investment growth in urban real estate development (IURED), state housing boom index (SHB), government leverage ratio (GLR), leverage ratio of resident sector (RLR), and leverage ratio of real economy department (RELR). The indicators of credit risk include commercial bank provision coverage rate (PCR) and nonperforming loan ratio of commercial banks (NPL). The indicators of liquidity risk include weighted average interest rate of interbank lending (WAIR), capital adequacy ratio of commercial banks (CAR), liquidity ratio of commercial banks (LRCB), loan-to-deposit ratio of commercial banks (LDR), and commercial bank excess reserve ratio (ERR). The indicators of market risk include exchange rate (ER), CSI corporate bond index (CBI), CSI 300 index (CSI 300), proportion of real estate loans (REL), and proportion of foreign exchange loans (FEL). The indicators of correlation risk include interbank asset dependence (IBA) and interbank liability dependence (IBL).

Since no systemic financial crisis has occurred in China, it is not appropriate to use 0 or 1 as a proxy for the change of systemic financial risk. In view of this, this section adopts the Chinese systemic financial risk (SRISK) data published by the Stern School of Business of New York University as a proxy variable for systemic financial risk, and the covariate data are obtained from the websites of National Bureau of Statistics, PBC, CBRC, and Wind. The time span is from February 2011 to December 2020.

2.2. Screening of Covariate Indicators. Commonly used indicator selection methods are filtering, wrapping, and embedding methods. Except for some algorithms in the filtering method, other algorithms are generally used for supervised learning, while the risk identification problem studied in this section is not easy to get the category to which the sample belongs in advance and should belong to unsupervised learning [29].

Based on the above analysis, this section adopts an indicator screening method that includes the variance threshold method and isolated forest method [29]. Isolated forests segment samples by random features and anomalies are more easily isolated compared to ordinary sample points so that anomalous sample points have shorter paths. Therefore, whether a sample point is an outlier can be determined by the following equation:

$$S(x, \psi) = 2^{-(E(h(x))/c(\psi))}, \tag{1}$$

where $h(x)$is the height of sample point x in each tree, $c(\psi)$ is the average of the path lengths for a given number of samples ψ, and S $(x\psi)$ denotes the anomaly score value of sample point x. If the anomaly score S is close to 1, then the sample point x must be an anomaly; if the anomaly score S is much less than 0.5, then the sample point x must not be an anomaly.

The specific steps of indicator screening combining the variance threshold method and isolated forest anomaly detection are as follows:(1) Find the variance of all indicators and use the mean of variance as the threshold to initially screen the indicator set. (2) Randomly select a subset of the indicator set obtained in the first step and standardize the indicator data of the subset. (3) Use isomap to downscale the standardized subset of indicator data. The isomap algorithm can better control the loss of data information and can represent the data in higher dimensions more comprehensively in the lower dimensional space. (4) Apply the isolated forest method to the downscaled data for anomaly detection and obtain the score of each sample point. (5) Repeat step (4) 10 times to obtain the average score of each sample point. (6) Obtain the Euclidean distance from each abnormal sample point to the normal sample centre in the subset and obtain the median of all Euclidean distances in the corresponding subset. (7) Repeat steps (2) to (6) n times to obtain n median Euclidean distances, and the subset with the largest median Euclidean distance is the best subset.

2.3. Systemic Financial Risk Early Warning Model. For time series forecasting problems, traditional statistical approaches focus on providing parametric models from the domain expertise level, such as autoregressive (AR), exponential smoothing, or structural time series models, but the application of traditional statistical methods is also limited by data validity and frequency. Autoformer models based on deep decomposition architecture provide a new research solution to address these issues. It can significantly improve long-time forecasting by coping with complex temporal patterns and information utilization bottlenecks through progressive decomposition and sequence-level connectivity.

Autoformer completely revolutionizes transformer as deep decomposition architecture, embedding sequence decomposition into the encoder-decoder as an internal unit of autoformer [26]. In the prediction process, the model alternately optimizes the prediction results and decomposes the sequence; that is, the trend item and the periodic item are gradually separated from the hidden variables to realize the gradual decomposition.

The sequence decomposition unit is based on the idea of moving average, smoothing the period term, and highlighting the trend term:

$$\begin{aligned} x_t &= \mathrm{AvgPool}(\mathrm{Padding}(x)), \\ x_s &= x - x_t, \end{aligned} \tag{2}$$

where x is the hidden variable to be decomposed, and x_t and x_s are the trend term and the period term, respectively. The above equations are collectively referred to as the series decomposition $x_t, x_s = SeriesDecomp(x)$ and will be embedded into autoformer layers.

In the encoder part, the model gradually eliminates the trend term and obtains the periodic terms $S_{en}^{l,1}, S_{en}^{l,2}$. Based on this periodicity, the model aggregates similar subprocesses of

different periods by designing an autocorrelation mechanism to achieve information aggregation. The relevant equations are as follows：

$$\begin{aligned} S_{en}^{l,1}, &= \text{SeriesDecomp}\left(\text{AutoCorrelation}\left(x_{en}^{l-1}\right) + s_{en}^{l-1}\right), \\ S_{en}^{l,2}, &= \text{SeriesDecomp}\left(\text{Fee dF orwar d}\left(S_{en}^{l,1}\right) + S_{en}^{l,1}\right). \end{aligned} \tag{3}$$

In the decoder part, the model shows the trend term and the period term separately, in which, for the period term, the autocorrelation mechanism uses the periodic nature of the sequence to aggregate subsequence with similar processes in different cycles; for the trend term, the trend information is gradually extracted from the predicted hidden variables using a cumulative approach. The correlation equation is as follows:

$$\begin{aligned} S_{de}^{l,1}, \mathcal{T}_{de}^{l,1} &= \text{SeriesDecomp}\left(\text{AutoCorrelation}\left(x_{de}^{l-1}\right) + x_{de}^{l-1}\right), \\ S_{de}^{l,2}, \mathcal{T}_{de}^{l,2} &= \text{SeriesDecomp}\left(\text{AutoCorrelation}\left(x_{de}^{l,1}, x_{en}^{N}\right) + S_{de}^{l,1}\right), \\ S_{de}^{l,3}, \mathcal{T}_{de}^{l,3} &= \text{SeriesDecomp}\left(\text{Fee dF orwar d}\left(S_{de}^{l,2}\right) + S_{de}^{l,2}\right), \\ \mathcal{T}_{de}^{l} &= \mathcal{T}_{de}^{l-1} + \mathcal{W}_{l,1} * \mathcal{T}_{de}^{l,1} + \mathcal{W}_{l,2} * \mathcal{T}_{de}^{l,2} + \mathcal{W}_{l,3} * \mathcal{T}_{de}^{l,3}. \end{aligned} \tag{4}$$

Based on the above progressive decomposition architecture, the autoformer model can gradually decompose the hidden variables in the forecasting process and obtain the forecasting results of periodic and trend components through autocorrelation mechanism and accumulation, respectively, so as to realize the alternate and mutual promotion of decomposition and prediction results optimization.

Based on the above analysis, this paper adopts autoformer to construct the early warning system of China's banking systemic financial risk. The early warning model takes systemic financial risk as the variable to be predicted and the best subset of indicators is screened based on the variance threshold method and isolated forest algorithm as covariates.

3. Results

3.1. Covariates Screening and Analysis

3.1.1. Result of the Variance Threshold Method. The principle of the variance threshold method is to first estimate the variance of each indicator and then screen the indicators according to a threshold value. If the variance of an indicator is small, it means that the overall fluctuation of the indicator is small; that is, it contributes less to the anomaly, and the indicator can be removed more safely. The calculation steps are as follows: (1) standardize the sample data; (2) calculate the variance of each indicator; and (3) filter the indicators according to a threshold value. The obtained calculation results are displayed in Table 1.

The mean value of variance (0.06679) is treated as threshold, and 15 indicators are selected for the systemic financial risk early warning model: commercial bank provision coverage rate (PCR), nonperforming loan ratio of commercial banks (NPL), leverage ratio of resident sector (RLR), interbank asset dependence (IBA), consumer confidence index (CCI), CSI corporate bond index (CBI), leverage ratio of real economy department (RELR), loan-to-deposit ratio of commercial banks (LDR), exchange rate (ER), liquidity ratio of commercial banks (LRCB), M2 growth rate (M2), state housing boom index (SHB), interbank liability dependence (IBL), government leverage ratio (GLR), and CSI 300 index.

Table 1: Variance of indicators.

Indicator	Variance	Indicator	Variance
PCR	0.13569	CSI300	0.06702
NPL	0.12835	CAR	0.06413
RLR	0.11398	REL	0.06120
IBA	0.09969	PIIE	0.05430
CCI	0.09319	ERR	0.04928
CBI	0.09303	FEL	0.04746
RELR	0.09038	CPI	0.03768
LDR	0.08995	IECI	0.03276
ER	0.08632	IEPI	0.03062
LRCB	0.08626	IURED	0.02932
M2	0.08441	WAIR	0.02409
SHB	0.07815	FAI	0.02285
IBL	0.07685	TRSCG	0.01378
GLR	0.06792	GDP	0.01138

3.1.2. Result of Isolated Forest. Isolated forest anomaly detection is applied for further screening indicators. All subsets are selected using the exhaustive method, and the median Euclidean distance of each abnormal sample point is calculated for all subsets. The subset of indicators corresponding to the largest median Euclidean distance includes commercial bank provision coverage rate (PCR), nonperforming loan ratio of commercial banks (NPL), leverage ratio of resident sector (RLR), interbank asset dependence (IBA), consumer confidence index (CCI), CSI corporate bond index (CBI), loan-to-deposit ratio of commercial banks (LDR), liquidity ratio of commercial banks (LRCB), state housing boom index (SHB), and interbank liability dependence (IBL). The top five subsets with the largest median Euclidean distances are shown in Table 2.

Table 3 shows the definition of the selected indicators. Consumer confidence index, leverage ratio of residential sector, and state housing boom index belong to macroeconomic risk; CSI corporate bond index belongs to market risk; nonperforming loan ratio and provision coverage ratio of commercial banks belong to credit risk of commercial banks; loan-to-deposit ratio and liquidity ratio reflect liquidity risk of commercial banks; and interbank asset and liability dependence reflects the correlation risk of commercial banks.

3.1.3. Logical Analysis of the Selected Indicators and Systemic Financial Risk. The indicator screening system based on the variance threshold method and the isolated forest anomaly detection method provides the selected indicators for the systemic financial risk early warning system in a purely data-driven manner, but the monetary policy and macroprudential supervision policy for preventing, controlling,

Table 2: The five largest subsets of Euclidean distance.

Index	Subset	Distance
1	PCR, NPL, RLR, IBA, CCI, CBI, LDR, LRCB, SHB, IBL	1.668514
2	PCR, NPL, RLR, IBA, CCI, CBI, LRCB, SHB, IBL	1.649928
3	PCR, NPL, IBA, CCI, CBI, ER, LRCB, SHB, IBL	1.622538
4	PCR, NPL, IBA, CCI, CBI, RELR, LDR, M2, IIBL	1.604764
5	PCR, NPL, RLR, IBA, CBI, RELR, LDR, LRCB, SHB, IBL, GLR	1.599431

Table 3: Definition of selected indicators.

Indicator	Definition
Consumer confidence index	Reflects consumers' expectations to overall economic development and the consumer sector
Leverage ratio of residential sector	It is the ratio of residents' debt to disposable income and an important indicator of the level of residents' debt
State housing boom index	A comprehensive index reflecting the boom status of the real estate industry
CSI corporate bond index	A comprehensive index reflecting the overall price movement trend of corporate bonds in the exchange market
Commercial bank provision coverage rate	Reflects bank's financial stability and loss absorption capacity
Nonperforming loan ratio of commercial banks	Reflects the safety of the bank's credit assets
Loan-to-deposit ratio of commercial banks	The ratio of total bank loans to total deposits
Liquidity ratio of commercial banks	The ratio of liquid asset to liquid liability, which measures the overall level of commercial banks' liquidity
Interbank asset dependence	Ratio of interbank assets to total assets
Interbank liability dependence	Ratio of interbank liabilities to total liabilities

and resolving systemic financial risks need to clarify the drivers behind the risks in order to accurately block and defuse risks. Therefore, this subsection analyses the logical relationship between the selected indicators and systemic financial risk from the perspective of the correlation between modern money supply and bank liquidity.

Under the modern credit monetary system, the central bank issues the base currency through asset expansion, and commercial banks create broad money through asset expansion [32]. In general, a broad money supply system with the base currency as the reserve is formed. Under this monetary supply system, the central bank injects liquidity into commercial banks by expanding and adjusting assets, and commercial banks provide liquidity to the banking system and the real economy through asset-side creation and active debt behaviour. Under the modern credit monetary system, China has now formed the following liquidity transmission structure: the central bank injects liquidity into commercial banks through debit or open market operations (liquidity facilitation tools) ; commercial banks with primary traders receive funds first, and then can lend them to small and medium-sized banks that do not have the qualifications of primary traders through interbank certificates of deposit and lending; commercial banks put credit into three areas such as government platform, real estate, and nonbank financial institutions; and the liquidity of nonbank financial institutions also flows to government platform and real estate.

Based on the above analysis, the logical relationship between the selected indicators and China's systemic financial risk is as follows: expectations about the economy (consumer confidence index) influence the leveraging behaviour of the residential sector (residential sector leverage ratio), which in turn influences commercial banks' liquidity allocation to the real estate market and the real economy (State Housing Boom Index and CSI Corporate Bond Index); commercial banks' liquidity allocation behaviour influences their credit risk (nonperforming loan ratio and provision coverage ratio) and liquidity risk (loan-to-deposit ratio, liquidity ratio). Under the external disturbance, large commercial banks replenish their liquidity gap through wholesale funding, while small and medium-sized banks rely on interbank business (interbank assets and interbank liabilities dependency) to obtain liquidity. Under the influence of residents' leveraging behaviour and commercial banks' liquidity placement and acquisition behaviour, China's systemic financial risks show a clustering and accumulation situation.

3.2. Systemic Financial Risk Forecasting and Analysis

3.2.1. Model Prediction Based on the Predictor and Covariates Indicators. The prediction indicators and covariate indicators are input into the autoformer model. In modelling, the data from February 2011 to December 2018 are divided into training sets, the data from January 2019 to December 2019 are divided into verification sets, and the data from January 2020 to December 2020 are used as test sets. The prediction error (MSE, MAE) of January 2020 (1 step), March 2020 (3 steps), June 2020 (6 steps), September 2020 (9 steps) is used to measure the model test error. The smaller the value of MSE and MAE, the better the prediction effect of the model.

Table 4: Prediction results based on the predictor.

Model		Autoformer		Transformer		DeepAR		LSTM		ARIMA	
Metrics		MSE	MAE	MSE	MSE	MAE	MAE	MSE	MAE	MSE	MAE
Prediction length	1	**0.0498**	**0.1883**	0.0702	0.2166	0.0885	0.2262	0.2561	0.4354	0.1903	0.3630
	3	**0.0503**	**0.1946**	0.0782	0.2260	0.0976	0.2525	0.2599	0.4269	0.2076	0.3360
	6	**0.0527**	**0.1991**	0.1019	0.2651	0.1734	0.3230	0.2293	0.4074	0.2207	0.3391
	9	**0.0850**	**0.2468**	0.1778	0.3228	0.2263	0.3796	0.4464	0.5174	0.2797	0.4176

Table 5: Prediction results based on the predictor and covariates indicators.

Model		Autoformer		Transformer		LSTM	
Metrics		MSE	MAE	MSE	MAE	MSE	MAE
Prediction length	1	**0.1997**	**0.3047**	0.7443	0.6600	2.0166	1.0742
	3	**0.2230**	**0.2417**	0.9108	0.7229	2.2230	1.0677
	6	**0.2744**	**0.2718**	1.2039	0.7880	2.4971	1.0996
	9	**0.3496**	**0.3739**	3.0996	1.3835	2.6220	1.2000

Autoformer was used to predict SFR, and the prediction results are displayed in Table 4. For comparative analysis, this subsection also uses transformer and LSTM to predict SFR, where LSTM is used as the baseline model.

As shown in Table 5, the autoformer model predicts the systemic financial risk much better than the other two models. The steady change of autoformer as the prediction length increases indicates that autoformer has long-term robustness.

3.2.2. Model Prediction Based on the Predictor. Early warning models are mostly constructed based on traditional econometrics, such as ARIMA and GARCH. The input of those models is only the predictor. Thus, we also only input the predictor into autoformer.

The results of the univariate prediction evaluation based on the predictor are shown in Table 5. The autoformer model still achieves better results in the long-term prediction.

4. Conclusion

Forward-looking early warning of systemic financial risks is necessary to ensure the sound operation of the banking system, which in turn lays the foundation for the safety and stability of the financial system.

This paper adopts the isolated forest algorithm and autoformer model to construct a systemic financial risk early warning system, which mainly solves the following problems: (1) Providing a theoretical basis for the selection of machine learning indicators. Based on the theoretical analysis of the mechanism of systemic financial risk generation in banks, this paper collates 28 indicators from five dimensions such as macroeconomic risk, market risk, liquidity risk, credit risk, and correlation risk, and then screens a total of 10 systemic financial risk early warning covariate indicators based on the variance threshold method and isolated forest algorithm, and analyses the logical relationship between the selected indicators and systemic financial risk. (2) Improving the accuracy of long-run forecasting of systemic financial risks. Compared with other machine learning models and traditional econometric models, this paper improves the prediction accuracy of systemic financial risk based on the autoformer model and also maintains robustness in long-order time prediction, which provides a better reference for regulatory policy expectation management.

This paper argues that in order to effectively prevent systemic financial risks, regulators can consider relying on administrative means to regularly obtain data on changes in the total amount of business transactions among financial institutions in addition to focussing on liquidity and credit risks of financial institutions so as to facilitate the analysis of correlation changes among financial institutions. In addition, according to the documents published by the regulators, the current additional regulation for systemically important banks in China mainly provides requirements in terms of additional capital, leverage ratio, large risk exposure, corporate governance, recovery and disposal plan, information disclosure and data reporting, etc., and involves less on liquidity management. Therefore, regulators should appropriately increase the content of liquidity management, develop liquidity management tools or indicators, and strengthen liquidity supervision of systemically important banks.

Data Availability

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

Conflicts of Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work was supported by School of Management, China University of Mining and Technology, Beijing.

References

[1] R. Barrell, E. P. Davis, D. Karim, and I. Liadze, "Bank regulation, property prices and early warning systems for banking crises in OECD countries," *Journal of Banking & Finance*, vol. 34, no. 9, pp. 2255–2264, 2010.

[2] G. Caggiano, P. Calice, L. Leonida, and G. Kapetanios, "Comparing logit-based early warning systems: does the duration of systemic banking crises matter?" *Journal of Empirical Finance*, vol. 37, pp. 104–116, 2016.

[3] C. Filippopoulou, E. Galariotis, and S. Spyrou, "An early warning system for predicting systemic banking crises in the Eurozone: a logit regression approach," *Journal of Economic Behavior & Organization*, vol. 172, pp. 344–363, 2020.

[4] T. Knedlik and R. Scheufele, *Three Methods of Forecasting Currency Crises: Which Made the Run in Signaling the South African Currency Crisis of June 2006?*, IWH Discussion Papers, Halle, Germany, 2007.

[5] G. L. Kaminsky, S. Lizondo, and C. M. Reinhart, "Leading indicators of currency crises," *IMF Working Papers*, vol. 97, no. 79, pp. 1–48, 1997.

[6] J. D. Sachs, A. Tornell, and A. Velasco, *Financial Crises in Emerging Markets: The Lessons from 1995*, National bureau of economic research Cambridge, Massachusetts, MA, USA, 1996.

[7] R. Duttagupta and P. Cashin, "Anatomy of banking crises in developing and emerging market countries," *Journal of International Money and Finance*, vol. 30, no. 2, pp. 354–376, 2011.

[8] Q. Chen, R. Gerlach, and Z. Lu, "Bayesian Value-at-Risk and expected shortfall forecasting via the asymmetric Laplace distribution," *Computational Statistics & Data Analysis*, vol. 56, no. 11, pp. 3498–3516, 2012.

[9] G. De Nicolò and M. Lucchetta, "Forecasting tail risks," *Journal of Applied Econometrics*, vol. 32, no. 1, pp. 159–170, 2017.

[10] W. Tang and F. Su, "An analysis of the effects of extreme financial events on systemic risk:evidence from China's banking sector," *Economic Research Journal*, vol. 04, pp. 17–33, 2017.

[11] T. Y. Wang, S. M. Zhao, G. X. Zhu, and H. T. Zheng, "A machine learning-based early warning system for systemic banking crises," *Applied Economics*, vol. 53, no. 26, pp. 2974–2992, Jun 2021.

[12] K. Bluwstein, M. Buckmann, A. Joseph, M. Kang, S. Kapadia, and Ö. Şimşek, *Credit Growth, the Yield Curve and Financial Crisis Prediction: Evidence from a Machine Learning Approach*, ECB Working Paper, Frankfurt am Main, Germany, 2020.

[13] D. Wang and Y. Zhou, "Application of random forest model in macro prudential regulation - an empirical study based on the data of 18 countries," *Studies of International Finance*, vol. 11, pp. 45–54, 2020.

[14] K. Ristolainen, "Predicting banking crises with artificial neural networks: the role of nonlinearity and heterogeneity," *The Scandinavian Journal of Economics*, vol. 120, no. 1, pp. 31–62, 2018.

[15] E. Tolo, "Predicting systemic financial crises with recurrent neural networks," *Journal of Financial Stability*, vol. 49, no. 19, Article ID 100746, Aug 2020.

[16] O. Hong-bing, H. Kang, and Y. Hong-ju, "Prediction of financial time series based on LSTM neural network," *Chinese Journal of Management Science*, vol. 28, no. 4, pp. 27–35, 2020.

[17] A. Takeda, S. Fujiwara, and T. Kanamori, "Extended robust support vector machine based on financial risk minimization," *Neural Computation*, vol. 26, no. 11, pp. 2541–2569, 2014.

[18] D. Zhao and J. Ding, "A research on banking systemic risk prediction in China——a modeling analysis based on support vector machines," *International Business*, vol. 04, pp. 100–113, 2019.

[19] C. Liu and Y. Zhu, "Research on the measurement framework of banking systemic risk," *Journal of Financial Research*, vol. 12, pp. 85–99, 2011.

[20] C. Wang and L. Hu, "An empirical research on early-warming of financial risk in China," *Journal of Financial Research*, vol. 9, pp. 99–114, 2014.

[21] C. Liu, J. Li, H. Yu, and Q. Xie, "Early warning of financial systemic risk in China: based on Markov regime-switching model," *Journal of Systems Engineering*, vol. 04, pp. 515–534, 2020.

[22] S. Tang and X. Zhou, "An empirical study on China's systematic financial risk and security early warning," *Macroeconomics*, vol. 03, pp. 48–61+117, 2018.

[23] L. Tao and Y. Zhu, "On China's financial systemic risks," *Journal of Financial Research*, vol. 06, pp. 18–36, 2016.

[24] Y. Fang and Z. Deng, "Research on contagion paths of systemic risk in the interbank system: based on network model with common assets holdings," *Studies of International Finance*, vol. 06, pp. 61–72, 2016.

[25] Z. Yang and D. Li, "Do systemic risk indicators have prospective predictive capability?" *China Economic Quarterly*, vol. 02, pp. 617–644, 2021.

[26] H. Wu, J. Xu, J. Wang, and M. Long, "Autoformer: decomposition transformers with auto-correlation for long-term series forecasting," *Advances in Neural Information Processing Systems*, vol. 34, 2021.

[27] Q. Sun and G.-h. Cui, "China banking systemic risk stress index design and empirical analysis," *Journal of Central University of Finance & Economics*, vol. 02, pp. 43–51, 2017.

[28] H. Zhu and X. Zhang, "The Early-warning System and Prevention Mechanism of Banking Systemic Risk," *Financial Regulation Research*, vol. 04, pp. 71–86, 2018.

[29] L. Chenying, "Identification and early warning of systemic risk of Chinese commercial banks," *Journal of Central University of Finance & Economics*, vol. 10, pp. 36–53, 2020.

[30] M. Li and S. Liang, "Monitoring systemic financial risks: construction and state identification of China's financial market stress index," *Journal of Financial Research*, vol. 06, pp. 21–38, 2021.

[31] Z. Zhang and Y. Chen, "Research on dynamic measurement of systemic financial risk and cross-sector network spillover effect," *Studies of International Finance*, vol. 01, pp. 72–84, 2022.

[32] J. Zhao, Q. Li, and S. Feng, "Money supply, liquidity fluctuation and systematic financial risk - micro-behavior, macro-structure and conduction mechanism," *Chinese Review of Financial Studies*, vol. 06, pp. 15–121, 2019.

Hindawi
Computational Intelligence and Neuroscience
Volume 2022, Article ID 8937084, 7 pages
https://doi.org/10.1155/2022/8937084

Research Article

Construction of Safety Early Warning Model for Construction of Engineering Based on Convolution Neural Network

Changge Zhao

Henan University of Engineering, Henan, Zhengzhou 45000, China

Correspondence should be addressed to Changge Zhao; zcg10801@haue.edu.cn

Received 23 June 2022; Revised 12 August 2022; Accepted 16 August 2022; Published 16 September 2022

Academic Editor: Ning Cao

In recent years, China's engineering construction management level has been greatly improved, but compared with other industries, the construction industry still has low production efficiency, serious waste, and low level of information problems, and especially in the process of engineering management practice, schedule delay has become the focus of engineering management problems. With the continuous development of science and technology, computer and information technology have been continuously applied in engineering, among which deep neural network (DNN) technology, lean management, information visualization technology, and other technologies have become the hot spot of industry research, and the application of these emerging technologies to improve the level of project schedule control has become an urgent demand of the industry. Therefore, on the basis of deep learning, this paper analyzes the principle and application of object detection and feature extraction constructed by neural network and combines text feature extraction and image feature extraction methods. This application provides a new idea for the development of the construction industry.

1. Introduction

Cross-media unified expression is an important research direction in the multimedia field [1], which aims to connect the semantic gap between different modal data, such as images and texts, and establish a unified semantic expression. Based on the theory and method of cross-media unified expression, some popular branch directions are derived, for example, visual question answering, image caption, and visual grounding.

With the development and in-depth research of deep learning in recent years, many difficult problems in the field of computer, especially those related to computer vision and natural language processing, have made great breakthroughs. Even on some problems, algorithms based on deep learning have surpassed human capabilities. Due to the powerful representation capability of DNN [2], deep learning algorithms have become the mainstream research direction in the field of cross-media.

Cross-media target retrieval has a wide range of application scenarios [3, 4]. Text or voice with dialogue question answering system has been widely used in mobile terminals and PC operating systems, as an important way of human-computer interaction, for example, Siri of Apple, Cortana of Microsoft, and Alexa of Amazon. Some wearable smart devices enable human-computer interaction, such as Microsoft HoloLens. In the future, text-based image object retrieval systems may become an important way of human-computer interaction, changing people's traditional production and lifestyle. This technology can greatly improve the efficiency of text-based image retrieval and object retrieval in the image.

At present, DNN technology in domestic architecture is still in the initial stage of exploration, and the theoretical technology and experimental research are still lagging behind the market demand for engineering practice. Therefore, we need to combine DNN technology software and project management methods in foreign buildings with China's national conditions to find out the applicable technology methods and standards suitable for China's construction industry. In the practice of architectural design and engineering management, how to improve the

efficiency of architectural design, effectively control the schedule, dynamically and rationally allocate all kinds of resources needed, and improve the scientific, efficient, and accurate management of architectural design and construction project are the problems to be solved.

In the field of construction engineering, many researchers have used deep learning to implement helmet detection. For example, the use of the convolutional neural network (CNN) to automatically extract image features for helmet detection is a topic that is still in the exploratory stage, and the technology is not yet mature. In the helmet detection process, computer vision is used to automatically detect whether a worker at a construction site is wearing a helmet, including body detection, helmet detection, and matching. When a mismatch is detected between the spatial position of the human body and the helmet, a safety alert is issued [4, 5].

In addition, the interior spatial layout is a typical application of deep learning in construction engineering [6]. With the development of 3D visual information processing technology, it is important to apply it to architectural interior space layout in combination with digital image processing technology in interior space planning and design. A safety alert is issued; visual parameter feature analysis methods are used to construct a building interior space layout feature extraction analysis model to achieve building interior space layout detection, thereby improving the quality of building interior space layout planning [7]. For the optimisation scheme of building interior space, deep learning can effectively extract building interior space features, identify image edge pixels, and determine the visual image points of the building interior space layout. A visual feature parameter fusion model for interior spatial layout is established as shown in Figure 1.

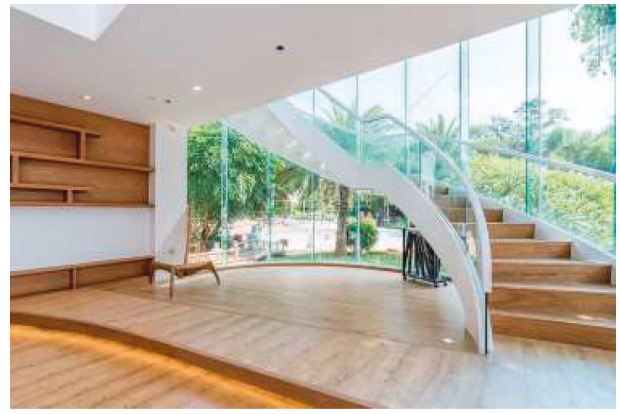


Figure 1: Indoor space layout sampling environment.

2. Related Work

2.1. Cross-Media Retrieval. Due to the basic characteristics of different media data, there is a problem of heterogeneity gap [8]. Many early researchers used traditional feature analysis methods such as principal component analysis and partial least square method. The subspace of the same dimension is obtained by dimensionality reduction of the underlying features of different media data. However, the subspace obtained by dimensionality reduction will lose some underlying features with potential correlation information, so it is difficult to accurately reflect the relationship between them in high-level semantics. Pay attention to the correlation of different media data contents, and use canonical correlation analysis [9] to find the correlation of different media data in the underlying feature space. At the same time, subspace mapping is carried out on this basis, which not only solves the heterogeneity of the underlying features but also ensures their correlation to the maximum extent.

In the topic of buildings, the building retrieval algorithm can be abstracted into an algorithm model containing three modules, for example, the method of building image location by reconstructing query phrases [10], the media retrieval method based on MCB [11], and the method of retrieving architectural image targets according to natural language [12, 13]. The general steps can be divided into three steps [14]. The first step is to generate candidate boxes and feature extraction. Given an input architectural image, a candidate box generation model is used to generate many candidate boxes and the corresponding visual features of each candidate box. DNN model is used to encode query text to extract text features. The second part is multimodal feature fusion. The visual features and query text features of each candidate box are fused using a cross-modal feature fusion algorithm. The simplest fusion methods, such as splicing and point multiplication, are used to complete cross-modal feature fusion. The third step is target positioning. A target box prediction model is designed to predict and output the most relevant target coordinate box of the query text on the image using the fusion feature obtained in the previous step. When processing building images or building texts, the general structure of cross-media retrieval is shown in Figure 2.

2.2. Application of DNN in Construction Engineering. Deep learning has a history of more than half a century. It first appeared in the early 1940s. From 1943 to now, deep learning has experienced three stages of development. The first is the beginning stage of deep learning, which is marked by the appearance of the MCP model, which lays the foundation for deep learning. The second is the rising stage of deep learning. Some researchers have put forward the concepts of BP neural network and activation function, making the neural network that has fallen into silence gradually become the focus of attention. The third stage, that is, the stage that deep learning is currently experiencing, is the outbreak of deep learning. The word "deep learning" has officially come into people's vision, and various deep neural network models shine in recognition tasks. Among them, the restricted Boltzmann machine, deep belief networks (DBN) [15], and CNN are excellent models in the field of deep learning.

In 2006, Hinton et al. proposed the concept of deep learning. Thus, in-depth learning has stepped into the third stage of development. In the 2012 ImageNet image

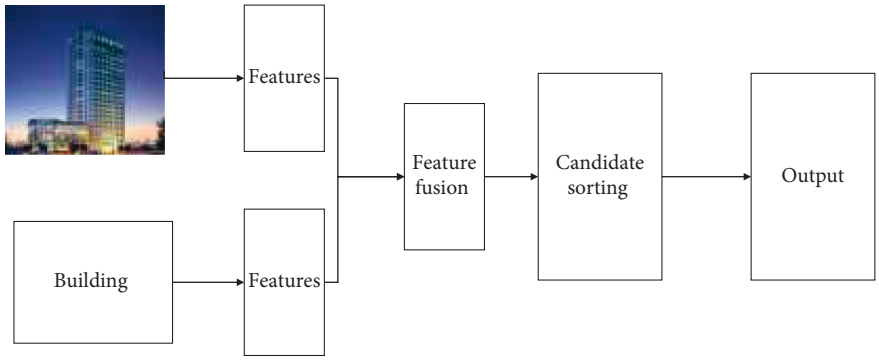


Figure 2: Cross-media retrieval and processing of construction engineering data.

recognition competition, Hinton's research group established the AlexNet model to reduce the error rate of ImageNet image classification to 15%. Taigman et al. proposed DeepFace model and achieved the accuracy of 97.35 in LFW face database [16]. In addition, the aldrop regularization method has achieved good results on image data sets. The development of deep learning has entered the fast lane.

In recent years, with the development of deep neural networks, various models have emerged CNN and recurrent neural networks, and other models perform well. CNN is one of the best models in deep learning. R-CNN can be used to improve the efficiency and visualization of 3d modeling. Mask R-CNN algorithm is used to segment architectural images, build 3D visualization models, and generate modeling images as shown in Figure 3.

Smart city has gradually become a new term in current urban architecture. The research focus in the process of smart city construction is to quickly realize the common three-dimensional visualization of indoor buildings and reduce the modeling cost. Many researchers have tried this. For example, 3D laser scanning combined with point cloud data is used for 3D modeling. Using the RCB-D indoor scene 3D space modeling giant [17, 18], the depth learning method is introduced into the RCB-D indoor modeling to effectively improve the flexibility and accuracy of indoor modeling.

2.3. Three-Dimensional Model Data. The 3D model is a rich and complex collection of information in a single point, line, and surface. In order to eliminate the effects of interference, conflicts, and noise, we need to preprocess the 3D model data to find the essential features of the model and minimise the extraction of nonessential features. In principle, the more key features are extracted from a model of the same species, the more accurate the model will be in terms of recognition and retrieval. Models have fixed shape features in the human subjective consciousness at different coordinates; however, for machine processing, they are very different. Therefore, the different processing of different coordinates and the subsequent retrieval of the model have a huge impact. The effectiveness of model preprocessing directly determines, in a simple and clear way, the accuracy of generalisation, classification, identification, filtering, and retrieval of models.

Figure 3: Animation 3D building visualization.

The experiments conducted in this paper are based on model data from the Princeton University shared 3D model standard library. This standard library contains several categories, and each category also contains a sufficient number of typical sample models, which were analysed, compared, identified, and selected by researchers from over 6000 sample models to form the Princeton University Library (PSB Library). The PSB contains 1814 sample models, and the two architectural 3D models in the library are shown in Figure 4.

The purpose of the 3D model standard library is to provide a reference evaluation system for clustering, classification, identification, and retrieval of 3D models. This benchmark library (PSB), for each 3D model, contains a detailed file (of the type off format with polygons) with a text file containing the original data information and a file with a jpg image containing a thumbnail of the 3D model. The 3D model standard library is a standardised and uniform format for normalising 3D models into off format (a very common polygon format at the University of Minnesota Centre for Advanced Mathematical Geometry). When the 3D model is

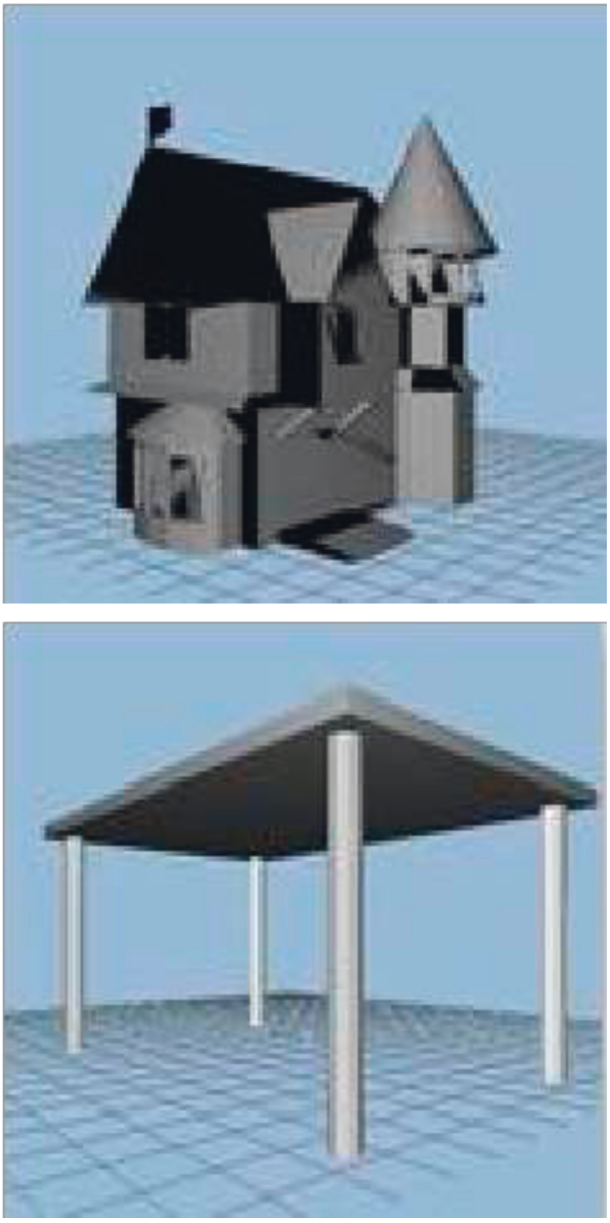


Figure 4: 3D model of a building.

converted uniformly, the scene information, textures, and color information about the 3D model is eliminated and simply stored as a set of vertices and polygons in a surface index set, with the core of the process being the geometric surface information. Saving the file in off format is very simple, the first line is prominently marked "OFF," and the second line clearly indicates the number of vertices, facets, and edges. The number of edges can usually be ignored automatically. The vertices are listed in x, y, and z coordinates, one per row. The facepiece information follows the list of vertices (the number of vertices is specified), which is distributed one per row, and finally the index list of vertices.

2.4. Deep Neural Networks. The important idea of DNN is that it has more "depth" than the common neural network models, i.e. a deep neural network contains multiple hidden layers, rather than just "input layer-hidden layer-output layer." The main types of deep neural networks are deep belief network and restrict Boltzmann machine and convolutional neural network. In this paper, we learn a new DNN by modifying its parameters and apply it to traffic sign recognition to observe the effect of different network structures on the recognition effect.

CNN is a new type of network based on the unique structure of a neuron that can reduce the complexity of feedback neural networks identified by researchers when observing the cat's brain cortex. CNNs are widely used in the field of pattern classification because they can avoid explicit feature extraction from images and can be trained directly on the original input image.

The network structure of a DNN alternates between multiple convolutional and subsampling layers, with each layer of neurons connected only to the nodes of the previous layer. As a multilayer feature extractor, the input to the DNN is the original pixel intensity of the image, which is extracted by the convolutional and subsampling layers to form a feature vector, which is then fully connected to the output neuron for classification.

3. Cross Media Building Detection Based on DNN

3.1. Core Ideas for Architectural Images in DNNs. At present, there are different problems with light, occlusion, and deformation when capturing images of buildings in realistic environments. Different preprocessing methods have their own advantages in highlighting different features or removing certain disturbances. In order to effectively avoid errors in the acquisition, effectively utilise the advantages of different preprocessing methods, and improve the recognition effect of the algorithm, it is proposed to use different preprocessing methods on the data set, form different training sets, overcome the disturbing information in the samples from different angles, and finally make a comprehensive evaluation to make the training results more accurate. At the same time, different structures of deep neural networks (different number of feature maps or different sampling methods) are built on the same preprocessed dataset to achieve fast recognition of distorted building images and to ensure the efficiency and correctness of the algorithm.

Before the start of each training session, the training set is preprocessed and input to different deep neural networks, thus forming the multicolumn deep neural network MCDNN. The MCDNN combines the DNNs trained on different samples linearly and then performs fuzzy judgement on the output results to improve the recognition effect.

3.2. Cross-Media Building Detection Based on DNN. For building image processing tasks in cross-media detection, VGGNet and RseNet [19] have excellent feature expression ability, so they are often used to encode visual features of images. VGGNet is an in-depth CNN developed by the

computer vision group of Oxford University. VGGNet explored the relationship between the depth of CNN and its performance and built a deep CNN by repeatedly stacking small convolution cores and pooling layers. In addition, the official website of VGGNet has open-source models trained on ImageNet datasets. Researchers can easily use VGGNet in their own research fields and retrain it. Therefore, VGGNet is often studied to extract the visual features of images, making it almost a standard for image feature extraction.

ResNet (Figure 5) is a deep neural network model proposed by Professor He[20]. One of the main problems solved by ResNet is the degradation of deep neural networks. When the network depth increases, the performance of the network becomes saturated or even decreases. However, the neural network with more upper layers can extract more complex feature patterns, that is, the deeper the network, the better the theoretical performance. He et al.proposed an identity mapping operation [21], which uses direct connection to connect the front and back layers of the network. In this way, the gradient vanishing problem in the network is well solved, and the number of layers of the deep neural network can be greatly increased.

The birth of CNN was initially influenced by the visual cognitive mechanism of biology. In 1962, in the experiment of studying the visual cortex cells of cats, it was found that each visual neuron only processes a small piece of visual image, which is called the receptive field. In 1984, the concept of neocognitron was proposed, and the first CNN model was born.

Although CNN is a neural network improved on the feedforward neural network, it is different from the traditional feedforward neural network in that the neurons between its adjacent layers are not fully connected. A complete CNN structure usually consists of five parts. The first is the input layer, which can directly use the original image as input and then repeatedly perform convolution and pooling operations on multiple convolution layers and pooling layers to abstract the information in the image into features with higher information content and then enter the full connection layer to synthesize the previously extracted features. Finally, with the calculation of softmax, the probability classification is output to complete the classification task.

3.3. Application of Cross-Media Object Detection in Buildings. In various media, construction engineering data are usually displayed in the form of pictures or text. Taking the sample of text construction engineering as an example, its essence is transformed into natural language processing. Query text in natural language is usually of variable length. In order to facilitate subsequent modeling and processing, it is usually necessary to convert the text into a fixed length vector in cross-media target retrieval methods. This method of converting natural language text into digital vectors is also called text feature coding. Traditional text feature coding models include one hot model [22], vector space model [23], and word2vec model [24, 25].

For the image samples of construction projects, with the maturity of the neural network, there are many image

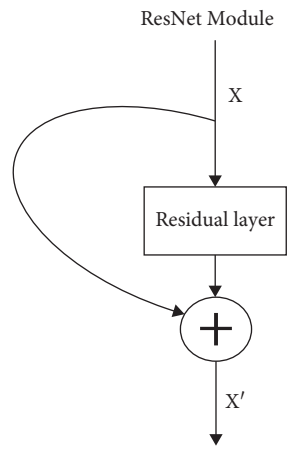


Figure 5: ResNet.

detection methods that can be selected. In general, AlexNet, VGg, and CNN are the mainstream algorithms. In 2012, Hinton et al. A 8-layer AlexNet won the championship in 2012. Compared with LeNet-5, it has a similar overall structure and deeper network layers. Since then, researchers have continuously put forward new improvement and improvement methods, and new CNN structures emerge in endlessly. It includes VGGNet from layer 16 to 19, GoogleNet from layer 22, and RESNET from layer 152. From these CNN structures, we can see that the remarkable feature of the development of the CNN model is that the number of layers is getting deeper and the structure is becoming more complex. The building target recognition based on CNN is shown in Figure 6.

4. Feature Extraction of Building Engineering Based on DNN

4.1. Building Text Feature Extraction. Since the emergence of human civilization, the text has existed in people's daily production and life. As one of the oldest information media, text still has an irreplaceable position in modern life. TF-IDF, LDA, and other models are commonly used for feature extraction of text.

TF-IDF refers to word frequency–reverse file frequency, where term frequency (TF) is word frequency and inverse document frequency (IDF) indicates the inverse text frequency index [26]. TF-IDF is a way to vectorize features and then extract them. It can evaluate the importance of a word or phrase in a document set or corpus. That is, if a word appears more in a document than in other documents, it means that the word can be used as a classification basis to distinguish the document from other documents.

Latent Dirichlet allocation (LDA) [27] is a document topic probability generation model, LDA can detect the potential topic probability distribution in the corpus. It is also a three-level Bayesian probability model. The three structural levels are document, topic, and word.

4.2. Cross Media Image Feature Extraction. As an intuitive and visible media type, image is the most direct source for

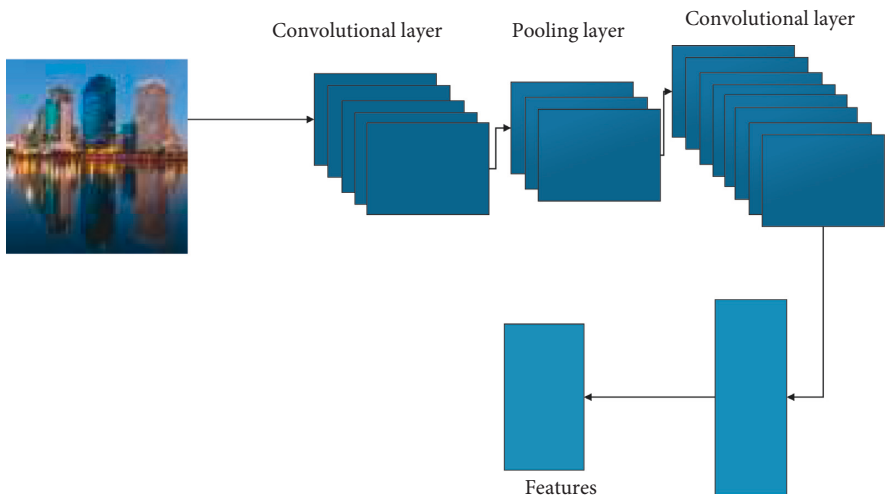


Figure 6: Object detection building with CNN.

people to obtain information from the outside world through their eyes. For the local areas and details of the image, the global feature cannot show its superiority, so we also need to study the local features of the image.

SIFT is a local feature descriptor for images proposed by David Lowers in 1999 [28]. The working principle of SIFT algorithm is to search for some key feature points distributed in the target image. These feature points are "special" in the whole image. They can also remain stable when the brightness of the image changes and the rotation angle and scale change significantly [29]. In short, searching the key feature points of images in different scale spaces is the core of the working principle of SIFT algorithm.

4.3. Application of Cross-Media Feature Extraction in Construction Engineering. The image features of buildings captured by human eyes in natural conditions are mainly area, shape, color, and spatial relationship. We see the height of the building, the floor area, and the texture of the building surface. These features describe the construction project from the overall relationship of the building, so it has the characteristics of intuitive performance and good invariance. Color histogram statistical algorithm, SIFT, and Gabor texture feature extraction algorithm are commonly used, which can realize global feature extraction algorithm. Among them, SIFT is a mature algorithm that can extract local features.

When the construction engineering data and samples are documents, the TF-IDF algorithm is used for processing. Suppose there is a construction engineering corpus D and a topic set t, document set D contains multiple documents D, topic set t contains multiple topic T, and a single document D also contains multiple topic t. The TF-IDF model uses topics as the middle layer of words and documents, rather than directly dealing with the relationship between words and documents, which can more accurately reflect the semantic information of documents.

When the construction samples are images, SIFT algorithm is used to extract features. That is to search the key points in the building image, such as segmentation line, unique shape, and appearance, these feature points are unique in the whole building. Therefore, it can remain stable when the brightness of the building image changes, and the rotation angle and scale change significantly. In other words, the key to SIFT feature extraction is to search the key feature points of building structures in different scale spaces.

5. Conclusion

Based on the background of construction engineering and the demand for cross-media, this paper introduces the application principle of multiclass deep network in cross-media detection from the perspective of deep learning, including VGg algorithm, ResNet, and CNN. When extracting building structure and building features, if the sample is a document, TF-IDF or LDA algorithm can be given priority. If the building sample is an image, CNN has the best processing effect. This paper has made a certain summary and application description on how to effectively identify and extract construction engineering, which has reference significance. In the future, we will collect enough building samples for verification experiments to deeply analyze the advantages and disadvantages of each depth neural network.

Data Availability

The dataset used in this paper is available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest regarding this work.

References

[1] J. Qi, Y. Peng, and Y. Yuan, "Cross-media Multi-Level Alignment with Relation Attention Network," 2018, https://arXivpreprintarXiv:1804.09539.

[2] H. Yu, J. Tao, C. Qin et al., "A novel constrained dense convolutional autoencoder and DNN-based semi-supervised method for shield machine tunnel geological formation recognition," *Mechanical Systems and Signal Processing*, vol. 165, Article ID 108353, 2022.

[3] B. Vidgen and T. Yasseri, "Detecting weak and strong Islamophobic hate speech on social media," *Journal of Information Technology & Politics*, vol. 17, no. 1, pp. 66–78, 2020.

[4] A. Li, J. Du, F. Kou et al., "Scientific and Technological Information Oriented Semantics-Adversarial and Media-adversarial Cross-media Retrieval," 2022, https://arXiv%20preprint%20arXiv:2203.08615.

[5] C. H. Wei, A. Allot, R. Leaman, and Z. Lu, "PubTator central: automated concept annotation for biomedical full text articles," *Nucleic Acids Research*, vol. 47, no. W1, pp. W587–W593, 2019.

[6] A. Latif, A. Rasheed, U. Sajid et al., "Content-based image retrieval and feature extraction: a comprehensive review," *Mathematical Problems in Engineering*, vol. 2019, Article ID 9658350, 21 pages, 2019.

[7] A. M. Oncescu, A. Koepke, J. F. Henriques, Z. Akata, and S. Albanie, "Audio Retrieval with Natural Language Queries," 2021, https://arXiv%20preprint%20arXiv:2105.02192.

[8] X. Zhou, Y. Zhang, H. Guo, and B. Jiang, "Towards bridging the structure gap in heterogeneous catalysis: the impact of defects in dissociative chemisorption of methane on Ir surfaces," *Physical Chemistry Chemical Physics*, vol. 23, no. 7, pp. 4376–4385, 2021.

[9] A. Haldorai and A. Ramu, "Canonical correlation analysis based hyper basis feedforward neural network classification for urban sustainability," *Neural Processing Letters*, vol. 53, no. 4, pp. 2385–2401, 2021.

[10] W. Shuai and J. Li, "Few-shot learning with collateral location coding and single-key global spatial attention for medical image classification," *Electronics*, vol. 11, no. 9, p. 1510, 2022.

[11] E. Agapaki and I. Brilakis, "Geometric Digital Twinning of Industrial Facilities: Retrieval of Industrial Shapes," 2022, https://arXiv%20preprint%20arXiv:2202.04834.

[12] S. Bai, Z. Zheng, X. Wang et al., "Connecting language and vision for natural language-based vehicle retrieval," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 4034–4043, Danvers, MA, June 2021.

[13] Q. Feng, V. Ablavsky, and S. Sclaroff, "Cityflow-nl: Tracking and Retrieval of Vehicles at City Scale by Natural Language Descriptions," 2021, https://arXiv%20preprint%20arXiv:2101.04741.

[14] C. Xiang, "Cross media target retrieval based on deep learning," Master's Thesis, Hangzhou University of Electronic Science and Technology, Chengdu City, 2019.

[15] M. Javeed, M. Gochoo, A. Jalal, and K. Kim, "HF-SPHR: hybrid features for sustainable physical healthcare pattern recognition using deep belief networks," *Sustainability*, vol. 13, no. 4, p. 1699, 2021.

[16] M. Wang and W. Deng, "Deep face recognition: a survey," *Neurocomputing*, vol. 429, pp. 215–244, 2021.

[17] H. Tan, X. Liu, B. Yin, and X. Li, "Cross-modal semantic matching generative adversarial networks for text-to-image synthesis," *IEEE Transactions on Multimedia*, vol. 24, pp. 832–845, 2022.

[18] E. Prasetyo, N. Suciati, and C. Fatichah, "Multi-level Residual Network VGG for Fish Species Classification Journal of King Saud University-Computer and Information Sciences," vol. 34, no. 8, pp. 5286–5295, 2021.

[19] R. Wightman, H. Touvron, and H. Jégou, "ResNet Strikes Back: An Improved Training Procedure in Timm," 10.48550/arXiv.2110.00476 https://arxiv.org/abs/2110.00476, 2021.

[20] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 770–778, 2016.

[21] K. He, X. Zhang, S. Ren, and J. Sun, "Identity mappings in deep residual networks," in *European conference on computer vision*, pp. 630–645, Springer, October 2016.

[22] J. Buckman, A. Roy, C. Raffel, and I. Goodfellow, "Thermometer encoding: one hot way to resist adversarial examples," in *Proceedings of the International Conference on Learning Representations*, Mountain View CA, 2018 February.

[23] G. Salton, A. Wong, and C. S. Yang, "A vector space model for automatic indexing," *Communications of the ACM*, vol. 18, no. 11, pp. 613–620, 1975.

[24] B. Jang, I. Kim, and J. W. Kim, "Word2vec convolutional neural networks for classification of news articles and tweets," *PLoS One*, vol. 14, no. 8, Article ID e0220976, 2019.

[25] K. Fukushima, "Neocognitron: a hierarchical neural network capable of visual pattern recognition," *Neural Networks*, vol. 1, no. 2, pp. 119–130, 1988.

[26] J. Ramos, "Using TF-IDF to determine word relevance in document queries," *Proceedings of the first instructional conference on machine learning*, vol. 242, no. 1, pp. 29–48, 2003, December.

[27] D. M. Blei, A. Y. Ng, and M. I. Jordan, "Latent dirichlet allocation," *Journal of Machine Learning Research*, vol. 3, pp. 993–1022, 2003.

[28] S. Gupta, K. Thakur, and M. Kumar, "2D-human face recognition using SIFT and SURF descriptors of face's feature regions," *The Visual Computer*, vol. 37, no. 3, pp. 447–456, 2021.

[29] H. Bay, T. Tuytelaars, and L. V. Gool, "Surf: speeded up robust features," in *Proceedings of the European Conference on Computer Vision*, pp. 404–417, Berlin, Heidelberg, 2006 May.

Hindawi
Computational Intelligence and Neuroscience
Volume 2022, Article ID 9346771, 10 pages
https://doi.org/10.1155/2022/9346771

Research Article

The Usage of Designing the Urban Sculpture Scene Based on Edge Computing

Junru Zhu

Academy of Arts, Sichuan University of Science & Engineering, Zigong, China

Correspondence should be addressed to Junru Zhu; zhujunru@suse.edu.cn

Received 12 July 2022; Revised 16 August 2022; Accepted 30 August 2022; Published 14 September 2022

Academic Editor: Ning Cao

To not only achieve the goal of urban cultural construction but also save the cost of urban sculpture space design, EC (edge computing) is combined with urban sculpture space design and planning first. Then it briefly discusses the service category, system architecture, advantages, and characteristics of urban sculpture, as well as the key points and difficulties of its construction, and the layered architecture of EC for urban sculpture spaces is proposed. Secondly, the cloud edge combination technology is adopted, and the urban sculpture is used as a specific function of the edge system node to conduct an in-depth analysis to build an urban sculpture safety supervision system architecture platform. Finally, the actual energy required for implementation is predicted and evaluated, the specific monitoring system coverage is set up, and some equations are made for calculating the energy consumption of the monitored machines according to the number of devices and route planning required by the urban sculpture safety supervision system. An optimization algorithm for energy consumption is proposed based on reinforcement learning and compared with the three control groups. The results show that when the seven monitoring devices cover detection points less than 800, the required energy consumption increases linearly. When the detection devices cover more than 800 detection points, the required energy consumption is stable and varies from 10000 to 12000; that is, when the number of monitoring devices is 7, the optimal number of monitoring points is about 800. When the number of detection points is fixed, increasing the number of monitoring devices in a small range can reduce the total energy consumption. The optimization algorithm based on the reinforcement learning proposal can obtain an approximate optimal solution. The research results show that the combination of edge computing and urban sculpture can expand the function of urban sculpture and make it serve people better.

1. Introduction

Urban sculptures are usually independent sculptures in public places such as city streets, squares, railway stations, docks, airports, stadiums, public green spaces, parks, and residential areas. With the development of modern cities, urban environments, and public environments, art has become the focus of urban planning. Educational culture, public environmental art, and urban environmental quality are some of the most important factors. The production of urban sculpture is determined by public artists and decision-makers based on specific environmental factors. In a sense, it is an urban carrier based on the local architectural style and its unique functions. It cannot be replaced by other cultural and artistic forms, especially in cultural and artistic forms that reflect the city's historical and cultural heritage, urban characteristics, and characteristic cultural art form, etc. Urban public art is the local art culture and public space art [1]. In recent years, with the rapid development of higher education in sculpture in China, more and more colleges have offered sculpture-related majors. However, judging from the current urban construction process in China, the sculpture majors offered by existing colleges still cannot meet the needs of urban development, and the quality of sculpture education needs to be improved [2]. It is precise because other professional technology in various aspects has also been applied to this field, such as biotechnology, material science, and technology [3–5].

There are two kinds of definitions approved by the existing research on EC (edge computing). One is to reduce

the cost of service delivery and network operation and to provide edge intelligent services in the storage and application of the network edge near data sources. The other is to execute calculations on the edge of the network. The operations of EC need the downlink data from cloud services and the uplink data from the server Internet. And the edge of EC refers to any calculations and network resources between the conversion calculation center paths from data sources to cloud calculations, and it is a continuous system. The emergence of fog computing and EC enlarges the application of cloud computing to pervasive, that is, computationally intensive research [6]. The combination of EC and deep learning based on AI (artificial intelligence) can reduce network risks and protect network safety. Radanliev et al. (2020) studied the relationship between AI and IoT (Internet of Things) and proposed a new dynamic network risk analysis prototype with the edge support of AI [7].

Nowadays, EC and computer technology tend to be applied to urban sculpture and make the style of urban sculpture more diverse and its theme richer. The development of digital technology provides many new technologies and methods for the creation of urban sculpture. Due to the rapid development of communication technology, it is easier for people to acquire cultural specific artistic knowledge, which promotes the internationalization of urban sculpture, making the language and function of urban sculpture more abundant, and changing the manifestation from static to dynamic. This study combines urban sculpture, edge computing (EC), and video structure technology to propose the security monitoring system of urban sculpture scenes and focus on the new forms of urban sculpture in function and service.

Based on the abovementioned analysis, this study combines urban sculpture with edge technology and proposes a safety monitoring system for urban sculpture scenes based on EC and video structure technology. The main contents are as follows: (1). the basic concepts and operation models of EC and the video-structured monitoring system are analyzed, and the concept of edge processor is used to improve the excessive and complex monitoring elements and provide links to the existing security monitoring system of urban sculpture scenes; (2). because of the dispersion and functionality of urban sculpture, a monitoring method combining UAV (unmanned aerial vehicles) with urban sculpture is proposed to improve the performance of the safety monitoring system of urban sculpture; and (3). the model of the safety monitoring system of urban sculpture is constructed, and the experimental data are analyzed to provide a theoretical reference for optimizing the safety monitoring system of urban sculpture. The innovation lies in: (1) the data transmission distance is shortened using EC, and the speed and accuracy of the information in transmission are improved; (2) the energy consumption is calculated and the optimal path is obtained by modeling the path from urban sculpture to UAV under EC; and (3) simulation experiments are carried out in tourist attractions to test the practical function of the system and verify its feasibility.

2. Analysis of the Theory and Method of EC Applied to Urban Sculpture

2.1. The Relationship between EC and Cloud Computing. EC cannot be applied individually, and it needs the support of various information technologies and technical systems. With the support of IoT, AI, and other technologies, the function and application scope of EC have greatly developed. It is often used to obtain real-time traffic data in cities and metropolitan areas and their surroundings. EC added to the edge application system, the edge storage system, and a series of framework algorithms for application programs and management to meet the needs of different service subjects. These programs and framework algorithms are often combined with blockchain technology to provide safety services for the network [8]. In a sense, it can be regarded as a safety technology that leads to the development of the EC network system. EC can also be combined with cutting-edge technologies. For example, foreign scholars integrate EC and 5G networks to do research, and some scholars apply EC to the UAV, and they have made outstanding achievements in the research [9, 10].

By comparison, EC and cloud computing are mutually complementary [11]. Literally, although cloud computing is powerful, its speed and efficiency of data processing will be subject to constraints and restrictions on different levels. In other words, the energy of cloud computing is limited. Only a certain amount of information can be processed within a certain time limit. However, EC is generated to screen information. Simple information with specific rules is processed at the edge of the network instead of being uploaded to the cloud. In this way, the workload of the cloud can be greatly reduced and work efficiency can be improved. This is the meaning of EC [12].

Common EC and cloud computing application systems are "center + edge" and "no center." They are adapted to different occasions. Based on its flexibility, the former is mostly used in large systems such as enterprise device management, while the latter is mostly used for simple occasions such as data collection and processing. The combination of the two can bring greater advantages into play, improve efficiency, and guarantee service quality [13].

2.2. Application Analysis of EC in Urban Sculpture. EC is powerful and adaptable. It has a wide range of applications in different fields. The application in the urban sculpture industry can also be based on the actual situation, and the core nodes can be flexibly set according to the needs (for example, the data center is the core node, the network device is the core node, the central sculpture is the core node, the platform is the central node, and the system is the core node) [14]. The urban sculpture is considered to have three different application modes at various nodes in cities and industries. They are the urban sculpture as the component node of EC of the urban building system, the urban sculpture as the component node of EC of the urban cultural characteristic dissemination system, and the urban sculpture as the component node of EC of the public service.

When the urban sculpture is used as a component node of EC of the urban sculpture system, the existing urban sculpture resources, and service resources need to be integrated to build the edge urban sculpture system platform [15], and the urban sculptures in a certain area are unified and synchronized management.

City sculptures as the city edge characteristic culture dissemination system calculation, the compositions of the node are the same EC applied to the construction of city characteristic culture when the system, mainly the city cultural buildings, monuments, museums, art museum, the children's palace, data center, financial institutions, knowledge, or data service enterprises such as all kinds of resources integration, formed through the calculation of edge knowledge dissemination system, for the city's cultural construction support [16].

There are four difficulties in the application of urban sculpture. The distribution is a strategic level, capital level, implementation level, and talent reserve level.

At the strategic level, as an emerging technology, EC has powerful functions. But because it is so new, most people are extremely lacking in their understanding of edge technology. The degree of acceptance is even more difficult to talk about. The urban sculpture is a traditional cultural building that has always served the masses. The first problem is how to formulate a reasonable strategic plan to organically integrate urban sculpture and edge technology, to improve the ability of urban sculpture to serve the public with the application of edge technology and ensure that its aesthetics and artistry are not affected, and can be widely accepted by the public [17].

At the financial level, the urban sculpture is located in a superior geographical location, but most of its value comes from the cultural and spiritual values that people attach to it. Such values are established and difficult to change. To make new plans and designs for urban sculptures, whether it is a breakthrough in technical barriers or the use of human resources, new capital injection is required. Therefore, new value points need to be found to bring enough value support to the work [18].

At the implementation level, the fringe transformation of urban sculpture requires a combination of software and hardware. The distribution of urban sculptures is usually numerous and scattered. Moreover, the surrounding environment of the urban sculpture is complicated and the flow of people is large. It takes a lot of time to investigate the situation on the spot during the transformation and planning to ensure that the plan can be implemented.

At the level of talent reserve, after the completion of the fringe city sculpture system, it is necessary to carry out regular inspection and maintenance to ensure its normal work as an edge node. Therefore, a large number of people familiar with sculpture and computer technology are required to enter the reform field and provide service and work, so a large pool of talents in related fields is needed.

2.3. System Construction of EC in the Application of Urban Sculpture. To construct the EC construction system of the urban sculpture industry, the composition and links of the edge construction system need to be clarified. That is, city sculptures can provide what services for people, how these services are converted into data, how these data are classified and stratified, how to classify and layer these data, which data and problems should be uploaded to the cloud computing center, what kind of data should be sink to edge computing for direct processing, and which businesses should be the first to start a certain scale of investment and construction [19]. Based on the analysis of these issues, a frame diagram for the construction of the urban sculpture fringe system is proposed, as shown in Figure 1, which mainly includes the core layer, the core application layer, the EC layer, and the front contact layer.

The core layer is the core of the urban sculpture edge system architecture, which mainly includes the basic data center, the infrastructure center, and the basic component center. The core layer is responsible for the management and evaluation of various data collected in the urban sculpture fringe system [20]. For example, in the entire urban sculpture fringe system, the core layer can collect data such as which sculpture location has the highest traffic volume in a certain period, which location has the highest safety incident rate and find out the corresponding reasons to optimize the system. The core application layer consists of five parts, architecture platform, storage deployment, exception management, resource allocation, and data processing. The five applications correspond to different service directions and play a role in data classification and screening [21, 22]. The EC layer is divided into three modules, namely, service applications, cache management, and edge device management. And it is responsible for the preliminary processing and storage of information collected by edge nodes and providing simple feedback to client programs for preliminary management. This is the key to the entire EC architecture. The last one is the front contact layer, which mainly includes data collection and data pre-processing, which is usually combined with the client to directly provide services to customers.

Based on the completion of the construction of the urban sculpture EC framework, it needs to be combined with reality for further application and deepening, and one of the specific modules safety monitoring is selected. The safety monitoring module combines video surveillance and cloud computing technology and uses EC to build a video-structured analysis system platform for the safety monitoring status of urban sculptures. The frame is shown in Figure 2.

The environment where the urban sculpture is located is complex and the flow of people is large. It is difficult and challenging in terms of safety, management, and monitoring. The figure shows that the fringe city sculpture divides the management elements into five key points, namely, characters, methods, devices, environment, and sculpture materials [23]. Characters include clothing, physical state, and movement state. The device includes factors such as the appearance of the vehicle in the entire environment and the state of motion. The material is the characteristics of the sculpture itself, the size and texture, etc., [24]. The specific scene and location are determined through the identification of the

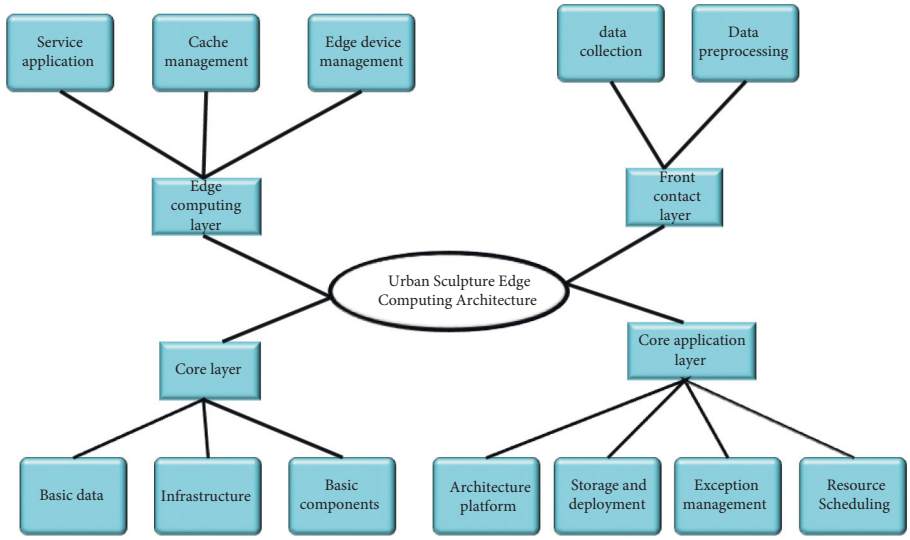


Figure 1: Urban sculpture edge system architecture.

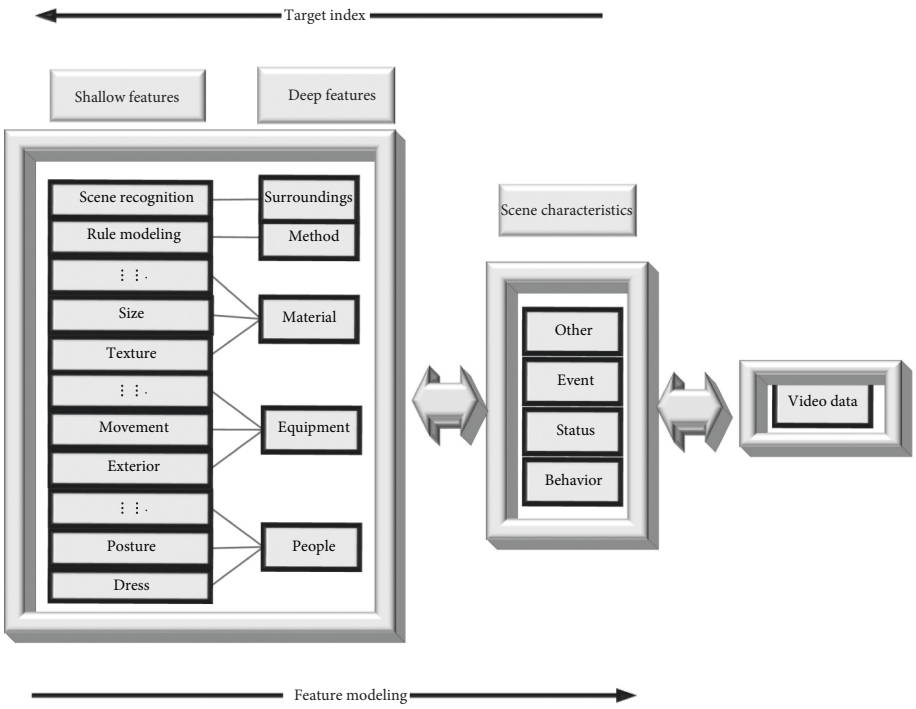


Figure 2: Analysis of the safety monitoring system.

sculpture. The method is the processing method of the EC architecture to deal with the outside world obtained through rule modeling. Environment refers to the cognition and analysis of the environment outside the main body of the sculpture. The purpose of safety monitoring is achieved by grasping and analyzing these five key points [25].

The architecture of the entire system needs to use video structured analysis technology as support, which can be divided into two parts, namely, feature modeling and target recognition [26]. Feature modeling refers to the core platform of the urban sculpture edge system, which uses data collected from specific scenes for feature analysis and matching modeling. To cope with the complex environment near the urban sculpture, the modeling is divided into two levels: the basic level and the advanced level. Feature modeling at the basic level mainly refers to the texture of the sculpture and the geometric characteristics of the external environment. The high-level reference refers to the establishment of a database for dynamic information such as human behavior and movement trajectory. Target recognition is to share these databases, summarize the inherent characteristics of the target, match the information that has existed in the scene, make reminders and reflect, and achieve the purpose of avoiding danger and reducing threats.

In terms of data identification and analysis, it is necessary to make the edge nodes have the function of screening, which can determine which data need to be uploaded to the cloud platform and which data can be directly processed by decentralization. Therefore, the data are given two attributes, namely, the delay sensitivity and the strong semantics. The specific analysis of the system is shown in Figure 3.

Figure 3 shows that the key factors of identification analysis are divided into five indicators for the construction of the database in the edge structure of the urban sculpture. The delay-sensitive task is divided into short periodic events, simple events, and urgent events, and it needs to be treated urgently. It usually refers to the safety accident and important note.

In certain scenarios, the edge nodes of the city sculpture are allowed to automatically contact nearby medical institutions for emergency treatment on the spot. Such events usually require timeliness and are suitable for decentralization to EC nodes; strong semantic tasks refer to tasks that are periodic and complex environments that require long-term analysis to make improvements. For example, tourists evaluate the satisfaction of urban sculptures and make suggestions for improvement [27]. Or due to long-term reasons, the traffic section near a certain city sculpture is blocked, and traffic safety incidents occur frequently. This type of data processing is defined as strong semantic data events and needs to be uploaded to the cloud platform to make decisions and analyses.

Strong semantic data will be uploaded, and time-sensitive data will be decentralized. Based on this, the structured analysis platform of urban sculpture cloud edge collaborative video is constructed as shown in Figure 4.

The platform includes four different working modules: data access, intelligent AI platform, container mirroring service, and intelligent edge platform. The data access platform is responsible for classifying and processing the data sent from the edge side. The intelligent AI platform performs big data analysis on the data processed by the data access platform, performs calculations and modeling, and builds a historical database [28]. The container mirroring platform is responsible for backing up various data and performing functional analysis. The intelligent edge platform is responsible for the management, control, and adjustment of all edge nodes to ensure the healthy operation of edge nodes [29].

The last link is the actual contact with tourists. The behavior status of tourists needs a series of monitoring to ensure that there are sufficient data sources and event basis, and it can also provide better services to tourists. The specific architecture is shown in Figure 5.

Various types of vision sensors, such as infrared and thermal sensors, need to be fully utilized and combined with the behavior of tourists near the city sculptures to model their operating conditions across modal characteristics, such as shape, size, and thermal imaging distribution, to perform correlation analysis and adaptive weighted fusion of cross-modal features. A knowledge base of tourist behavior information data based on historical information is constructed to analyze the temporal and spatial topological relationship through feature mapping [30]. Finally, metric learning is used to query and compare the tourist information data knowledge base to realize the joint analysis and comprehensive evaluation of tourist behavior.

2.4. Optimal Path Planning Model and Simulation Evaluation under Edge Technology. According to the framework, a deployed mobile EC scene is set up, in which a series of wireless access points, mobile edge micro clouds, and wireless charging points are arranged, and a mathematical model is established to evaluate the structure [31].

Due to the complexity of optimal three-dimensional detection path planning, a heterogeneous detection path planning algorithm is proposed based on reinforcement learning. The equation is as follows:

$$F(\mathrm{x},\mathrm{y}) = \mathrm{F}(\mathrm{x},\mathrm{y}) + \mathrm{a}\left(\mathrm{m} + n\max_{y1}\mathrm{F}(\mathrm{x1},\mathrm{y1}) - \mathrm{F}(\mathrm{x},\mathrm{y})\right). \quad (1)$$

In (1), x, y is the current state and action of the detection device, x1, y1 is the state of the detection device in the next second, m is a reward for completing an action when the monitoring device is in condition y, x is the reward when the action is completed, a is the learning rate of reinforcement learning, n is the recession rate of m. The calculation process of the algorithm is shown in Figure 6.

The energy consumption is calculated by three mathematical modeling algorithms, namely, optimal, greedy, and depth-first search (DFS). They are the members of the control groups, and each of them has its advantages. The relationships between the total energy consumption and the number of device nodes, as well as the number of detection devices, can be obtained by different algorithms. First, the principle of the optimal algorithm is to use IBM ILOG CPLEX to solve the heterogeneous path planning problem realized by AMPL language, and find the optimal solution of the equation by branch and pruning methods. The algorithm helps to obtain the optimal solution accurately and easily, but it also has the disadvantage of long solving time, which is difficult to operate in practical applications and can only be used as a theoretical standard. Second, the greedy algorithm is always the best choice to get the optimal solution at present. That is to say, the algorithm obtains the local optimal solution in a sense, and cannot obtain the overall optimal solution for all problems. The key to getting the optimal solution lies in selecting the appropriate greedy strategy.

The greedy algorithm is a simpler and faster technique for the optimal solution of some problems. It is often solved problems step by step according to the current situation, and makes the best choice from the optimization measures and the best solutions, without considering all possible global situations. This algorithm overcomes the limitation of the time required to make all possible hypotheses, improves efficiency, and reduces the workload. It uses the iterative top-down method to carry out continuous greedy selection. When the selection is made, the original problem is changed into a subproblem. The best solution in the local sense can be achieved by selection. Although the optimal local solution

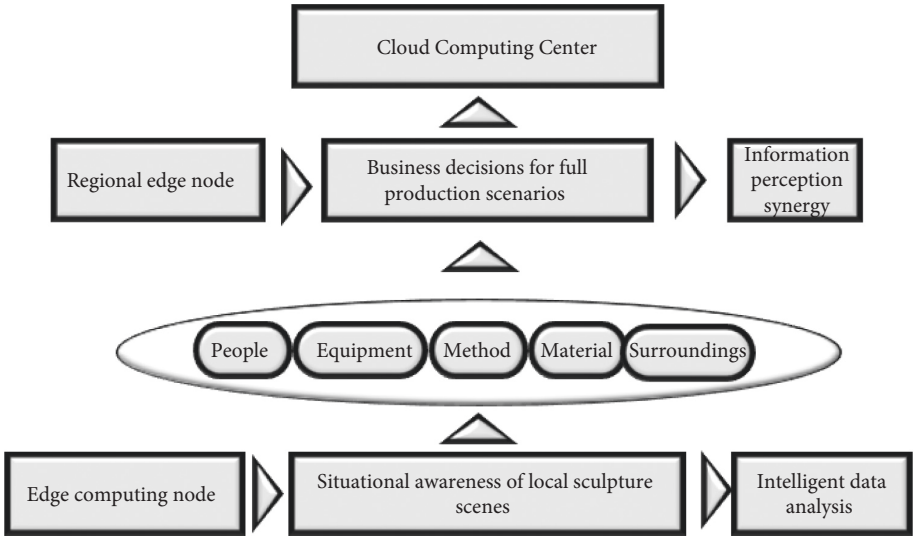


Figure 3: Cloud edge collaborative architecture.

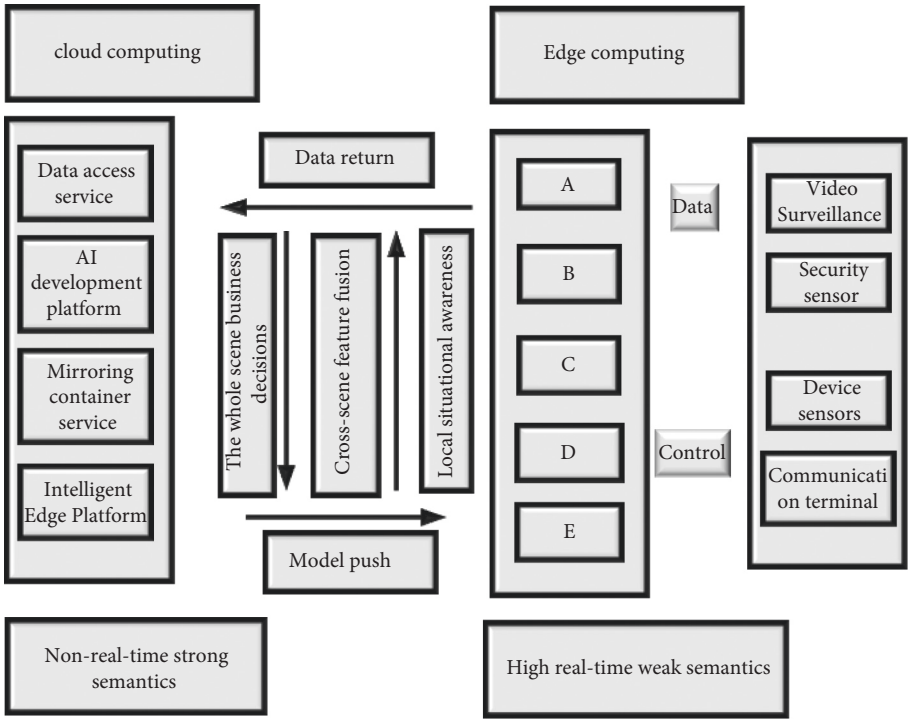


Figure 4: Urban sculpture edge supervision platform (A-E in the figure represent target detection, target re-recognition, behavior recognition, semantic segmentation, scene recognition).

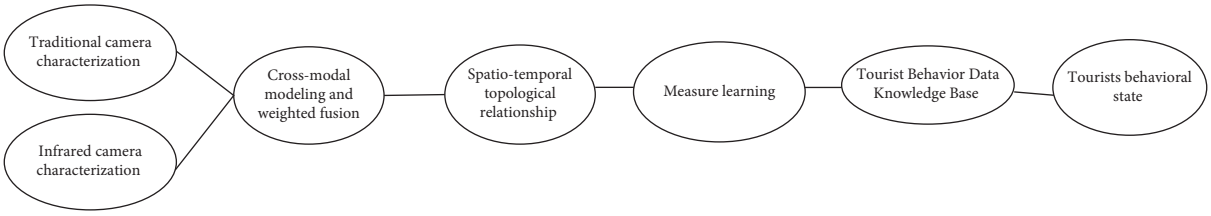


Figure 5: Video joint control and analysis technology principle.

can be obtained after selection, it cannot be ensured that the final optimal solution is global. Each detection device accesses the nearest uncovered area each time until all areas are detected. The greedy algorithm is a simpler and faster technique for some optimal problems. It is characterized by selection step by step, which is often based on the current situation, and makes the optimal choice according to an optimization measure, without considering the overall situation, thus saving a lot of time spent to find the optimal solution from all possible situations. It adopts the top-down method to make successive greedy choices iteratively. Each greedy choice simplifies the problem into a subproblem. Through each choice, an optimal solution to the problem can be obtained. Although it is necessary to ensure that the local optimal solution can be obtained in each step, the overall solution is not necessarily optimal. Here, the greedy algorithm is used to make each detection device access the uncovered area closest to it every time until all areas are detected.

The DFS algorithm is another control algorithm used. The basic idea of DFS is: (1) the vertex v is accessed; (2) the depth-first traversal of the graph is performed from the uncovered adjacent point v until the vertex in the graph with path v is accessed; (3) if a vertex in the graph is not accessed at this time, the depth-first traversal of the graph is performed from an unaccessed vertex until all vertices in the graph are accessed. When it is applied, all regions are connected into a path in the order of depth-first traversal, and then the path is evenly divided into multiple lines. Each detection device is responsible for the detection of one path. In the $1000^* * 1000$ map, 100~1000 detection points, 50 mobile access points, 50 charging devices, and 5~10 detection devices are randomly set to generate the starting position of the detection device. In addition, different initial power, moving speed, and power consumption rate under different states are set on each device. The parameters are adjusted through two experiments to detect the impact of the number change of the covered areas and detection devices on the energy consumption of the system. Here, a scenic spot is selected to conduct the study. According to the behavior of tourists and the safety accidents in the scenic spot, the statistical results are obtained to verify the feasibility of this study.

3. Analysis of Monitoring Energy Consumption Assessment Results

3.1. Detect the Relationship between the Number of Nodes and Energy Consumption. Figure 6 shows the change in the total energy consumption of the system when the number of detection nodes that need to be monitored increases from 100 to 1,000. In this set of experiments, the number of detection devices is set to 7.

Figure 7 shows that when the coverage target increases, the proposed reinforcement learning algorithm is compared with the other three algorithms for energy consumption. When the number of monitoring nodes in the coverage area reaches 800, the total mobile energy consumption varies from 10,000 to 12,000. This is because when the targets to be detected are denser, the total movement path changes less, and the required energy consumption is correspondingly reduced. In addition, as a benchmark for comparison, optimal can always get the best solution. Secondly, the planning path of the DFS algorithm is better than the greedy algorithm. This shows that the solution of reinforcement learning is closest to the optimal.

3.2. The Relationship between the Number of Detection Devices and Energy Consumption. In addition to the number of detection nodes, another factor that has a greater impact on energy consumption is the number of detection devices. The relationship between the number of detection devices and the required energy consumption is shown in Figure 8 below.

Figure 8 shows the influence of the number of detection devices on the total energy consumption of the system, and the x-coordinate represents the number of detection devices. In this set of experiments, the number of detection points in the total detection coverage area is fixed at 500, and the number of detection devices is between areas (5, 10). Figure 7 shows that when the number of detection devices increases, the total system energy consumption tends to decrease. This is because each detection device only needs to move within a part of the detection area, which reduces the moving distance between the areas.

It is found that reinforcement learning can obtain an approximate optimal solution compared with the other two algorithms. The total energy consumption of the path obtained by the algorithm is the least. In short, the optimal path planning algorithm for monitoring the energy consumption of detection devices based on reinforcement learning proposed can obtain an approximate optimal solution and consumes less energy than the other two algorithms.

3.3. Results of System Monitoring. The intelligent monitoring system of urban sculpture based on EC can provide different service functions to tourists according to their historical behavior. Figure 9 shows the tourists' travel trends from July to December.

Figure 9 shows that the tourists' travel trends in different types of scenic spots are drawn according to the flow at different times in different regions. The data in July show that the weather is hot and tourists tend to travel more watery places, so the flow of tourists around the water is large. As time passes by, their tourism destinations tend to be suburbs and villages, and tourists can plan their trips through instant information provided by edge sculpture nodes. The staff of tourist attractions can also allocate resources reasonably according to the statistical data, which is conducive to improving their work efficiency.

The safety monitoring system of urban sculpture based on EC can monitor and calculate the types of safety accidents in recent months, and the statistical results are shown in Figure 10.

Figure 10 shows that it is an important task to monitor the safety in densely populated areas like the area with urban sculpture. Figure 7 shows the time and type of safety

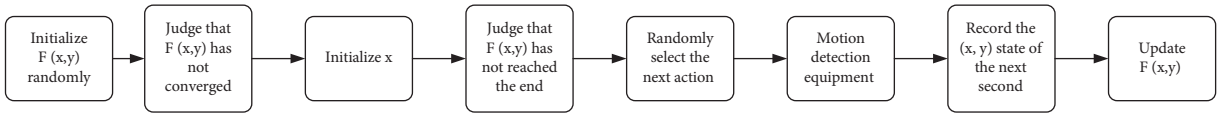


Figure 6: Calculation process of reinforcement learning.

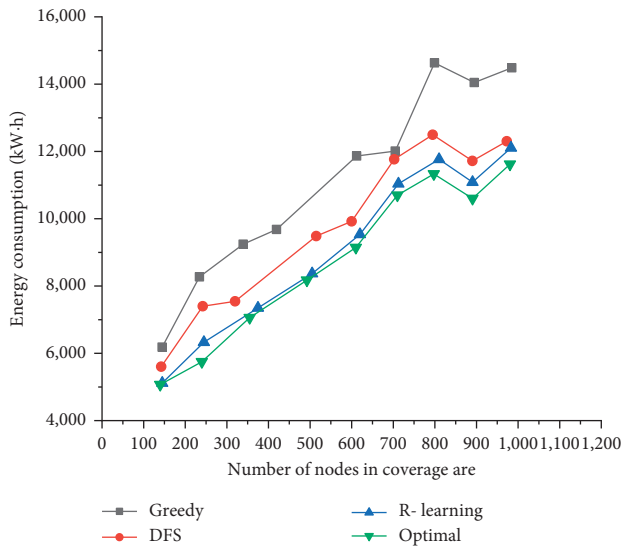


Figure 7: Relationship between the number of detection nodes and energy consumption.

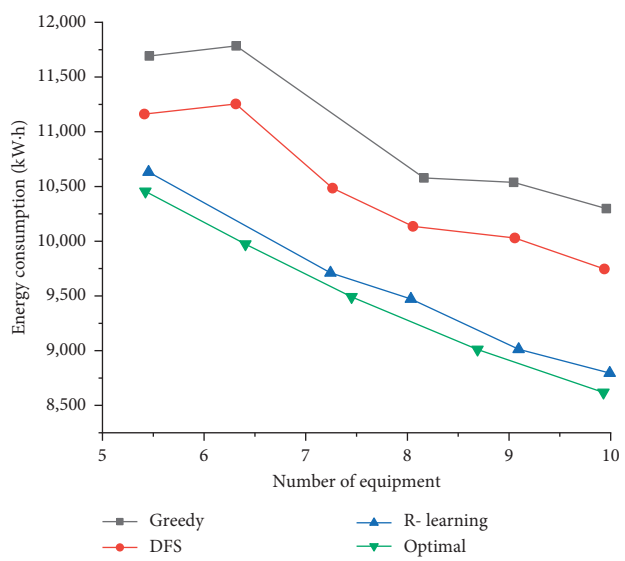


Figure 8: Relationship of energy consumption and the number of detection devices.

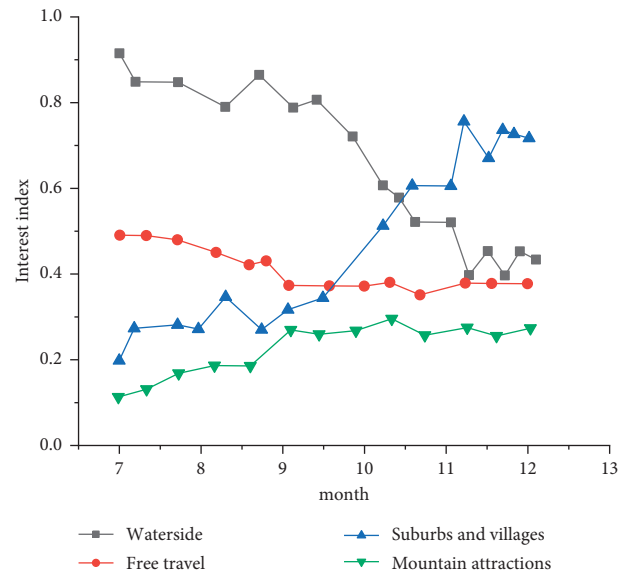


Figure 9: Tourists' travel trends in different types of scenic spots.

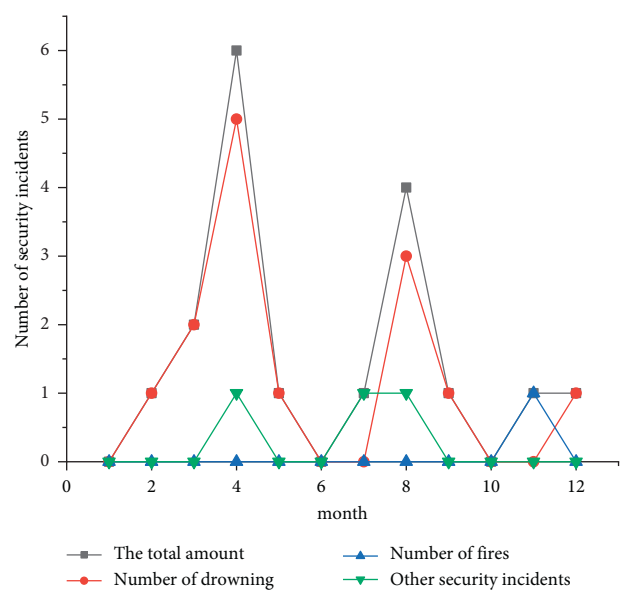


Figure 10: Statistics of safety accidents.

accidents in the tourist attraction where an urban sculpture is located last year. It is found that April and August are the high occurrences of accidents in scenic spot, and the main safety accident is drowning. In addition, November is a month in which fire disasters usually happen. The information is combined with the intelligent monitor system of urban sculpture based on EC can figure out safety and emergency measures in scenic spots.

4. Conclusion

The relevant functions and deployment methods of urban sculpture design and planning are constructed from the theoretical and technical levels and combined with the characteristics of the modern urban sculpture industry under the background of EC. Urban sculpture and EC are used to construct an edge system for modern urban sculpture. The general

technical route of EC applied to the urban sculpture system is listed, the platform architecture combining urban sculpture and EC is given, and the video joint control analysis technology principle based on the algorithm path planning of optimal, greedy, and DFS and the results of energy consumption analysis is also given. The results of energy consumption analysis have strong practicality. It also provides ideas and references for the application of similar urban architecture and EC. The results show that the intelligent safety monitoring system proposed based on EC can expand the functions of urban sculpture, provide more convenient services for the people, and make the functions of urban sculpture more intelligent and diversified. The relationship between the energy consumption of monitoring devices and the number of the edge node is studied by a simulation experiment. Through the comparison of three algorithms, it is found that as the number of the monitoring nodes covered increases, the energy consumption of the monitoring nodes will gradually increase. The optimal number of detection nodes should be maintained at about 800. The proposed reinforcement algorithm can obtain the approximate optimal solution. The simulation results show that the intelligent monitoring system of urban sculpture based on EC can improve the practical value of urban sculpture and provide better services for people.

Since multiple fields are involved, the combination of urban sculpture and EC is still in the process of exploration, and the research stays at the theoretical framework and technical level of system construction. Although the simulation experiment was carried out, further research is still needed due to the constraints of research funding and experience. In the follow-up research, more environmental factors will be considered to make the research better and applicable in real life.

Data Availability

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethical Approval

This article does not contain any studies with human participants or animals performed by any of the authors. Informed consent was obtained from all individual participants included in the study.

Conflicts of Interest

All authors declare that they have no conflicts of interest.

Acknowledgments

The authors acknowledge the help from their university colleagues.

References

[1] A. Tsagkaropoulos, Y. Verginadis, M. Compastié, D. Apostolou, and G. Mentzas, "Extending TOSCA for edge and fog deployment support," *Electronics*, vol. 10, no. 6, p. 737, 2021.

[2] N. J. B. Florita, A. N. M. Senatin, A. M. A. Zabala, and W. M. Tan, "Opportunistic LoRa-based gateways for delay-tolerant sensor data collection in urban settings," *Computer Communications*, vol. 154, pp. 410–432, 2020.

[3] J. Zhang, C. Lu, G. Cheng et al., "A blockchain-based trusted edge platform in edge computing environment," *Sensors*, vol. 21, no. 6, p. 2126, 2021.

[4] T. Le, "Multi-hop routing under short contact in delay tolerant networks," *Computer Communications*, vol. 165, pp. 1–8, 2021.

[5] P. Amos, P. Li, W. Wu, and B. Wang, "Computation efficiency maximization for secure UAV-enabled mobile edge computing networks," *Physical Communication*, vol. 46, p. 101284, 2021.

[6] E. Zhou, J. Zhang, and K. Dai, "Research on task and resource matching mechanism in the edge computing network," *International Core Journal of Engineering*, vol. 6, no. 4, pp. 94–104, 2020.

[7] P. Radanliev, D. De Roure, R. Walton et al., "Artificial intelligence and machine learning in dynamic cyber risk analytics at the edge," *SN Applied Sciences*, vol. 2, no. 11, p. 1773, 2020.

[8] N. Silva, R. C. Pullar, M. E. Pintado, E. Vieira, and P. R. Moreira, "Biotechnology for preventive conservation: development of bionanomaterials for antimicrobial coating of outdoor sculptures," *Studies in Conservation*, vol. 63, no. sup1, pp. 230–233, 2018.

[9] A. Ksentini and P. A. Frangoudis, "Toward slicing-enabled multi-access edge computing in 5G," *IEEE Network*, vol. 34, no. 2, pp. 99–105, 2020.

[10] A. Mukherjee, N. Dey, and D. De, "Edge Drone: QoS aware MQTT middleware for mobile edge computing in opportunistic Internet of Drone Things," *Computer Communications*, vol. 152, pp. 93–108, 2020.

[11] W. Li, S. Wang, and C. Koo, "A real-time optimal control strategy for multi-zone VAV air-conditioning systems adopting a multi-agent based distributed optimization method," *Applied Energy*, vol. 287, p. 116605, 2021.

[12] X. Zhang, Z. Cao, and W. Dong, "Overview of edge computing in the agricultural internet of things: key technologies, applications, challenges," *IEEE Access*, vol. 8, pp. 141748–141761, 2020.

[13] Z. Chen, N. Xiao, and D. Han, "Multilevel task offloading and resource optimization of edge computing networks considering UAV relay and green energy," *Applied Sciences*, vol. 10, no. 7, p. 2592, 2020.

[14] N. Tritschler, A. Dugenske, and T. Kurfess, "An automated edge computing-based condition health monitoring system: with an application on rolling element bearings," *Journal of Manufacturing Science and Engineering*, vol. 143, no. 7, p. 1404, 2021.

[15] J. Li, G. Zhou, T. Tian, and X. Li, "A new cooling strategy for edge computing servers using compact looped heat pipe," *Applied Thermal Engineering*, vol. 187, p. 116599, 2021.

[16] W. Wang, H. Huang, L. Xue, Q. Li, R. Malekian, and Y. Zhang, "Blockchain-assisted handover authentication for intelligent telehealth in multi-server edge computing environment," *Journal of Systems Architecture*, vol. 115, p. 102024, 2021.

[17] K. Peng, J. Nie, N. Kumar et al., "Joint optimization of service chain caching and task offloading in mobile edge computing," *Applied Soft Computing*, vol. 103, p. 107142, 2021.

[18] L. Mohan, D. Farooq, and S. Rajesh, "GEESE: edge computing enabled by UAV," *Pervasive and Mobile Computing*, vol. 72, p. 101340, 2021.

[19] X. Liu and Y. Z. Jiang, "A real-time detection method for abnormal data of internet of things sensors based on mobile edge computing," *Mathematical Problems in Engineering*, vol. 2021, pp. 1–7, 2021.

[20] Y. Fu, X. Yang, P. Yang et al., "Energy-efficient offloading and resource allocation for mobile edge computing enabled mission-critical internet-of-things systems," *EURASIP Journal on Wireless Communications and Networking*, vol. 2021, no. 1, p. 26, 2021.

[21] K. Nithya, "Geographic routing in WSN for measuring coverage constraints and energy consumption in cloud environments," *International Journal of Innovative Technology and Exploring Engineering*, vol. 9, no. 3, pp. 3069–3072, 2020.

[22] N. Balamuralidhar, S. Tilon, and F. Nex, "MultEYE: monitoring system for real-time vehicle detection, tracking and speed estimation from UAV imagery on edge-computing platforms," *Remote Sensing*, vol. 13, no. 4, p. 573, 2021.

[23] N. Qin, B. Li, D. Li, X. Jing, C. Du, and C. Wan, "Resource allocation method based on mobile edge computing in smart grid," *IOP Conference Series: Earth and Environmental Science*, vol. 634, no. 1, pp. 012054–012675, 2021.

[24] H. Zhang, Z. Liu, Y. Zhang et al., "Research on deployment method of edge computing gateway based on microservice architecture," *IOP Conference Series: Earth and Environmental Science*, vol. 675, no. 1, pp. 012164–634, 2021.

[25] C. Liu, X. Su, and C. Li, "Edge computing for data anomaly detection of multi-sensors in underground mining," *Electronics*, vol. 10, no. 3, p. 302, 2021.

[26] R. Pahič, Z. Lončarević, A. Gams, and A. Ude, "Robot skill learning in latent space of a deep autoencoder neural network," *Robotics and Autonomous Systems*, vol. 135, p. 103690, 2021.

[27] G. Zhang and G. H. Peng, "Research on the stabilization effect of continuous self-delayed traffic flux in macro traffic modeling," *Physica A: Statistical Mechanics and Its Applications*, vol. 526, p. 121012, 2019.

[28] M. Cui, H. Zhang, Y. Huang, Z. Xu, and Q. Zhao, "A fountain-coding based cooperative jamming strategy for secure service migration in edge computing," *Wireless Networks*, vol. 1, pp. 1–14, 2021.

[29] S. Guo, K. Zhang, B. Gong, W. He, and X. Qiu, "A delay-sensitive resource allocation algorithm for container cluster in edge computing environment," *Computer Communications*, vol. 170, pp. 144–150, 2021.

[30] n Li, "Construction of landscape architecture art design based on streaming media data processing," *International Journal of Arts and Technology*, vol. 12, no. 4, p. 335, 2020.

[31] K. M. Chowdary and V. Kuppili, "Enhanced clustering and intelligent mobile sink path construction for an efficient data gathering in wireless sensor networks," *Arabian Journal for Science and Engineering*, vol. 46, no. 9, pp. 8329–8344, 2021.

Hindawi
Computational Intelligence and Neuroscience
Volume 2022, Article ID 2727512, 12 pages
https://doi.org/10.1155/2022/2727512

Research Article

The Big Data Model for Urban Road Land Use Planning Is Based on a Neural Network Algorithm

Sunan Tu[1] **and Ming Zhang**[2]

[1]*School of Architecture, Southeast University, Jiangsu, Nanjing 210096, China*
[2]*Keltai Real Estate Development (Nanjing) Co., Ltd., Jiangsu, Nanjing 210005, China*

Correspondence should be addressed to Sunan Tu; sunan@seu.edu.cn

Received 20 July 2022; Accepted 22 August 2022; Published 13 September 2022

Academic Editor: Ning Cao

The spatial differentiation of land use induces traffic demand and guides the construction of traffic supply; traffic conditions are an important influencing factor in determining the nature of land use, and there is a close interaction between the two. This study uses a neural network-based approach at the urban grid level to portray representative phenomena of urban development and analyze the interaction between transportation and land use. The results reflect the model's effective simulation of urban laws, and the case study reveals the differences in the laws of different cities, to guide the benign development of cities and transportation. This article firstly conducts a study on the theoretical foundation; compares the development history, planning, and design methods and practical experience of road planning and resilient planning; summarizes the experience of resilient road system design; and analyzes the future development trend, based on the above basic theoretical research, to develop research ideas and methods. Secondly, the scenario analysis method is explicitly applied to analyze various scenarios that may occur in the future development process of simulated urban roads and rank the scenarios based on the probability of occurrence. For the impact of traffic on land use, the concepts of vitality and potential are introduced, and a multidimensional long and short-term memory network (MDLSTM) model is established. The model takes into account land use lags and potential transfer and has relatively higher prediction accuracy. The results show that larger cities with urban dominant industries and tertiary industries also have higher land use potential and the more significantly influenced by traffic.

1. Introduction

Transportation activities are an important part of modern urban life, and transportation facilities are the yardstick and link to economic and social development and urban growth. With accelerated urbanization, the problem of urban traffic congestion has become increasingly serious. The spatial distribution of land use generates transportation demand, and location and transportation conditions are important influencing factors of urban land differentiation, and modeling and quantifying the mutual influence process between them has been a key topic in the field of sustainable urban development [1]. From a macroscopic point of view, the evolution of urban development is always accompanied by changes in land transactions and transportation patterns. The high concentration of residential neighborhoods and the separation of jobs and residences have profoundly affected the spatial distribution patterns of traffic in many large cities at present, and the monocentric and borderless spatial structure of some cities has led to an unstoppable slide toward long-distance, high-frequency commuting, and car-led traffic patterns. Urban transportation and land use have a close coupling relationship. In the short term, the spatial distribution of urban land use is relatively static, and various areas with different land-use properties complete the exchange of people and products through urban transportation links, that is, urban land use patterns determine the generation and spatial distribution of urban transportation [2]. In the long term, the construction of transportation infrastructure and the supply of transportation

services improve accessibility, and the increased commuting costs of traffic congestion reduce the land value of the area, and this difference in value further adjusts the land development process through the individual choice behavior of urban residents, thus affecting the urban spatial structure [3]. The macroscopic relationship between these has been systematically studied by a series of more mature theories such as accessibility theory, transportation planning theory, and transportation supply-demand balance theory from a general perspective.

The complexity of urban laws and influencing factors also pose a challenge for the identification of traffic and land use integration laws. Cities have a large number of nonlinear, multi-factor simultaneous effects and other complex development and change laws. In parts of time, there is often a lag in the construction of various types of facilities in cities; land use types do not respond in real-time to make changes after traffic conditions change, and it takes time to plan and build transportation facilities after land-use types change, which objectively affects the accuracy of data-driven models [4]. Spatially, different regions, spatial units, land use types, and influencing factors within the city have different spatial interactions, and the spatial correlation of land use and traffic demand is often significantly correlated with distance, while the spatial correlation of traffic supply is directional and often distributed along the road network, bus network, or rail network, which makes it similar to meta-cellular automata, convolutional. This makes it difficult to accurately simulate models like beta cellular automata, convolutional neural networks, and other models with "translation-invariant" properties. In terms of influencing factors, the number or size of most commercial facilities reflects the demand of urban residents for this type of urban service, but in terms of facility quality, the demand for high-level facilities is not easily reflected by static indicators such as floor area and floor area ratio. In this article, we use convolutional neural network-multidimensional long and short-term memory network (CNN-MDLSTM) to fit the basic attributes of the network, road network density, and various types of land use property scores of the parcel as the input and traffic generation attraction as the output, and the prediction accuracy of traffic generation attraction and the influence of various types of land use and indicators are analyzed. The factors of the prediction accuracy of traffic attraction and the influence of each type of land use and each index on the prediction accuracy of traffic attraction were analyzed. Finally, we used the convolutional neural network and cross-long and short-term memory network (CNN-CLSTM) to fit the basic attributes of the network, the scores of various land-use properties and traffic attraction as input, and the road network density, the total number of surrounding bus stops, and traffic facilities as output and analyzed the relative influence of each index on the output.

2. Related Work

The complexity of urban laws and influencing factors also pose a challenge for the identification of traffic and land use integration laws. Cities have a large number of nonlinear, multi-factor simultaneous effects, and other complex development and change laws. In terms of time, there is often a lag in the construction of various facilities in cities; land use types do not respond in real-time to make changes after traffic conditions change; and it takes time to plan and build transportation facilities after land-use types change, which objectively affects the accuracy of data-driven models.

Some scholars have preliminarily studied the theory of intensive land use, put forward the relevant concepts and connotations of intensive land use, and determined the basic contents of intensive land use. Jun [5] put forward reasonable improvement methods from urban land use structure and layout to promote coordinated and sustainable urban development. Qi et al. [6] considered the connotation of land-intensive utilization from three different perspectives, such as the perspective of matching the scale of urban development and the scale of industrial structure, the perspective of the structure and reasonable distribution of the functions of each land use in the city, and the perspective of the input-output requirements of each land use. From different levels of the literature, [7] argues that different cities present different urban conflicts and characteristics, and need to combine different urban development levels and positioning, and puts forward intra-city land rectification programs and measures. Mu et al. [8] construct and evaluate land-use efficiency evaluation systems from macroscopic to microscopic perspectives for commercial, industrial, and residential land use types in municipal districts. Wu et al. [9] propose that the land type of urban low-utility land is urban construction land first, and on this basis, urban low-utility land can be judged by two indicators: building capacity and land output efficiency. Some government departments not only consider the level of land use and input-output efficiency, but also consider whether urban land is low-utility land in terms of land layout, level of supporting facilities, and ecological and environmental protection. The article [10] evaluates the efficiency of industrial land use in terms of land use output and social services based on the micro-plot scale and explores the relationship between land use structure and its efficiency. The literature [11] studied the effect of far-left-turn traffic implementation at signalized intersections and proposed a microscopic traffic model applicable to typical urban areas, and evaluated the left-turn organization of different signal-controlled intersections. The literature [12] described the operation law of mixed traffic flow at signalized intersections by establishing a microscopic traffic simulation model and studied the before-and-after effect of left-turn traffic flow, left-turn vehicle ratio, and dedicated lane length based on the waiting time of vehicles after the stop line. Chen et al. [13] used detectors to collect left-turn traffic flow and delay under various types of phases with constant periods and phases.

The data relating to the errors and the relationship model between the flow and delays were established. The literature [14] investigated the conditions for setting left-turn lanes and their application in practice proposed that left-turn lanes can effectively organize the left-turn traffic flow at planar intersections, and evaluated the safety and efficiency of left-turn lanes. Chen et al. [15] emphasize that the road

network form and streets should become an important public space reorganization linked to the urban design level in the design optimization of the slow walking system, while the spatial forms of streets and road networks are closely related to other spatial design contents and deep social, economic, and environmental contexts. Saralioglu and Gungor [16] studied the development process of the green transportation concept and proposed that, firstly, the construction of horizons on the human sightline should be considered from the scale of people themselves. Secondly, it is necessary to create a traffic information platform using cell phone terminal data. Finally, it is proposed that effective measures should be taken to improve the competitiveness of urban public transportation.

3. Optimization Algorithm of Urban Road Land Use Planning Based on a Neural Network Algorithm

3.1. BP Neural Network Land Allocation Optimization Algorithm Based on Entropy Power Method. Since this article needs to quickly identify low-utility sites in cities for further screening by subsequent planners based on actual conditions and further determination of urban renewal timing and renewal strategies, reducing human subjective involvement is the main way to improve the low-utility site identification process. The complex evolution mechanism of geographic scenes determines that giving physical meanings to indicators requires the introduction of machine learning algorithms with autonomous learning capability and self-adaptive ability, which can be used to mine the data and learn the intrinsic complex connections behind the data, to calculate the indicator weights based on the geographic meanings. Therefore, this article adopts the method of combining the entropy method and machine learning algorithm to objectively reflect the importance of index attribute information in the whole evaluation system by entropy method and then combines with machine learning to correct the weights determined by the entropy method [17]. The entropy weight method is an objective method of calculating weights evolved from the discipline field of the signal system, which is developed by measuring the amount of information in the system, called information entropy, and the higher the information entropy in the system, the more orderly the system represents; on the contrary, the more chaotic the system is. For the evaluation system, the information size of the evaluation index also represents the degree of the orderliness of the system. Spatially, different regions, different spatial units, different land use types, and different influencing factors within the city have different spatial interactions, and the spatial correlation between land use and transportation demand is often significantly correlated with distance. The higher the information size of the evaluation index, the higher the information entropy, which indirectly represents that the index occupies an important position in the system.

The indicator weights determined by the entropy weight method depend on how much attribute information the indicators contain among all indicators, and their weight values rely heavily on the sample data, lacking horizontal comparison among indicators, resulting in the indicator weights contradicting the importance of the actual impact on land-use efficiency, and finally obtaining unreliable low-utility land identification results. Land use is a complex system, and the assessment of land-use efficiency is also a complex nonlinear process. To accurately assess the land-use efficiency situation and identify low-utility land, the learning mechanism of indicator weights should be designed based on the complex system, and the systematic characteristics of the indicator factors should be explored. In this article, the initial weights of evaluation indicators obtained by the entropy weight method are used as input data, and the learning mechanism of the BP neural network is used to reverse calculate the node values of each node in the output layer and the implied layer as well as the corrected amount of node weights, and the neural network is used to identify the unknown connections among indicators and the positive knowledge inference to correct the weights obtained by entropy weight method and reduce the deviation of the weights obtained based on the pure information confusion so that the weights are closer to the actual situation [18]. Based on the initial weights of indicators obtained initially by the entropy weighting method, the land inefficiency identification index is calculated and normalized with the initial land inefficiency identification index as the learning sample, that is, the input layer, which is represented by the matrix as

$$A_{mn} = \begin{pmatrix} a_{11} & \dots & a_{1k} & \dots & a_{1n} \\ \dots & \dots & \dots & \dots & \dots \\ a_{i1} & \dots & a_{ik} & \dots & a_{in} \\ \dots & \dots & \dots & \dots & \dots \\ a_{m1} & \dots & a_{mk} & \dots & a_{mn} \end{pmatrix}. \quad (1)$$

Denote the desired weight by w, the number of nodes in the hidden layer by $p(x)$, and the number of nodes in the hidden layer by δ, where

$$w = \frac{kx + \delta}{p^2(x)} + Cp(x). \quad (2)$$

The vector of weights of the connections between the nodes in the implied layer and the output layer is denoted by (x), where k denotes the connection weight of the kth node. The output layer is the weight of each indicator with a vector $\overrightarrow{O}(x)$, with V denoting the connection weight between the input layer and the implied layer, indicating the connection tightness between the nodes. The transfer function is selected as the Sigmoid function to train the samples.

$$G(x) = V\int \left[kx^2 + p(x)\right]dx + \overrightarrow{O}(x). \quad (3)$$

The specific process and basic steps of the road land weight correction model using the BP neural network in this article are shown in Figure 1.

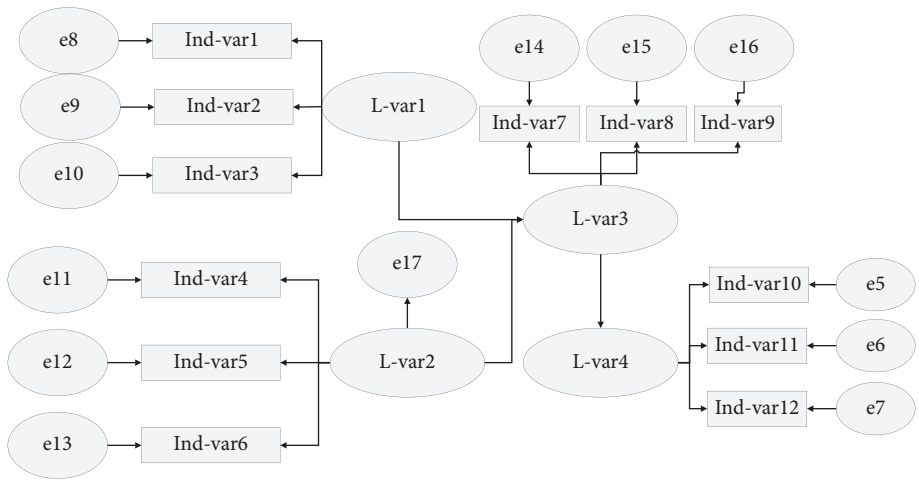


Figure 1: Flow chart of weight correction model based on BP neural network.

The network error for each sample during training and the total cumulative network error is calculated as

$$g_k = T[f_k] = \sum_{i=0}^{k} P \cdot f_i. \tag{4}$$

The model is generally divided into three layers of structure: an input layer, an output layer, and an implicit layer, where the input layer is the initial value of internal and external indices and index standardization obtained by using the initial weights, and the output layer is the initial weight of entropy weight method. For the two types of land use, the indicators of internal properties and external characteristics need to be optimized separately, so four times neural network weight optimization is carried out according to the model structure, that is, the six indicator values and the initial index of internal properties of road use are used as the input layer nodes after standardization of internal properties of road use, and the weights of indicators of internal properties of road use are used as the output layer; the seven indicator values and the initial weights of external characteristics of road use after standardization of external characteristics of road use are used. In the same way, the same training is carried out for all industrial land parcels. In the BP neural network algorithm, the S-type function is generally used to ensure that the activation function is derivable everywhere. The logistic function is suitable for the activation function required to calculate the weights in this article because the output values of the training samples are the weights between 0 and 1, and all the values are guaranteed to be in the interval [0, 1] compared with the general linear function.

3.2. Urban Road Land System Planning Scheme. The establishment of the evaluation index system of the new area road network system requires the selection of suitable evaluation indexes from the technical evaluation indexes of the road network. The evaluation of technical indicators can be done from several aspects such as construction level, the geometric topology of the road network, and connection quality [19]. However, without a large-scale traffic survey, many evaluations may not be analyzed due to the lack of data for the current situation of a new city with almost no traffic demand. Therefore, in this article, the technical indicators of the road network are divided into four levels: road characteristics, access depth, service characteristics, and traffic characteristics, some of which may reflect two or more levels at the same time and there are some duplications. Road characteristics reflect the nature of the planned road function and construction level in the case of clear road traffic planning.

Under can be described from the perspective of road network technical level, per capita road area, road network density, arterial road network density, road network connectivity, road network capacity, and pavement paving rate. For new urban areas generally in the construction process, there is the problem of pavement rate road network is completed and there is no problem of poor pavement rate. Access depth reflects the connectivity and convenience of the road network. In the case of clear traffic demand and planning, road network can be described from the perspective of road network density, road network connectivity, connectivity index, road network accessibility, road network nonlinear coefficient, unit traffic accessibility, unit traffic nonlinear coefficient, etc.

The method of integration degree uses spatial syntax indicators and models to study the spatial morphology of road networks to find the interaction between the spatial morphology of road networks and urban spatial morphology. Integrating the pulse of Changsha's urban growth with the fitting of cab big data, it is proved that the method of aggregation degree can be used to reflect the spatial morphology of the road network and analyze its relationship with the urban spatial morphology [20]. The indicator weights determined by the entropy weight method depend on how much attribute information the indicators contain among all indicators, and their weight values rely heavily on the sample data, lacking horizontal comparison among indicators, resulting in the indicator weights contradicting

the importance of the actual impact on land-use efficiency, and finally obtaining unreliable low-utility land identification results. The spatial sentence method integration index can be used at the general planning level to identify the central tunnels, districts, and box shot axes, provide traffic and morphological basis for general planning zoning, and provide data support for zoning differentiation requirements; at the control planning level, it is mainly used to identify roads with high integration degree, check and optimize the layout of convenient land, and check and optimize the number of road lanes, rail lines, and conventional bus lines; at the revision planning level, it is mainly used to identify roads with high integration, find the preferred path for encrypting the branch road network, and optimize the integration and layout of related buildings, supporting facilities and sites in the project. The "system quantification" assessment can be used to comprehensively assess the supply of the road network carrying capacity (also includes the comprehensive traffic carrying capacity of the road network as a carrier), the traffic demand, and traffic patterns corresponding to the land-use planning, of which the grid method is the supply-side assessment, compared with the previous single-density assessment, there are more density indicators, and indicators reflecting non-homogeneity. The integration degree method reveals the non-homogeneity of the road network in a more systematic way. The saturation method directly assesses the supply and demand, which is the key to recognizing and predicting the traffic supply and demand corresponding to the planning and to plan the traffic supply and demand balance.

Among them, road network density is an important indicator of the quality of urban traffic road network layout can reflect the number and level of urban road network construction reflects the quality of urban road network layout. Service characteristics reflect the road users according to the traffic state from the speed, comfort, convenience, economy, safety, and other aspects of the degree of service received. In the case of clear road traffic planning and road, traffic flow can be described from the perspective of road network service level, road network ratio, and road network load degree uniformity. Among them, the level of service reflects the level of quality of operational services provided under certain traffic conditions road network load uniformity reflects the degree of difference in congestion on each route in the region. Traffic characteristics reflect the performance of traffic flow on the planned road. Based on a clear section of traffic demand, it can be described from the perspective of road network traffic speed, road network traffic density, and road network flow [21]. Among them, road network traffic speed includes simple average traffic speed, flow or distance weighted average traffic speed, and flow distance weighted average traffic speed. It is assumed that accurate prediction of traffic demand and clear road network planning follow the principle of indicators as few as possible and indicators all quantitative. This article extracts 15 indicators from 4 levels of road characteristics, access depth, service characteristics, and traffic characteristics to establish a three-tier indicator system. These indicators can reflect the essential characteristics of the road network to achieve our goal of road optimization. The structure of the evaluation index system of the new area road network system planning scheme is shown in Figure 2. The definition of each indicator in this evaluation system is consistent with the model in the conventional technical evaluation.

Due to the different scales used for each technical indicator, it is necessary to standardize the indicators to calculate the selected indicators. Here, the linear extremum method is used to establish the affiliation function to normalize each indicator. The expressions are as follows for the larger and better indicators.

$$H(x) = \sum_{i-1}^{n} p_i \ln \frac{1}{p_i}. \tag{5}$$

The consistency ratio of the judgment matrix can be obtained by the above method and then determine whether it passes the consistency test. The weight coefficients that pass the test can be used as the weights of each index. In the process of moving from the current situation to the future, various spatial and temporal scales of planning and self-organization come into play at the same time. Generally speaking, the long-term planning, the larger the spatial scale considered, the more it favors the structural order, and the more recent planning, the more it needs to be detailed and precise and operable, and the more it favors the functional order. Structure and function are two sides of the same coin for a specific space, and any construction behavior and planning can find its corresponding spatio-temporal scale and target, and make corresponding macro and micro, near and long-term integration [22]. From the point of view of planning regulation, the larger the spatio-temporal scale is, the larger the planning is in the knotty framework, only the objectives need to be specified for the lower level planning, thus the structural framework and the objectives of the lower level are its focus and the focus of planning assessment, the indicators of structural elements and the spatial layout and implement are the elements to be assessed. Planning at smaller spatial and temporal scales, within the framework of upper level planning, also needs to fully correlate macro-structure and context with relevant details to better play the role of self-organization and better realize spatial functions.

4. Dynamic Balance of Urban Road Land-Use Planning Based on a Neural Network Algorithm

The goal of urban road network planning is to provide a planning blueprint and construction basis for achieving a balance between traffic supply and demand, and the corresponding assessment is whether a balance between traffic supply and demand can be achieved, and the whole process of planning should be a process to achieve a dynamic balance. From the research of the current situation, the formulation of the planning scheme to the decision-making, implementation, feedback, and even adjustment of the planning scheme, a set of scientific and reasonable

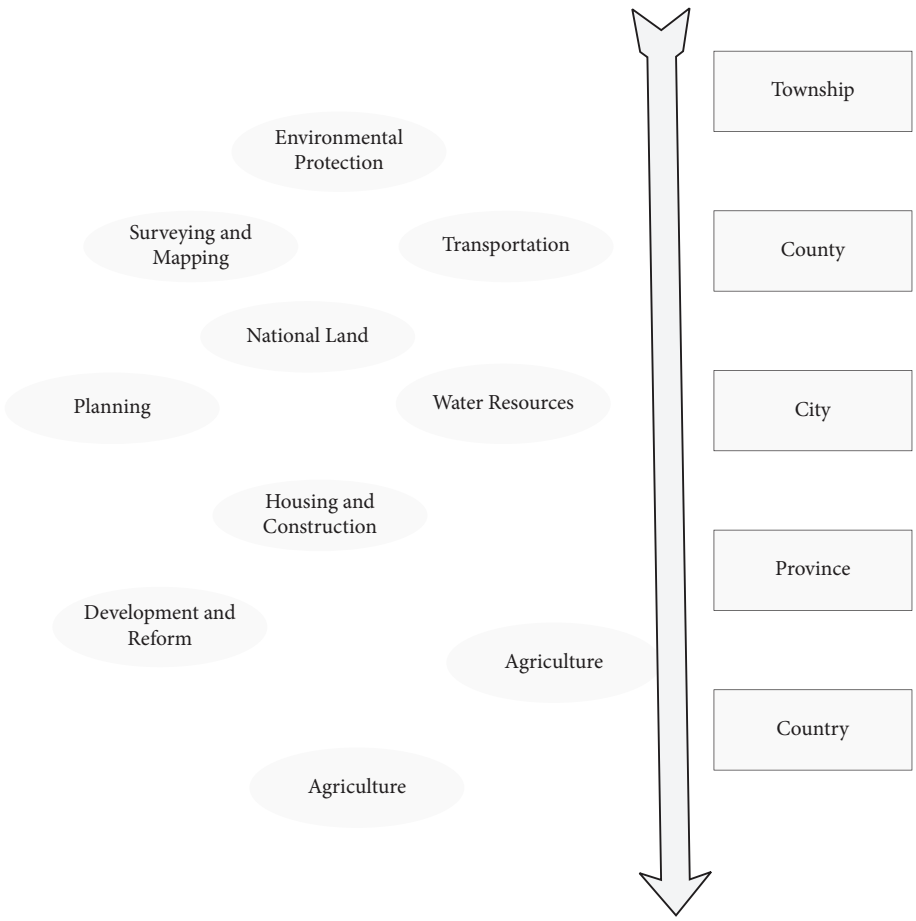


Figure 2: Evaluation index system of road network planning scheme in the new area.

evaluation methods are needed to assess the current situation, the rationality of the planning scheme, and the adaptability of the plan after implementation. From the perspective of demand management, land-use planning should interact with road traffic planning and match each other to ensure land for road traffic and realize its traffic carrying capacity, but also scientifically plan the nature and intensity of land parcels and rationally plan traffic demand. Therefore, the introduction of evaluation in the planning process is the need to improve the planning process, and the planning evaluation flow chart is proposed by drawing on the neural network algorithm of complex systems, as is shown in Figure 3.

In physical planning, the object of planning is mainly physical space, focusing on the study and arrangement of block land space of various natures and network road space made up of a combination of linear space, and there is an interactive relationship between road network planning and the land planning of the areas it serves. From the perspective of traffic supply and demand, the land planning of other spaces corresponds to the corresponding traffic demand, the amount of traffic demand and spatial distribution is due to the nature of the land and intensity of other spaces layout triggered by the supply of the road network that the road network carrying capacity for (including motor vehicle traffic carrying capacity for, dependent on the road network of rail transit and The balance between the volume and spatial distribution of demand (including motor traffic carrying capacity, rail and conventional bus carrying capacity attached to the road network, and slow traffic carrying capacity) depends on the balance of planned land supply. The assessment of the dynamic balance of traffic supply and demand through the above process contributes to the preparation of urban planning and its implementation.

From the viewpoint of supply carrying capacity, the self-organization of adaptive subjects and their elasticity of multi-modal combination of travel provide the basis for the dynamic balance of urban traffic, and the road network planning at all levels of statutory planning should provide a framework for self-organization and provide target requirements and spatial guidance for the planning of public transport line network, especially the rail line network among them. From the point of view of demand management, land-use planning should interact with road traffic planning, matching each other, not only to ensure that road traffic land achieves its traffic carrying capacity, but also scientific planning of land parcel land nature and intensity and rational planning of traffic demand. From the point of view of urban form, the road network form as a whole should match the urban spatial form. From the same circle

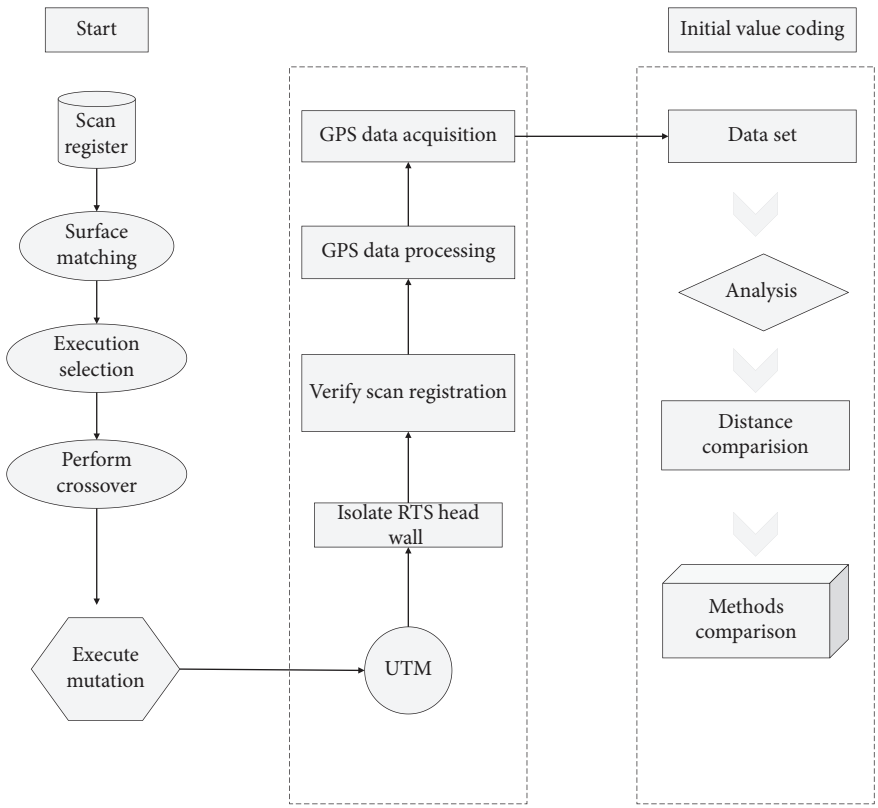


Figure 3: Framework diagram of dynamic planning of urban road land based on a neural network algorithm.

theory, the city is also a district, and box shot axis integration is high, good economic benefits, traffic demand, should be more bearing capacity supply [23]. The middle is the area often the land price highland, the proportion of road occupation is not likely to be too high, the difficulty of building higher grade roads, and thus the need for a much higher proportion of public transport sharing than other areas, the need for road space in as much as possible to plan the rail and other public transport facilities.

From the perspective of complex adaptive system theory, the node is a space enclosed by its surrounding road network dividing other spaces, and its size and shape are directly influenced by the density and shape of the road network. Completeness is a measure of the success of a retrieval system in detecting relevant results from a sample set, that is, the ratio of relevant results detected to the total number of objects in the retrieval system; accuracy is a measure of the signal-to-noise ratio of a retrieval system. The nature of the land and volume ratio of the node space determines the demand and characteristics of the road traffic of the "main body," that is, the "traffic flow." The relationship between the road and the public transport capacity it carries and the "traffic flow" affects the utilization of the road and thus the utilization of the node space. The grid method uses the method of comparing standard indicators to assess the size, car-carrying capacity, and shape of the road network as a physical network, derives the standard indicator system and general intersection spacing, proposes a statistical calculation table to facilitate the comparison of indicators, and then draws conclusions from the table calculation and intersection spacing comparison. The grid method is based on the norms and experience of the road supply perspective assessment method, and also needs to cooperate with other methods of assessment. From the research progress of the national standard, the urban morphology has an increasingly influential role in road network planning, and the non-homogeneity of urban space indicates that the national standard index cannot guide the control plan of all lots, much less the node revision plan.

A comprehensive and systematic comparison with an index system based on the needs of each level of planning is the basis for evaluation, and more empirical research is needed in various places to accumulate experience and data. The saturation method uses traffic indicators and traffic models to analyze the relationship between traffic demand and supply and proposes an overall saturation table to grasp the relationship between traffic supply and demand in the statutory urban planning stage, while the "four-stage" traffic model can output the saturation of each road section, which is a visual reflection of the traffic supply and demand in the future years under the implementation of the plan [24]. In the general plan, the overall saturation table mainly calculates and grasps the average saturation of the backbone road network, while constructing as detailed a traffic model as possible to grasp the saturation forecast of each road as accurately as possible. In the revision and road planning and design, the evaluation results from the master plan and control plan preparation and decision-making are used to

analyze the system connection between the project and the lot and the city, and to improve the design level and project value.

5. Experimental Verification and Conclusion

5.1. Algorithm Parameter Testing. After screening, the main parameters selected for the algorithm include the sampling window size is 9, the evaluation region size is 3, the number of convolutional kernels is 4, the convolutional kernel edge length k$d = 7$, the data segment length is 1500, the optimizer is RMSProp, the learning rate is moving average exponential decay mode, and the segmentation mode is "window-to-window" mode with high model prediction accuracy and training speed. In the long term, the construction of transportation infrastructure and the supply of transportation services improved accessibility, and the increased commuting costs of traffic congestion reduce the value of regional land use. The model hyperparameters to be further validated are the number of neurons in the hidden layer of the cross-long-short memory network and the orbit-long-short memory network. In the cross-long-short memory network, the number of hidden layer neurons from 1 to 20 was tested considering the model training time reason. As shown in Figure 4, for each value of the number of hidden neurons, 100 cycles of training were performed, the percentage error of the test set was collected, and the trough of each 100 values corresponded to the highest value of the test accuracy, and it was found that the model with 11 hidden neurons had better results.

The number of hidden layer neurons from 1 to 20 was tested in the railroad long and short memory network, and each peak shows the training process when the test accuracy reaches its highest point, and the model with 7 hidden layer neurons was found to have better prediction results. The CNN-CLSTM model was compared with various neural network models including feedforward neural networks (FFNN), convolutional neural networks (CNN), and multidimensional long and short-term neural networks (MDLSTM) in terms of training speed and model effectiveness. In addition to the constituent models, feedforward neural networks were also used as control models for the comparison of model effects due to their simple model structure. The training speed was reflected by the performance of each model on the training set, and the model effects were evaluated by the performance of individual models on the test set. About 100 steps of training were performed for each model, and the percentage errors of the training and test set data are shown in Figure 5.

From the percentage error results of the training set, it can be seen that the CNN-MDLSTM has a relatively fast convergence speed among all models within 30 training steps, and the improved CNN-MDLSTM model converges faster than other models after 30 training steps. From the prediction effect, the improved CNN-MDLSTM significantly improves the accuracy of the results. Adding the directional prior information reduces the prediction model error from 16.36% to 13.30%, with an accuracy improvement of about 3%; adding the rail traffic prior information reduces the error from 16.32% to 13.30%, with an accuracy improvement of about 3%.

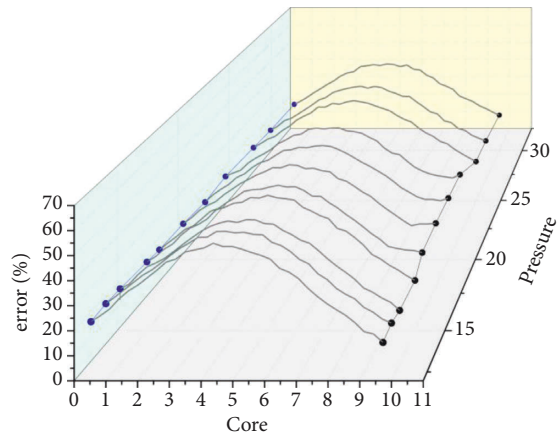


Figure 4: Percentage error of the model corresponding to different number of neurons in the hidden layer of the crossover neural network.

5.2. Evaluation Index Weight Optimization Example Validation. Completeness rate and accuracy rate are important metrics commonly used in the field of information retrieval to reflect the effectiveness of retrieval. Completeness rate is a measure of the success of a retrieval system in detecting relevant results from a sample set, that is, the ratio of relevant results detected to the total number of objects in the retrieval system; accuracy rate is a measure of the signal-to-noise ratio of a retrieval system, that is, the percentage of correct results detected to the total number of results detected. As shown in Figure 6, the detection rate and accuracy rate of residential land use are 57.14% and 85.71%, respectively, and the detection rate and accuracy rate of road use are 67.35% and 86.84%, respectively, with the accuracy rate reaching more than 80%, but the detection rate is low, indicating that the low-utility land evaluation and identification method proposed in this article can identify the low-utility land within the city more accurately, and some errors in some of the identification levels and a very small percentage of them are not identified, but the overall identification is good. Land-use types do not respond in real-time to make changes after traffic conditions change, and it takes time to plan and build transportation facilities after land-use types change, which objectively affects the accuracy of data-driven models. The main goal of this article is to first identify the parcels that may belong to low-utility land as much as possible, and then the staff will conduct field research on the specific parcels to determine whether they are classified as urban renewal objects, thus reducing the research workload and therefore focusing more on the accuracy rate, which leads to some parcels being classified below the moderate level.

In this article, taking the road land type as an example, the optimized weights were calculated for 186 road land parcels, and the internal property index and external characteristic index of the road land were calculated by

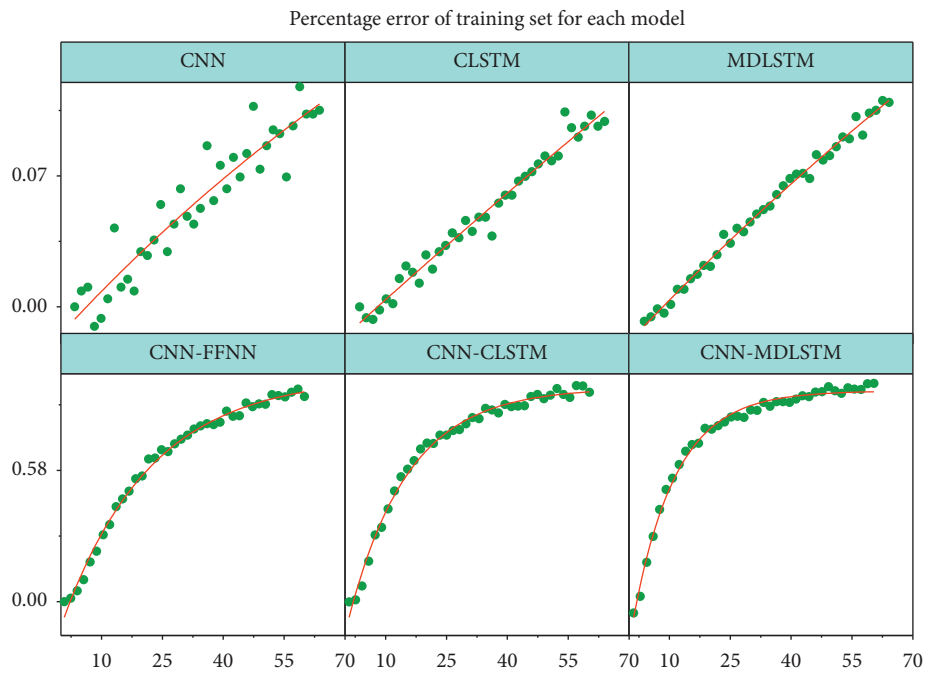


Figure 5: Percentage error of training set for each model.

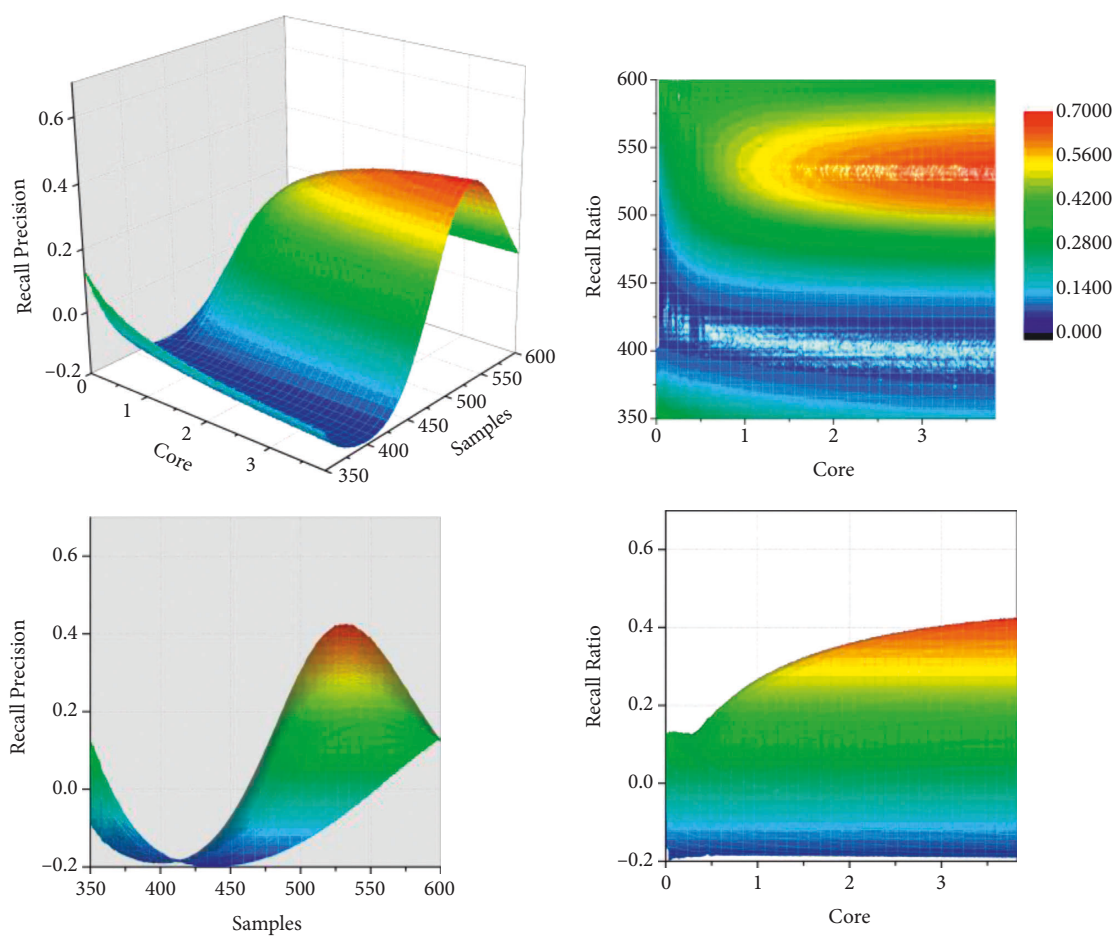


Figure 6: The search completion rate and accuracy rate of different site types.

weighting the initially determined weights, and the standardized internal and external initial index values and the internal and external indices were used as the input nodes of the BP neural network, respectively, and the initial weights were used as the output nodes to obtain the corrected values of the weight of each index in the land type. The results are shown in Figure 7.

As can be seen from Figure 7, the initial values of the normalized internal and external indicators of the road land and the internal and external indices are used as input to the initial weights for the BP neural network model correction, and the overall operational accuracy R values reach 0.9942 and 0.97882, respectively, and the training learning accuracy is good, which can realize the mining of the data features between the internal indicators based on the existing objectively determined weights, and the learning training process In the process of learning and training, the weights are adjusted, and the weights before and after the final adjustment are good. As can be seen from the figure, the optimization error (Error) of the internal property and external feature indicators of road land is much smaller than the set target error of 0.001, and the training results are good. The adjusted weights of the internal property indexes of the BP neural network are 0.1411, 0.1464, 0.2012, 0.2037, 0.1536, and 0.1540, and the initial weights obtained by the entropy weighting method can be seen, to reduce the weight of the building layer index in all the indicators, increasing the weight of the building density and building age indicators, and the weight of the site property and The adjusted weights of external characteristics of the BP neural network are 0.1067, 0.1911, 0.1241, 0.1086, 0.1149, 0.1635, 0.0476, and 0.1435 in order. Neural network modified weights, the weights of the infrastructure completeness index are more prominent. It can be seen that the index weights modified by the BP neural network algorithm retain the original meaning of the entropy weight method and also make the index weights more objective, so the weight modification method based on BP neural network can be applied in the evaluation of land-use efficiency.

Low-utility land is identified for all road land type parcels in a district, and the overall distribution is shown in Figure 8 after the division according to the low-utility evaluation grade. From the figure, it can be seen that the heavy low-utility land in the district is mainly distributed in the northeast corner of the district, and the heavier low-utility land is mainly distributed in the northeast and south areas, and the land-use situation of the whole district is mainly lighter low-utility grade and medium low-utility grade, and the overall land-use efficiency situation is good.

Through statistical analysis of the obtained results, there were 38 low-utility land parcels with moderately low-efficiency grade or above, accounting for 20.32% of the total number of road land parcels, including 8 heavy low-utility parcels, 30 heavier low-utility parcels, 56 moderately low-utility and lighter low-utility parcels each, and 37 light low-utility parcels, with the grade above moderately low-utility being the focus of this article. This article randomly selects the streetscape data of parcels with heavy inefficiency grades

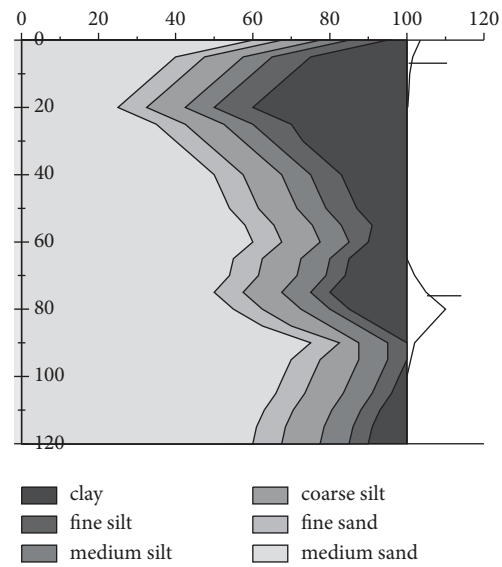


Figure 7: Road land-use internal property index weighting correction.

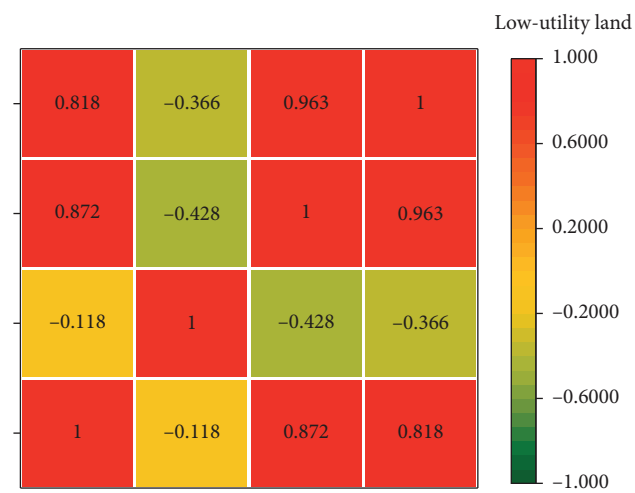


Figure 8: Distribution of evaluation results for identification of low-utility land for road use.

and heavier inefficiency grades in the study area for viewing and verification and finds that the overall identification effect is good using the low-utility land identification and evaluation method proposed in this article, and can accurately identify parcels with serious inefficiency in land use. From the overall spatial distribution, the inefficient road land above the medium level shows a more obvious clustering feature and a circling structure decreasing from the center of the area to the outer edge step by step.

6. Conclusion

This article compares the concept and characteristics of the urban central area and low-utility land in the central area, discusses the formation mechanism and influencing factors of low-utility land in the urban central area, and designs a big

data model for urban road land planning. The concept and criteria for defining low-utility land in urban central areas are determined, and land-use data, road data, river and water system data, service POI data, statistical yearbook data, and other auxiliary data are selected as basic data. The grid method uses the method of comparing standard indicators to assess the size, small car-carrying capacity, and morphology of the road network as a physical network, derives the standard indicator system and general intersection spacing, proposes a statistical calculation table that is convenient for comparing indicators, and then draws assessment conclusions by comparing the table calculation and intersection spacing. Finally, the feasibility of the proposed index system and the index weight determination model is verified by the case 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] B. Chen, Y. Tu, Y. Song et al., "Mapping essential urban land use categories with open big data: results for five metropolitan areas in the United States of America," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 178, pp. 203–218, 2021.

[2] B. Chen, B. Xu, and P. Gong, "Mapping essential urban land use categories (EULUC) using geospatial big data: progress, challenges, and opportunities," *Big Earth Data*, vol. 5, no. 3, pp. 410–441, 2021.

[3] X. Lanqin, "Intelligent multimedia urban planning Construction based on spectral clustering algorithms of large data mining," *Multimedia Tools and Applications*, vol. 79, no. 47-48, Article ID 35183, 2020.

[4] S. Alqadhi, J. Mallick, A. Balha, A. Bindajam, C. K. Singh, and P. V. Hoa, "Spatial and decadal prediction of land use/land cover using multi-layer perceptron-neural network (MLP-NN) algorithm for a semi-arid region of Asir, Saudi Arabia," *Earth Science Informatics*, vol. 14, no. 3, pp. 1547–1562, 2021.

[5] M. J. Jun, "A comparison of a gradient boosting decision tree, random forests, and artificial neural networks to model urban land use changes: the case of the Seoul metropolitan area," *International Journal of Geographical Information Science*, vol. 35, no. 11, pp. 2149–2167, 2021.

[6] Y. Qi, S. Chodron Drolma, X. Zhang et al., "An investigation of the visual features of urban street vitality using a convolutional neural network," *Geo-Spatial Information Science*, vol. 23, no. 4, pp. 341–351, 2020.

[7] A. R. Honarvar and A. Sami, "Towards sustainable smart city by particulate matter prediction using urban big data, excluding expensive air pollution infrastructures," *Big data research*, vol. 17, pp. 56–65, 2019.

[8] L. Mu, L. Wang, Y. Wang, X. Chen, and W. Han, "Urban land use and land cover change prediction via self-adaptive cellular based deep learning with multisourced data," *Ieee Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 12, no. 12, pp. 5233–5247, 2019.

[9] X. Wu, X. Liu, D. Zhang, J. Zhang, J. He, and X. Xu, "Simulating mixed land-use change under multi-label concept by integrating a convolutional neural network and cellular automata: a case study of Huizhou, China," *GIScience and Remote Sensing*, vol. 59, no. 1, pp. 609–632, 2022.

[10] H. Elmannai and A. D. AlGarni, "Classification using semantic feature and machine learning: land-use case application," *TELKOMNIKA (Telecommunication Computing Electronics and Control)*, vol. 19, no. 4, pp. 1242–1250, 2021.

[11] X. Deng, P. Liu, X. Liu et al., "Geospatial big data: new paradigm of remote sensing applications," *Ieee Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 12, no. 10, pp. 3841–3851, 2019.

[12] A. Jamali, "Land use land cover mapping using advanced machine learning classifiers: a case study of Shiraz city, Iran," *Earth Science Informatics*, vol. 13, no. 4, pp. 1015–1030, 2020.

[13] H. Chen, T. Cheng, and X. Ye, "Designing efficient and balanced police patrol districts on an urban street network," *International Journal of Geographical Information Science*, vol. 33, no. 2, pp. 269–290, 2019.

[14] M. Mohammadi and A. Sharifi, "Evaluation of convolutional neural networks for urban mapping using satellite images," *Journal of the Indian Society of Remote Sensing*, vol. 49, no. 9, pp. 2125–2131, 2021.

[15] Y. Chen, Q. Weng, L. Tang, Q. Liu, X. Zhang, and M. Bilal, "Automatic mapping of urban green spaces using a geospatial neural network," *GIScience and Remote Sensing*, vol. 58, no. 4, pp. 624–642, 2021.

[16] E. Saralioglu and O. Gungor, "Semantic segmentation of land cover from high resolution multispectral satellite images by spectral-spatial convolutional neural network," *Geocarto International*, vol. 37, no. 2, pp. 657–677, 2022.

[17] A. Jamali, "Improving land use land cover mapping of a neural network with three optimizers of multi-verse optimizer, genetic algorithm, and derivative-free function," *The Egyptian Journal of Remote Sensing and Space Science*, vol. 24, no. 3, pp. 373–390, 2021.

[18] D. He, Q. Shi, X. Liu, Y. Zhong, and X. Zhang, "Deep subpixel mapping based on semantic information modulated network for urban land use mapping," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 59, no. 12, Article ID 10628, 2021.

[19] D. Zhang, X. Liu, X. Wu, Y. Yao, X. Wu, and Y. Chen, "Multiple intra-urban land use simulations and driving factors analysis: a case study in Huicheng, China," *GIScience and Remote Sensing*, vol. 56, no. 2, pp. 282–308, 2019.

[20] X. Zhang, L. Liu, L. Xiao, and J. Ji, "Comparison of machine learning algorithms for predicting crime hotspots," *IEEE Access*, vol. 8, Article ID 181302, 2020.

[21] K. L. M. Ang and J. K. P. Seng, "Big data and machine learning with hyperspectral information in agriculture," *IEEE Access*, vol. 9, Article ID 36699, 2021.

[22] A. Al Kafy, A. Al Rakib, A. Al Rakib et al., "The operational role of remote sensing in assessing and predicting land use/land cover and seasonal land surface temperature using machine learning algorithms in Rajshahi, Bangladesh," *Applied Geomatics*, vol. 13, no. 4, pp. 793–816, 2021.

[23] F. P. Diez, J. C. Sinca, D. R. Vallès, and J. M. Campos Cacheda, "Evaluation of transport events with the use of big data, artificial intelligence and augmented reality techniques," *Transportation Research Procedia*, vol. 58, pp. 173–180, 2021.

[24] V. Nasiri, A. A. Darvishsefat, R. Rafiee, A. Shirvany, and M. A. Hemat, "Land use change modeling through an integrated Multi-Layer Perceptron Neural Network and Markov chain analysis (case study: arasbaran region, Iran)," *Journal of Forestry Research*, vol. 30, no. 3, pp. 943–957, 2019.

Hindawi
Computational Intelligence and Neuroscience
Volume 2022, Article ID 6321101, 11 pages
https://doi.org/10.1155/2022/6321101

Research Article

Optimization of Residential Landscape Design and Supply Chain System Using Intelligent Fuzzy Cognitive Map and Genetic Algorithm

Tingyin Deng

Sichuan University of Science & Engineering, Zigong, Sichuan 643000, China

Correspondence should be addressed to Tingyin Deng; tingyind@suse.edu.cn

Received 12 July 2022; Revised 18 August 2022; Accepted 31 August 2022; Published 13 September 2022

Academic Editor: Ning Cao

This work intends to optimize residential landscape design and Supply Chain (SC) network systems. First, Fuzzy Cognitive Map (FCM) intelligent assistance and genetic algorithm (GA) are used to study residential landscape design and its integration with SC deeply. Weight matrix interactions are employed to implement iterative inference for FCM. The functions are transformed to unify variables of different scopes. Subsequently, a weighting method is proposed to deal with the disadvantage of the simple average method being too general. In addition, the Hebbian learning algorithm is used to adjust the state nodes and the connection weights. Finally, according to the fitness function of the GA and logistic regression (LR) model, residential landscape design and SC are combined. The simulation experiment results show that the causal relationship analysis between SC networks under fuzzy cognition shows that the state errors of each specific situation are 0.21, 0.16, and 0.24, respectively. The total average error is 0.21 in the case of multiple iterations. The average error of the result vector under fuzzy cognition and the operation of the actual result is 0.20, 0.15, and 0.24, respectively, and the error value is much reduced. The simulation accuracy of the GA-LR method for residential landscape design is improved from 77% to 84.7%. The "kappa coefficient" is also improved to 82.3%. The conclusion shows that the weight matrix is used to analyze the high-quality performance of landscape design according to the specific situation of SC. For each specific case, FCM is effective in reducing errors over multiple iterations. Under the GA-LR method, fewer geographic location types and larger accuracy deviations can improve the simulation accuracy.

1. Introduction

As human needs evolve, techniques and styles change, and so do landscape design trends. Landscape architecture in the future will not only be limited to landscape design, but will also provide a series of solutions around social, economic, environmental, and other issues. Technological innovation and sustainable development means can improve people's quality of life. The technology used in the smart landscape is mainly modern information technology, but also digital technology. The digital landscape is the main body of the smart landscape. The application of digital technology and methods can break through the limitations of traditional design techniques and construction materials and greatly release the creativity of landscape planners. Many scholars believe that Supply Chain (SC) [1–3] is a process often used by enterprises. The actual process is that the raw materials are processed by the enterprise and finally transferred to the customers through various channels such as sales and various forms of business activities. However, with the rapid development of science and technology and the rapid progress of society, SC has long been not only used in business. At present, the SC functional network structure model is very helpful for residential landscape design [4–6]. Compared with developed countries in Europe and America, the development of residential landscape design in China is still at a low level [7–9]. For further development, China needs to make up the gap with foreign residential landscape design standards [10–12] and optimize the practical application of SC systems in related industries.

International scholars have important research theories on optimizing residential landscape design and SC systems [13–15]. There are more than 10,000 articles about landscape design in the Chinese master thesis. These research directions are standardized residential landscape design. In addition, contemporary real estate enterprises are investigated on the spot, and some historical and common problems in residential landscape design are further summarized. Chinese residential landscape design is too simple and lacks humanistic considerations. It is more inclined to residential attributes and pays less attention to the spatial, humanistic, and temporal value of landscape facilities. Foreign scholars have conducted special studies on the quality of living space [16–18] and European standards. They studied landscape design, planted seedlings, and developed norms and standards for ornamental plants. In addition, related scholars have also specialized in research strategies for 3D landscape representation. They perform 3D modeling [19–21] to seek a standardized study of landscape design. In water-constrained countries, waterwise landscaping is considered an important landscaping method to conserve water. However, there is still a concern and a lack of knowledge about people's preferences and the factors that influence them.

Nazemi Rafi et al. [22] examined the effect of landscape factors of plant combination (level 6) and cover type (level 3) on the preferences of 207 respondents. After obtaining the preferences that participants assigned to each plot, they used regression analysis to determine the effect of landscape attributes on overall preferences. Results showed that flowering plants were more popular ($p \leq 0.01$) and less expensive ($p \leq 0.01$) in water landscapes containing only herbaceous plants compared to other designed landscapes. The Visible Green Index (VGI) reflects the degree of greening from human vision to affect health and well-being and has gradually become a new type of urban green space index. Zhu et al. [23] designed and conducted an experiment to evaluate human responses to scenes from panoramic pictures of residential green spaces projected by virtual reality. This study provides a certain theoretical basis for the landscape planning of residential areas, which is helpful to improve the VGI theory and promote its establishment as a new type of urban green space index. O'Neill and Maravelias [24] mentioned in *Towards Integrated Landscape Design and Biofuel Supply Chain Optimization* the integration of landscape design with SC networks [25–27] to determine the synergy of the systems. The research of relevant international scholars makes it possible to optimize the residential landscape design and SC system.

In the research results of relevant scholars in this field, although the theory of the combination of residential landscape standardization design and SC system network is proposed, the research direction is also standardized residential landscape design. These theories do not involve optimizing residential landscape design and SC systems. Therefore, this study innovatively uses the intelligent assistance of a Fuzzy Cognitive Map (FCM) and genetic algorithm (GA) further to optimize the residential landscape design and SC system. The innovation lies in the optimization of the SC system, a new design for the residential landscape.

2. Methods

2.1. FCM. The function of FCM is to solve complex problems. Uncertain decision-making can impact decision-makers, and FCM constructs an important process of people's decision-making and analyzes its construction patterns to explain the external world. Cognitive maps have a variety of ways to build high-quality and often accurate results. The general principle of FCM is that the path of adjacent connection points in the cognitive graph is a vector arc segment, which expresses the causal relationship between nodes [28–30]. If any two connection points in the cognitive graph extend through a directed path, and the sign between the connection points is positive, it means that the relationship between them is positive. Otherwise, their relationship is an inverse relationship. Figure 1 shows a simple FCM.

In Figure 1, a fuzzy cognitive graph with n nodes, each node represents a concept in the system. This concept can be events, goals, trends of the system, etc. Each concept characterizes its properties using state values. Directed arcs represent the causal influence relationships between concepts. The weight w_{ij} reflects the degree of causal influence. $w_{ij}>0$ means that the increase in C_i will lead to the increase of Cj. There is a positive causal relationship between C_i and C_j, which is indicated by "+" in Figure 1; $w_{ij}<0$ means that the increase of C_i will increase C_j decreases. There is a negative causal relationship between C_i and C_j, that is, the part indicated by "–" in Figure 1; $w_{ij} = 0$ means that there is no causal relationship between C_i and C_j. C_i and C_j are represented by A, B, C, D, and E in Figure 1.

Based on the scenario studied in the present work, the discrete-time Markov decision process with continuous action space can be used to represent the optimal stochastic control problems of residential landscape allocation and location factor scheduling. Because of the complex transformation of the external environment, the residential landscape module cannot obtain accurate transformation information. Besides, the optimal solutions with low complexity are difficult to obtain by the general algorithm or equation through the previous methods. Therefore, the resource fuzzification of FCM under the reinforced deep learning (DL) algorithm is studied below to optimize the residential landscape design further. In the case of discrete time, the optimization aims at the conventional reinforcement learning problem composed of a residential landscape design environment and intelligent fuzzy cognitive assistance. The real-time reward is based on the timely action taken by the intelligent module after receiving the observation results in each process. The state space is the product of the resource allocation scenario of residential landscape design, that is, the current environmental state of the intelligent module. According to the DL model of the large-scale residential landscape, in equation (1), cx represents the data transmission direction and pointer; w_{ij} represents the weight; $x_i(t)$ represents the object's value at the current moment; $x_i(t+1)$ represents the value of the object at the next moment. The derivation process of the cognitive graph is not complicated, and its key function is

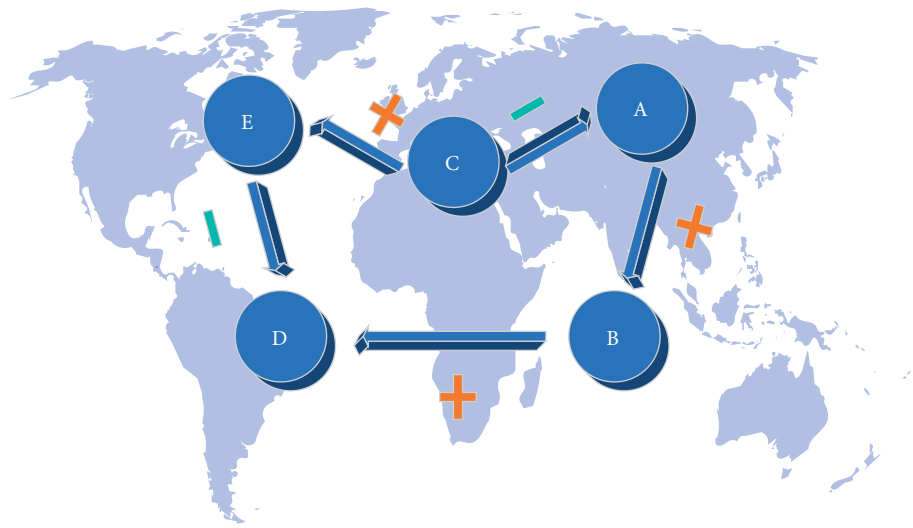


Figure 1: Simple FCM.

to grasp the concept nodes and weight matrix. Equation (1) represents a specific iterative function.

$$x_i(t+1) = f\left[\left(\sum_{\substack{j=1 \\ j\neq i}}^{n} w_{ij}x_j(t) + cx_i(t)\right)\right] \tag{1}$$

The transformation function [31] is used to change the data useless to the system. The variation range of the transformation function is [0,1]. Equation (2) is a threshold function; equation (3) is a hyperbolic tangent function; equation (4) is a sigmoid function [32–34]; Equation (5) is a saturation function. In equations (2)–(5), x represents a real number between 0 and 1.

$$f(x) = \begin{cases} 1, x \geq 0 \\ 0, x < 0 \end{cases} \tag{2}$$

$$f(x) = \tanh(x) \tag{3}$$

$$f(x) = \frac{1}{1+e^{-x}} \tag{4}$$

$$f(x) = \begin{cases} 1, & x \geq \frac{1}{k} \\ kx, & -\frac{1}{k} \leq x \leq \frac{1}{k}. \\ -1, & x < -\frac{1}{k} \end{cases} \tag{5}$$

Based on the general problem of simple average, the weight method is proposed, mainly used to conduct the weighted average on the simple average. Equation (6) demonstrates the principle of the method.

$$W = f\left(\sum_{k=1}^{n} b_k W_k\right) \tag{6}$$

In equation (6), W means the final weight; n denotes the number of experts; W_k refers to the weight matrix according to the kth expert; f stands for conversion function, which is used to change all data in the weight matrix to the range between −1 and 1; b_k means the accuracy of matching expert experience and knowledge.

Hebbian learning algorithm. The two most important factors in Hebbian learning of FCM are the dynamic characteristics of process and environment. The value of connection weight can be further adjusted by adjusting the state node. Wij denotes the weight between node cj and node ci, when xj represents the state of node cj and xi represents the state of ci. Therefore, when xj and xi change, the adjustment function is shown in:

$$\Delta w_{ji} = \gamma w_{ji}^{k-1} + \eta A_i^{k-1}\left(A_j^{k-1} - \operatorname{sgn}\left(w_{ji}^{k-1}\right)w_{ji}^{k-1}A_i^{k-1}\right) \tag{7}$$

In equation (7), η means learning factor; γpoints to the attenuation factor.

GA algorithm is used to select the values with good adaptability and pass the values with high-quality characteristics to the next generation. In equation (8), M stands for the population size; F_i accords to the fitness of monomer i. Equation (8) illustrates the calculation of the probability that i is selected.

$$P_i = \frac{F_i}{\sum_{i=1}^{M} F_i} \quad (i = 1, 2, \ldots, \mathrm{M}). \tag{8}$$

The minimum deviation e means that the fitness of GA has been reduced to the minimum state. Then, its termination condition is expressed as:

$$\left|f_{\max} - f^*\right| < e \tag{9}$$

In equation (9), f^* represents the fitness goal and $f_{\max}$ denotes the maximum fitness. The fitness function is set to the inverse of the prediction error to train the FCM. Equation (10) manifests its expression.

$$f = \frac{1}{e} = \frac{1}{\sqrt{(1/m)\sum_{i=1}^{m}(s_i - \hat{s}_i)^2}}. \tag{10}$$

In equation (10), $s_i$designates the actual value and $\hat{s}_i$ stands for the predictive value of FCM.

$$Probability\,(gene\ j\ is\ selected) = \frac{f_j}{\sum_{k=1}^{m} f_k}. \tag{11}$$

In equation (11), $f_j$refers to the fitness value of the jth gene.

$$\mathrm{minG}\,(Q, R) = \frac{k\lambda}{Q} + h\left(\frac{Q}{2} + R - \lambda\tau\right) + \frac{Pn(R)\lambda}{Q} \tag{12}$$

In equation (12), $n(R)$ means the prediction of landscape design land; Q accords to the amount of landscape land occupation; R stands for the current land stock; and H represents the market interest rate.

$$g\left(s_i^t\right) = \begin{cases} 0, if\, s_i^t < a_i \\ \left(s_i^t - a_i\right)/\{2\left(m_i - a_i\right), if\, a_i \le s_i^t \le m_i \\ 0.5 + \left(s_i^t - m_i\right)/\{2\left(b_i - m_i\right), if\, m_i \le s_i^t \le b_i \\ 1, if\, s_i^t \ge b_i \end{cases}. \tag{13}$$

In equation (13), the state value of fuzzy cognition is generally set as a real number between 0 and 1. The purpose of introducing a nonlinear function approximator into the algorithm is to better deal with the problems of continuous space and multi-dimensional state space. The ordinary linear function approximators cannot properly deal with the problems of the frequency resources of the unauthorized frequency band. In the following algorithm, the residential landscape design is the core, the SC is the main structure of the algorithm, and the supply network inputs the state into St and outputs the result. Landscape design is triggered by the subsequent trigger action of the intelligent module under the action of the surrounding environment of the residence and feeds back to the new state. Therefore, it is necessary to standardize the numbers in this range. The standardized method is the historical minimum, maximum, and average values obtained in time t. In equation (13), s_i^t means the ith state value at time t.

$$\begin{aligned} a_i &= \min_{t\in T}\left\{s_i^t\right\} \\ b_i &= \max_{t\in T}\left\{s_i^t\right\} \\ m_i &= \underset{t\in T}{average}\left\{s_i^t\right\} \end{aligned} \tag{14}$$

The state vector can be obtained based on the operation of fuzzy cognition [35–37], which can be expressed as an equation:

$$S_t = f\left(S_{t-1} \times E\right), t = 1, 2, \ldots. \tag{15}$$

In equation (15), $f(S_t)$ refers to a transfer function whose specific purpose is to set the state values between 0 and 1; s_i^t represents the ith state value at time t; $\hat{s}_i^t$ accords to the results of the fuzzy cognitive calculation. Then, the difference between $\hat{s}_i^t$ ands_i^t is the error in the prediction. Furthermore, equation (16) displays the square root of all state values.

$$e_t = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(s_i^t - \hat{s}_i^t\right)^2}. \tag{16}$$

Equation (17) discloses the expression of the total prediction error of the GA.

$$\min \sum_{t\in T} e_t \tag{17}$$

$$Probability\,(gene\ j\ is\ selected) = \frac{f_j}{\sum_{k=1}^{m} f_k}. \tag{18}$$

In equation (18), f_j represents the fitness value of the jth weight. In the Euclidean distance, prediction results of $\hat{S}_G^C$ are closest to the actual target vector S^T. The calculation is shown in the equation:

$$\hat{S}_G^C = arg \quad \min\left\{distance\left(S^T, \hat{S}^T\right)\right.. \tag{19}$$

The emergence of the definition of $\hat{S}_H^C$ means that the analysis stops because the data collected at this time are not the category of historical data anymore.

$$\hat{S}_H^C = arg \quad \min\left\{distance\left(S, \hat{S}_G^C\right)\right. \tag{20}$$

2.2. Genetic Algorithm. GA is an optimization algorithm with a high utilization rate in the academic field, based on Darwin's *Theory of Evolution* and related elements in genetics. It can optimize the target. Its composition is generally divided into four aspects: maximum number of iterations, crossover rate, control parameters, fitness function, and chromosome coding. The foundation of GA is based on mathematical theorems. GA has great advantages for individuals with good fitness. The definition of a genetic algorithm is that genes and chromosomes are elements. Genes are the smallest elements that assemble into chromosomes. The form that exists on chromosomes is string encoding. Characters in a string represent genes. Its algorithmic meaning is the relationship between variable solution set and variable solution. Meanwhile, these solutions are combined to form the population of GA. LR method is a common learning algorithm whose advantage is that it has characteristics to describe nonlinear problems. Simultaneously, the LR model can usually solve common problems, so it has wide practicability. Equation (21) indicates the expression form of the sigmoid function of LR:

$$f(t) = \frac{1}{1 + e^{-t}}. \tag{21}$$

In equation (23), t is a variable and represents a linear function. Equation (22) refers to the expression of fitness function *Fit(f(x))*, which is implemented to find the minimum value of objective function *f(x)*:

$$Fit(f(x)) = \begin{cases} c_{\max} - f(x), & f(x) < c_{\max} \\ o, & else \end{cases}. \quad (22)$$

Equation (21) denotes the expression of fitness function *f(x)*, which is adopted to find the maximum value of objective function *f(x)*.

$$Fit(f(x)) = \begin{cases} f(x) - c_{\min}, & f(x) > c_{\min} \\ o, & else \end{cases}. \quad (23)$$

In equations (22) and (23), $C_{\max}$ and $C_{\min}$ mean the maximum and minimum values of the objective function, respectively. The fitness function in the present work can effectively identify all chromosomes and find unsolved and normal chromosomes. If the identification degree between them is not high, the value of the fitness function should be adjusted to ensure normal identification. This is a very important step in the design of fitness functions.

Simple genetic algorithm (SGA) has three operators: crossover operator, selection operator, and mutation operator. The selection operator refers to the operator that can simulate biological evolution, which can decide whether to retain the individual according to the value of the fitness function. Generally speaking, the probability of being selected is closely related to fitness. The higher the fitness, the greater the probability of being selected. Equation (24) signifies the specific expression of the selection operator of SGA:

$$p_i = \frac{f_i}{\sum_{i=1}^{n} f_i}. \quad (24)$$

In equation (24), *Pi* accords to the probability that monomer i is selected, and *fi* represents the fitness of monomer I:

$$Fit = \frac{p_1 + p_2}{p_{all}} \quad (25)$$

Equation (25) refers to the function of fitness *Fit*, where *P1* and *P2* represent the number of records of landscape land and non-landscape land. *Pall* stands for the total number of records in the data set of the sample. Equations (26) and (27) calculate adaptive variation rate *pm* and crossover rate *pc*, respectively.

$$p_c = pc_1 - (pc_2 - pc_1) * \frac{popsize - n}{popsize}. \quad (26)$$

$$p_m = pm_1 + (pm_2 - pm_1) * \frac{popsize - n}{popsize} \quad (27)$$

In equations (26) and (27), *pc1* and *pc2* point to the maximum and minimum cross rate, respectively. Parameter n represents the number of individuals. The change of n and *pc* is accompanied by the iterative process of the algorithm. The change of *pm* and *pc* is determined according to the change of n.

SGA [38–40] is the model's foundation in the present work. Its solution process is: first, the population is initialized, which is to determine the control parameters of GA and generate results. Second, the fitness value of the current monomer is calculated, the monomer is screened, and cross-operation is carried out. Third, the mutation operation is carried out and the iteration is judged immediately. Fourth, the optimal solution is output. Figure 2 displays the specific flow of GA.

The disadvantage of the rule-based method is that it is often difficult for people to formulate a good rule base in an unfamiliar scene. After the membership function is established, the corresponding processing rule base is indispensable. Finally, the proposed method is simulated, and GA is used for residential landscape design and SC system optimization. The experiment uses SYNTHIA-Dataset. This dataset is a large-scale dataset of photorealistic renderings of virtual cities with semantic segmentation information, which is proposed for scene understanding in research fields such as autonomous driving or urban scene planning and garden design. The dataset provides fine-grained pixel-level annotations for 11 categories of objects (empty, sky, buildings, roads, sidewalks, fences, vegetation, poles, cars, signal signs, pedestrians, and cyclists). Thirteen thousand four hundred seven training images were extracted from the video stream. This dataset is also known for its variability, including scenes (towns, cities, highways, etc.), objects, seasons, weather, etc. Experiments on this dataset can clearly show the improvement in the performance of the residential landscape design system after the optimization of the GA.

The software and hardware environment adopted in the experiment is Matlab. The hardware and software environment settings are on one PC. The operating system is Windows 10 version. The CPU is Intel(R) Core (TM) i9-9900K CPU @ 3.60 GHz, and the memory is 8G. The programming language adopts PyThon 3.0.

3. Results

3.1. Fuzzy Cognitive Analysis of Residential Landscape SC. Causality between SC networks under fuzzy cognition is analyzed. The storage status of landscape design land is assumed to be 1, which indicates that the actual value is much higher than that in general. In Figure 3, the interval values are the smallest when symbol values are between [0–0.1] and [0.5–0.6]. On the contrary, the interval values are the largest when symbol values are between [0.4–0.5] and [0.9–0.1]. Besides, according to the histogram, all the interval values show an increasing trend. Figure 3 shows the values of key points at the fuzzy state.

Based on simulation experiments, the weight matrix performs well in analyzing the landscape design according to the SC. Figure 4 demonstrates the specific situation of the simulation. For each specific case, the premise that the FCP can reach a fixed value is to carry out multiple iterations. Figure 4(a) depicts the state error in specific cases. The error

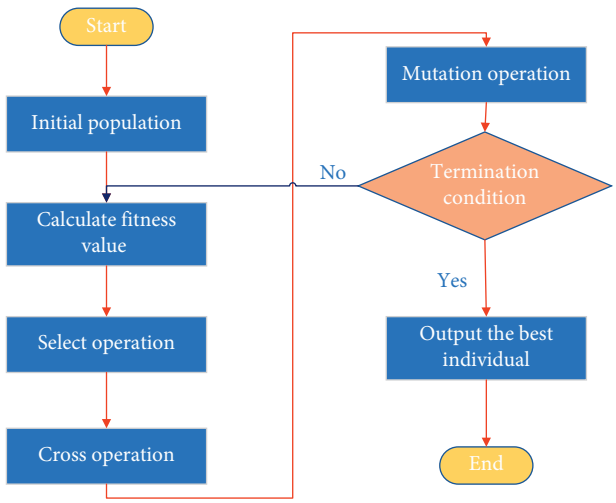


Figure 2: Flow of GA.

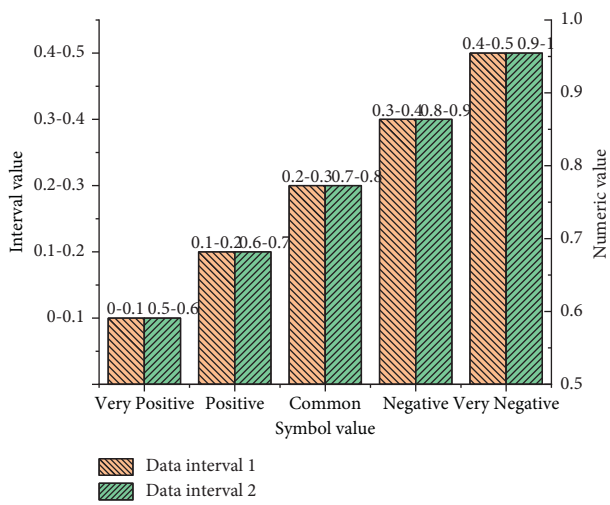


Figure 3: Values of key points at the fuzzy state.

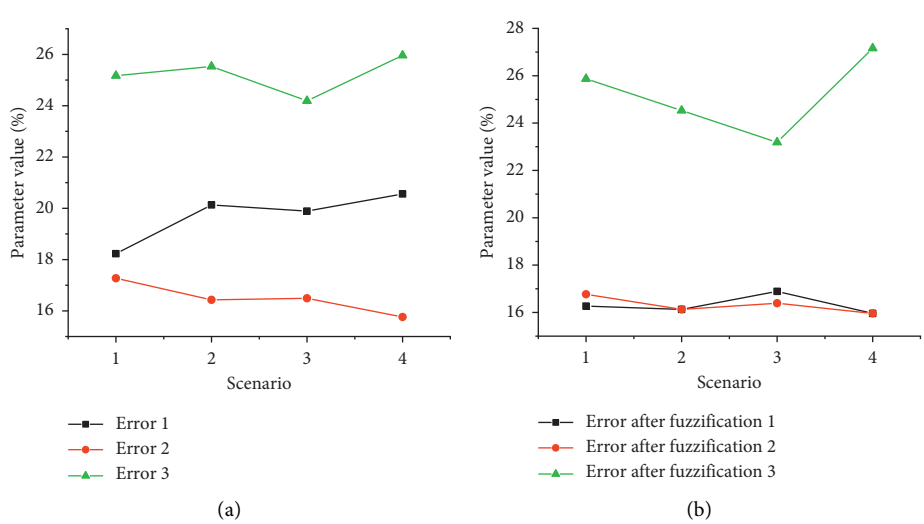


Figure 4: Simulation analysis of fuzzy cognition. (a) Error diagram; (b) error diagram after fuzzification.

is 0.21, 0.16, and 0.24 in turn, and the total average error is 0.21. According to the result vector under FCP and the operation of actual results, the error, in this case, is reduced a lot. According to the blurred result, the average error is 0.20, 0.15, and 0.24, in turn, according to Figure 4(b). Meanwhile, compared with the initial average error, the error value is reduced to 0.19. Figure 4 shows the fuzzy cognitive simulation analysis based on landscape design.

3.2. SC and GA Analysis. GA can achieve convergence based on multiple iterations, and the mean value of the objective function can be infinitely close to zero. Therefore, samples with prediction errors close to zero can be obtained through the optimal weight combination, which is the best regression coefficient of LR. Figure 5 shows the convergence curve of GA.

LR coefficients are obtained by genetic algorithm training. LR training can obtain specific regression coefficients. Figure 6(a) and 6(b) depicts the curves of the regression coefficients under different factors, respectively. The value of the elevation factor changes from −0.33 to −0. Twenty-two under two methods, while the other two data groups in Figure 6(a) change from −0.63 to −0.74 and from −0.71 to −0.79. In Figure 6(b), the value of the factor between the landscape design and the river changes from −0.43 to −0.45, and the other two groups of data have an upward trend. Only the elevation factor plays a positive role in landscape design.

Figures 7 and 8 depict the simulation results of the confusion matrix for the GA-based logistic regression (LR) and the traditional LR method, respectively. The matrix is used to calculate the simulation accuracy and Kappa coefficient.

In Figure 7, the actual values of landscape and non-landscape land are 5768 and 1621, respectively, and 9845 and 5169 under simulation conditions. In Figure 8, the actual values of landscape and non-landscape land are 5208 and 2691, respectively, and they are 1675 and 5163 under simulated conditions. In the research of residential landscape design, problems such as many geographical locations and insufficient sample richness should be considered to eliminate precision bias. In Figure 9, the comparison between the GA-optimized LR and the traditional LR method shows that the simulation accuracy of the proposed GA-based LR method in residential landscape design is significantly improved from 77% to 84.7%. The "Kappa coefficient" also increases to 82.3%. The proposed algorithm has better performance.

4. Discussion on the Application of Fuzzy Cognition and GA in Landscape Design

In Sections 3.1 and 3.2, the FCM and GA analysis results of residential landscape SC show that, through simulation experiments, the weight matrix performs well in analyzing the landscape design in the SC case. Bakhtavar et al. [41] pointed out that fuzzy cognitive graphs are widely used to analyze complex, causal-based systems in modeling,

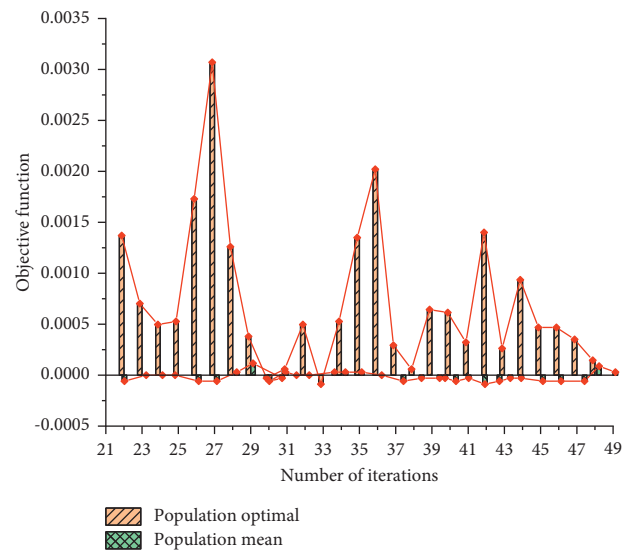


Figure 5: Convergence curve of GA.

decision-making, analysis, prediction, classification, etc. Managers cannot decide to implement and allocate resources to improvement projects based on the outputs of the risk management process. Rezaee et al. [42] attempted to accurately identify and prioritize potential failures in a production process with the help of crossover using an approach based on a multi-stage fuzzy cognitive graph approach and process failure mode and impact analysis techniques. Ladeira et al. [43] aimed to develop an FCM to identify and analyze the determinants of digital entrepreneurship. Iqbal et al. [44] proposed a new GA-based feature reduction technique. This hybrid approach reduced the feature set size to 42% without compromising accuracy.

The premise of a fuzzy graph reaching a fixed value is multiple iterations. In addition, the state error of the specific case is given. This error is relatively large, and the total average error is also relatively large. The operation error between the result vector under FCM and the actual result is much reduced. Furthermore, the error after blurring is reduced compared to the initial average error. GA plays a role in improving the accuracy of optimizing the eigenvalues.

In addition, the analysis of residential landscape SC in the GA direction can achieve convergence based on multiple iterations. The mean of the objective function can be infinitely close to zero. Therefore, the optimal weight combination obtains a sample with a prediction error close to zero, that is, the optimal regression coefficient of LR. The LR coefficients obtained by GA training and the conventional LR training obtained by the maximum likelihood estimation algorithm are used to obtain specific regression coefficients. The analysis shows that the regression coefficient values obtained by the two different methods are quite different. The regression coefficient values obtained by the two methods are quite different. The results show that only the elevation factor plays a positive role in residential landscape design. Comparing GA and conventional LR methods, the

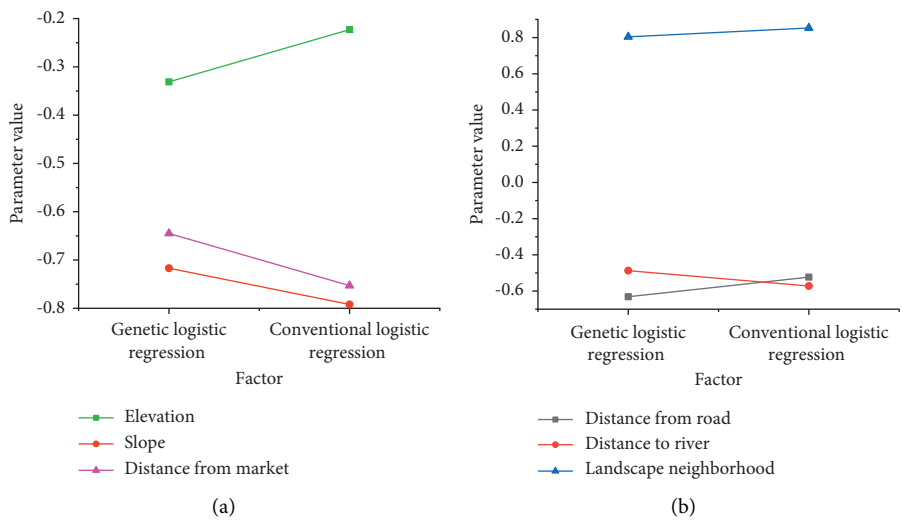


Figure 6: Changes of LR coefficient. (a) Genetic logistic regression coefficient. (b) Conventional logistic regression coefficient.

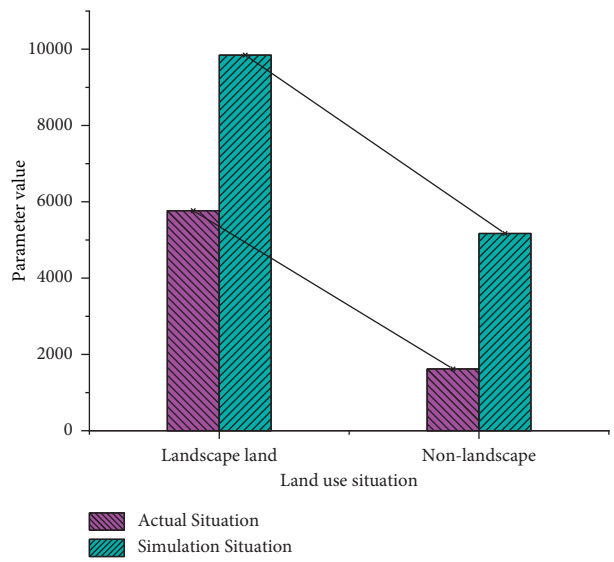


Figure 7: Confusion matrix for genetic LR.

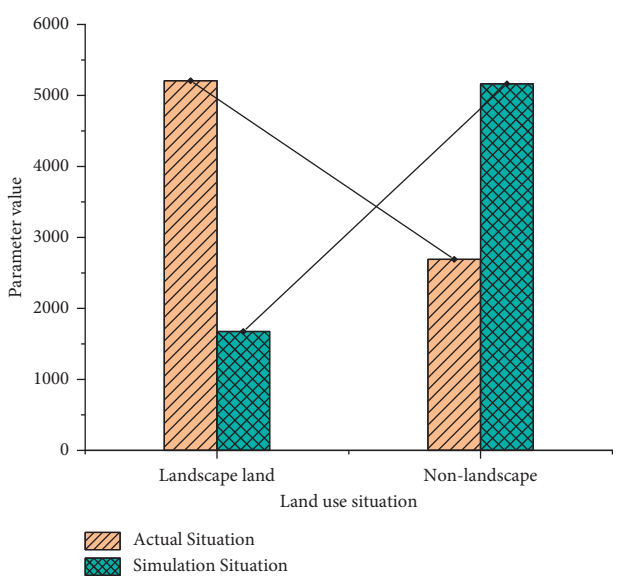


Figure 8: Confusion matrix for conventional LR.

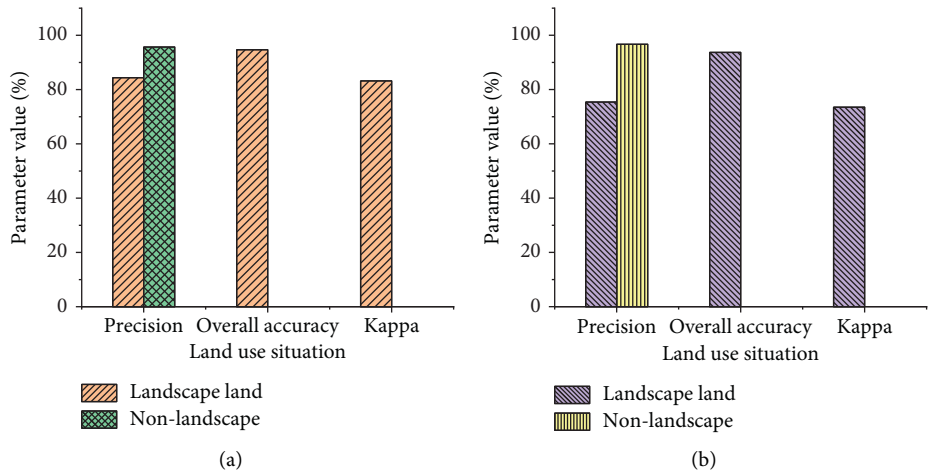


Figure 9: Comparison of simulation accuracy. (a) Genetic LR method; (b) conventional LR method.

simulation accuracy of the GA-LR method in residential landscape design is significantly improved. Therefore, simulated cognitive maps and GA play an irreplaceable role in residential landscape design and SC optimization.

5. Conclusion

The problem of combining residential landscape design with SC management and the current residential design in China is further optimized and improved to narrow the gap with foreign high-quality residential landscape design. The research work is to model, simulate and optimize residential landscape design. First, the previous research is the basis for the theoretical support. Classification and summarization are made according to the residential landscape design and the regional characteristics of geographical location, and LR is adopted as the model, which is combined with GA and the conventional LR method. Afterward, the fitting degree of this model is not enough, and the fitting ability is still weak due to the deviation of regional data and the influence of geographical location characteristics. Based on this situation, the GA-LR model is used to analyze residential landscape design and further study it. Finally, the model eventually has a good fitting effect, and its reliability is confirmed by verification and analysis. Besides, based on the SC management analysis of residential landscape design and the simulation research of SC network oriented to FCM, the logical relationship of SC network is very strong, which can well analyze the causal relationship between various nodes. Simultaneously, the complex state nodes can also be handled well further to ensure the accuracy and universality of each link. The following conclusions are drawn: (1) the weight matrix can analyze the high-quality performance of landscape design according to the situation of SC; (2) for each specific case, in the case of multiple iterations, the total average error of model simulation is small; (3) this method is very helpful to study residential landscape design; and (4) under the genetic LR method, the simulation accuracy can be significantly improved in the face of fewer types of geographical location and large accuracy deviation. The simulation data can show that the FCM constructed in the present work and the combination with the SC network system can better analyze and optimize the residential landscape design. There are still some deficiencies in the simulation accuracy calculation under genetic and conventional LR methods. A few types are considered for the geographical location of residential landscape design, and the sample richness is not enough so the accuracy deviation may be large. This problem will be the focus of follow-up research.

Data Availability

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethical Approval

This article does not contain any studies with human participants or animals performed by any of the authors.

Consent

Informed consent was obtained from all individual participants included in the study.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

The authors acknowledge the help from their university colleagues. This work was supported in part by Sichuan 2021–2023 Higher Education Talent Training Quality and Teaching Reform Project, Project No. JG 2021–1056, by Sichuan Smart Tourism Research Base Planning Project, Project No. ZHYR21-02, by Education Reform Project of Sichuan University of Science and Engineering in 2022, Project No. JG-2250.

References

[1] S. Min, Z. G. Zacharia, and C. D. Smith, "Defining supply chain management: in the past, present, and future," *Journal of Business Logistics*, vol. 40, no. 1, pp. 44–55, 2019.

[2] H. Fatorachian and H. Kazemi, "Impact of Industry 4.0 on supply chain performance," *Production Planning & Control*, vol. 32, no. 1, pp. 63–81, 2021.

[3] R. Casado-Vara, J. Prieto, F. D. Prieta, and J. M. Corchado, "How blockchain improves the supply chain: case study alimentary supply chain," *Procedia Computer Science*, vol. 134, pp. 393–398, 2018.

[4] H. S. Cinar and N. K. Aktas, "Xeriscape analysis: a case study in a residential garden in Istanbul," *J Environ Prot Ecol*, vol. 19, no. 4, p. 1904, 2018.

[5] K. R. Colter, A. C. Middel, and C. A. Martin, "Effects of natural and artificial shade on human thermal comfort in residential neighborhood parks of Phoenix, Arizona, USA," *Urban Forestry and Urban Greening*, vol. 44, Article ID 126429, 2019.

[6] G. R. Roshan, R. Oji, and S. Attia, "Projecting the impact of climate change on design recommendations for residential buildings in Iran," *Building and Environment*, vol. 155, pp. 283–297, 2019.

[7] A. Hami, F. F. Moula, and S. B. Maulan, "Public preferences toward shopping mall interior landscape design in Kuala Lumpur, Malaysia," *Urban Forestry and Urban Greening*, vol. 30, pp. 1–7, 2018.

[8] S. M. H. Atwa, M. G. Ibrahim, A. M. Saleh, and R. Murata, "Development of sustainable landscape design guidelines for a green business park using virtual reality," *Sustainable Cities and Society*, vol. 48, Article ID 101543, 2019.

[9] P. Motealleh, W. Moyle, C. Jones, and K. Dupre, "Creating a dementia-friendly environment through the use of outdoor natural landscape design intervention in long-term care facilities: a narrative review," *Health & Place*, vol. 58, Article ID 102148, 2019.

[10] S. Khatibi, L. Zare, and M. H. Kaboli, "The effect of green space and landscape design on promoting the sense of space belonging to the residential complex," *Hoviatshahr*, vol. 12, no. 3, pp. 19–28, 2018.

[11] L. A. Warner, J. M. Diaz, and A. K. Chaudhary, "Informing urban landscape water conservation extension programs using behavioral research," *Journal of Agricultural Education*, vol. 59, no. 2, pp. 32–48, 2018.

[12] Y. Feng, "Design of landscape gardens in urban residential areas," *Journal of Landscape Research*, vol. 10, no. 3, pp. 24–34, 2018.

[13] M. R. Saavedra, C. H. de O Fontes, and F. G. M. Freires, "Sustainable and renewable energy supply chain: a system dynamics overview," *Renewable and Sustainable Energy Reviews*, vol. 82, pp. 247–259, 2018.

[14] S. Tiwari, C. K. Jaggi, M. Gupta, and L. E. Cárdenas-Barrón, "Optimal pricing and lot-sizing policy for supply chain system with deteriorating items under limited storage capacity," *International Journal of Production Economics*, vol. 200, pp. 278–290, 2018.

[15] M. L. Tseng, A. S. F. Chiu, G. Liu, and T. Jantaralolica, "Circular economy enables sustainable consumption and production in multi-level supply chain system," *Resources, Conservation and Recycling*, vol. 154, Article ID 104601, 2020.

[16] X. Feng and T. Astell-Burt, "Residential green space quantity and quality and symptoms of psychological distress: a 15-year longitudinal study of 3897 women in postpartum," *BMC Psychiatry*, vol. 18, no. 1, pp. 348–411, 2018.

[17] L. Rui, R. Buccolieri, Z. Gao, W. Ding, and J. Shen, "The impact of green space layouts on microclimate and air quality in residential districts of Nanjing, China," *Forests*, vol. 9, no. 4, p. 224, 2018.

[18] R. Aerts, S. Dujardin, B. Nemery et al., "Residential green space and medication sales for childhood asthma: a longitudinal ecological study in Belgium," *Environmental Research*, vol. 189, Article ID 109914, 2020.

[19] D. Antón, B. Medjdoub, R. Shrahily, and J. Moyano, "Accuracy evaluation of the semi-automatic 3D modeling for historical building information models," *International Journal of Architectural Heritage*, vol. 12, no. 5, pp. 790–805, 2018.

[20] V. Fantini, M. Bordoni, F. Scocozza et al., "Bioink composition and printing parameters for 3D modeling neural tissue," *Cells*, vol. 8, no. 8, p. 830, 2019.

[21] A. I. Benzer and B. Yildiz, "The effect of computer-aided 3D modeling activities on pre-service teachers' spatial abilities and attitudes towards 3D modeling," *Journal of Baltic Science Education*, vol. 18, no. 3, pp. 335–348, 2019.

[22] Z. Nazemi Rafi, F. Kazemi, and A. Tehranifar, "Public preferences toward water-wise landscape design in a summer season," *Urban Forestry and Urban Greening*, vol. 48, Article ID 126563, 2020.

[23] H. Zhu, F. Yang, Z. Bao, and X. Nan, "A study on the impact of Visible Green Index and vegetation structures on brain wave change in residential landscape," *Urban Forestry and Urban Greening*, vol. 64, Article ID 127299, 2021.

[24] E. G. O'Neill and C. T. Maravelias, "Towards integrated landscape design and biofuel supply chain optimization," *Current Opinion in Chemical Engineering*, vol. 31, Article ID 100666, 2021.

[25] H. C. Su, T. W. D. Kao, and K. Linderman, "Where in the supply chain network does ISO 9001 improve firm productivity?" *European Journal of Operational Research*, vol. 283, no. 2, pp. 530–540, 2020.

[26] F. Goodarzian, H. Hosseini-Nasab, J. Muñuzuri, and M. B. Fakhrzad, "A multi-objective pharmaceutical supply chain network based on a robust fuzzy model: a comparison of meta-heuristics," *Applied Soft Computing*, vol. 92, Article ID 106331, 2020.

[27] B. Zahiri, P. Jula, and R. T. Moghaddam, "Design of a pharmaceutical supply chain network under uncertainty considering perishability and substitutability of products," *Information Sciences*, vol. 423, pp. 257–283, 2018.

[28] P. Morone, P. M. Falcone, and A. Lopolito, "How to promote a new and sustainable food consumption model: a fuzzy cognitive map study," *Journal of Cleaner Production*, vol. 208, pp. 563–574, 2019.

[29] K. Papageorgiou, P. K. Singh, E. Papageorgiou, H. Chudasama, D. Bochtis, and G. Stamoulis, "Fuzzy cognitive map-based sustainable socio-economic development planning for rural communities," *Sustainability*, vol. 12, no. 1, p. 305, 2019.

[30] M. Alipour, R. Hafezi, E. Papageorgiou, M. Hafezi, and M. Alipour, "Characteristics and scenarios of solar energy development in Iran: fuzzy cognitive map-based approach," *Renewable and Sustainable Energy Reviews*, vol. 116, Article ID 109410, 2019.

[31] F. Gu, S. Chen, Z. Zhou, and Y. Jiang, "An enhanced normalized step-size algorithm based on adjustable nonlinear transformation function for active control of impulsive noise," *Applied Acoustics*, vol. 176, Article ID 107853, 2021.

[32] P. Goel and S. S. Kumar, "Certain class of starlike functions associated with modified sigmoid function," *Bulletin of the Malaysian Mathematical Sciences Society*, vol. 43, no. 1, pp. 957–991, 2020.

[33] Z. Qin, Y. Qiu, H. Sun et al., "A novel approximation methodology and its efficient vlsi implementation for the sigmoid function," *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 67, no. 12, pp. 3422–3426, 2020.

[34] Y. C. Wu, C. W. Chow, Y. Liu et al., "Received-signal-strength (RSS) based 3D visible-light-positioning (VLP) system using kernel ridge regression machine learning algorithm with sigmoid function data preprocessing method," *IEEE Access*, vol. 8, pp. 214269–214281, 2020.

[35] H. S. Firmansyah, S. H. Supangkat, A. A. Arman, and P. J. Giabbanelli, "Identifying the components and interrelationships of smart cities in Indonesia: supporting policymaking via fuzzy cognitive systems," *IEEE Access*, vol. 7, pp. 46136–46151, 2019.

[36] P. P. Groumpos, "Intelligence and fuzzy cognitive maps: scientific issues, challenges and opportunities," *Studies in Informatics and Control*, vol. 27, no. 3, pp. 247–264, 2018.

[37] I. Pluchinotta, D. Esposito, and D. Camarda, "Fuzzy cognitive mapping to support multi-agent decisions in development of urban policymaking," *Sustainable Cities and Society*, vol. 46, Article ID 101402, 2019.

[38] H. Liu, S. Shi, P. Yang, and J. Yang, "An improved genetic algorithm approach on mechanism kinematic structure enumeration with intelligent manufacturing," *Journal of Intelligent and Robotic Systems*, vol. 89, no. 3-4, pp. 343–350, 2018.

[39] B. Li, P. Hu, N. Zhu, F. Lei, and L. Xing, "Performance analysis and optimization of a CCHP-GSHP coupling system based on quantum genetic algorithm," *Sustainable Cities and Society*, vol. 46, Article ID 101408, 2019.

[40] B. Rolf, T. Reggelin, A. Nahhas, S. Lang, and M. Müller, "Assigning dispatching rules using a genetic algorithm to solve a hybrid flow shop scheduling problem," *Procedia Manufacturing*, vol. 42, pp. 442–449, 2020.

[41] E. Bakhtavar, M. Valipour, S. Yousefi, R. Sadiq, and K. Hewage, "Fuzzy cognitive maps in systems risk analysis: a comprehensive review," *Complex & Intelligent Systems*, vol. 7, no. 2, pp. 621–637, 2021.

[42] M. J. Rezaee, S. Yousefi, M. Valipour, and M. M. Dehdar, "Risk analysis of sequential processes in food industry integrating multi-stage fuzzy cognitive map and process failure mode and effects analysis," *Computers & Industrial Engineering*, vol. 123, pp. 325–337, 2018.

[43] M. J. M. Ladeira, F. A. F. Ferreira, J. J. M. Ferreira, W. Fang, P. F. Falcão, and A. A. Rosa, "Exploring the determinants of digital entrepreneurship using fuzzy cognitive maps," *The International Entrepreneurship and Management Journal*, vol. 15, no. 4, pp. 1077–1101, 2019.

[44] F. Iqbal, J. M. Hashmi, B. C. M. Fung et al., "A hybrid framework for sentiment analysis using genetic algorithm based feature reduction," *IEEE Access*, vol. 7, pp. 14637–14652, 2019.

Hindawi
Computational Intelligence and Neuroscience
Volume 2022, Article ID 6377043, 12 pages
https://doi.org/10.1155/2022/6377043

Research Article

Network Architecture for Intelligent Identification of Faults in Rabbit Farm Environment Monitoring Based on a Biological Neural Network Model

Hanjie Zhang [1,2,3] **and Shuqu Qian** [1]

[1] *Anshun University School of Mathematics and Computer Science, Anshun, Guizhou 561300, China*
[2] *Guizhou University of Finance and Economics, Guiyang, Guizhou 550026, China*
[3] *University of Sao Paulo, Brazil*

Correspondence should be addressed to Hanjie Zhang; zhanghj@mail.gufe.edu.cn

Received 27 June 2022; Revised 9 August 2022; Accepted 16 August 2022; Published 10 September 2022

Academic Editor: Ning Cao

Currently, livestock and poultry farming is gradually developing towards modernization and scale, and closed livestock and poultry farms are widely used for poultry feeding management, but at the same time, the farming risks of large-scale farms are increasing. In this paper, based on the study of wireless sensor networks and biological neural network models, the environmental factors that mainly affect the growth of domestic rabbits are analyzed as an example, and the technology is used to design and implement an environmental monitoring system for modern farms. The design of the system is divided into three main parts: hardware design of each node, software design, and upper computer monitoring software design. The hardware part of the system uses coordinator nodes, router nodes, sensor nodes, and control nodes to form a wireless sensor network in the farm, carries out the hardware circuit design of each node, and based on the protocol stack, designs the software program of each node to realize the collection, transmission, and regulation of environmental information in the farm. In the upper computer part, the design and development of the upper computer monitoring software interface are used to complete the real-time display of environmental data, historical query, database storage, and curve drawing, and to design a remote client data query system based on the architecture to realize the query of environmental data of the farm by remote users and to carry out monitoring fault intelligent identification alarm. At the same time, the paper investigates the optimal deployment of wireless sensor network nodes and searches for the optimal location of sensor nodes through an improved biological neural network algorithm to maximize the network coverage and reduce the coverage of blind areas, and conducts simulation experiments with the coverage rate of a rabbit farm as the optimization target.

1. Introduction

In the breeding process, livestock and poultry are particularly sensitive to the environment of farm areas (rabbit farms, cattle sheds, pig pens, etc.), and an unsuitable environment can lead to livestock diseases. It is difficult to perceive environmental information in a timely and effective manner by manual means alone. A low level of information in the farming industry can reduce animal welfare and increase animal disease rates [1]. Many farms have substandard feeding environments and insufficient levels of information technology, ignoring the impact of the environment on the growth of livestock, which leads to poor quality of farm products. The requirements of livestock on the growth environment are very strict; in the case of rabbits, for example, rabbits in different growth cycles have different requirements for temperature, humidity, light, and harmful gases. Especially for female rabbits during pregnancy, the requirements for air quality, relative humidity, ventilation, wind speed, and lighting are extremely stringent. An unsuitable environment can lead to low immunity and resistance, slowed growth, and a lower reproduction rate in

rabbits. It may even lead to dry and cracked skin and mucous membranes, causing various skin and respiratory diseases, etc., increasing the rate of disease and mortality in rabbits and greatly reducing the yield of farmed products [2]. The merits of the livestock farm environment directly affect the occurrence and spread of livestock and poultry diseases, and the environment around the farm is also closely related to the healthy growth of livestock and poultry and the improvement of product quality. Actually, closed farms need to be insulated in cold weather, but if the air in the poultry house is not smooth, it will cause the index of harmful gases such as carbon dioxide, hydrogen sulfide, and ammonia to exceed the standard, and the temperature and humidity fluctuations in the house will also be large, and the exceedance of these environmental indicators will cause various stress reactions in animals, resulting in lower immunity and triggering various infectious diseases. Therefore, it is necessary to monitor the environmental conditions of farms in real-time and accurately, which not only provides a reference basis for the future regulation of the poultry house environment but also can effectively prevent and control the spread of animal diseases and improve the quality of livestock and poultry products [3].

In this project, we adopt wireless sensor network technology combined with a biological neural network model to design and implement a farm environmental monitoring system with the functions of real-time display, storage, history query, and over-limit alarm for environmental information. The CSPNet architecture is used as a backbone to extract features at three different scales from the rabbit target, and the PANet + architecture is used to generate a feature pyramid in the neck section to aggregate features, which is finally sent to the head module to obtain the output with confidence and box coordinate information using anchored boxes on the feature map. The optimal deployment of wireless sensor network nodes is also studied, and the optimal location of sensor nodes is searched by an improved bio-neural network algorithm to maximize the network coverage and reduce the coverage blind area, and a simulation experiment is conducted with the coverage rate of a rabbit farm as the optimization target.

2. Related Work

In general, environmental factors such as light, temperature, humidity, and harmful gas concentration in the farm area are important to the growth of livestock and poultry, so comprehensive monitoring and reasonable control are required. Sensor networks usually have a large number of sensor nodes. How to optimize the deployment of sensor nodes to achieve maximum network coverage, reduce network resource waste, and improve performance is of great importance. The network quality is also a concern.

Air [4] was collected from livestock farms and analyzed for its quality. The results showed that changes in the concentration of carbon dioxide in farms can have different effects on the growth and reproduction of livestock and poultry, etc. Finally, a reasonable mathematical model was proposed based on mathematical modeling to estimate its emission. A powerful chicken house control software package [5] was designed to meet a variety of needs, consisting of a variety of control application modules, which can flexibly regulate the chicken house environment; a wireless sensor detection device for rabbit farm odor was developed [6], which can be used not only to monitor the breeding environment but also to analyze the health status of livestock and poultry-based on odor, health status, etc. To reduce the occurrence of diseases such as foot-and-mouth disease in rabbits, Bumanis et al. [7] prevented the occurrence of diseases by installing wireless sensor networks in their poultry houses to analyze and compare the environmental parameters collected in real-time. In the literature [8], in response to the problem that most current greenhouse monitoring system solutions are limited to small-area temperature control system applications and do not apply their routing functions to extend to large-area applications, a solution design of wireless greenhouse monitoring system based on multi-level routing is proposed to solve the problem of monitoring large-area greenhouses for agricultural use. The multi-parameter recorder of livestock barn designed in literature [9] can monitor a variety of environmental data at the same time, is easy to use, and can save historical data for a long time, but its monitoring range of the environment is small; Long et al. [10] [11] Included the temperature, humidity, atmospheric pressure, and harmful gases in the barn into the monitoring range by deploying the barn environment monitoring system. The system has more sensor nodes, but has not studied the deployment location of sensor nodes; at the same time, the environment monitoring system of livestock and poultry houses based on bus is studied. Compared with the traditional environment monitoring system of livestock houses, the system improves the real-time and reliability of the monitoring network, but the layout of communication lines is more cumbersome, which is not conducive to network updating and easy to be damaged by animals. In the literature [12], a wearable sensor was used to identify the presence of diseases in individual cattle by recognizing their behaviors such as lame walking and head shaking, so that early intervention could be made to reduce the morbidity rate. In the literature [13], a system based on WSN technology was designed to collect and monitor livestock activity data for outbreaks of livestock diseases such as foot-and-mouth disease and avian influenza, and sensors were deployed on farms to diagnose and prevent the occurrence of diseases by comparing the environmental parameters of livestock before and after the disease. Chen and Yang [14] developed a suspected sick rabbit behavior monitoring system using UWB technology. The system uses a hardware platform constructed with an infrared thermometer and indoor precise positioning to monitor the walking trajectory, speed, and direction of rabbits around the clock and monitor whether there are abnormal behaviors such as blind movement and violent advance and retreat to confirm whether the rabbits are suspected to be sick. However, such rabbit identification studs can cause alarm and stress reactions to rabbits when worn, and may also cause discomfort when worn, which reduces animal welfare and is not conducive to welfare

farming. Chen [15] designed an automatic monitoring system for sick rabbits' behavior, which monitors rabbits' excretion behavior, records the number of rabbits' excretion, and integrates wireless communication, image processing, and embedded technology to locate suspected sick rabbits with abnormal behavior through algorithms. The images are eventually transmitted to the monitoring center, replacing the traditional human recording method. Weng et al. [16] integrated an individual model of poultry determined in advance and a house regulation model to establish an intelligent monitoring management system for livestock and poultry breeding, which can collect environmental parameters while comprehensively evaluating the optimal solution for temperature regulation under different price costs of feed, fuel, and electricity in the growing environment of chicken houses, calculating the ventilation duration as well as the heating time to reduce capacity costs and expand economic benefits. The spray cooling system developed in the literature [17] effectively reduces the concentration of respirable dust and ammonia in chicken houses by spraying water when the concentration of harmful gases is monitored to exceed the standard, which greatly reduces the exposure of farm workers to poultry dust and stops the occurrence of comprehensive inflammatory diseases in the lungs such as endotoxin.

3. The Overall Design of Network Architecture for Intelligent Identification of Environmental Monitoring Faults in Rabbit Farms

3.1. A Multiobjective Rabbit Detection Network Based on Biological Neural Networks. Achieving accurate and fast detection of multi-target domestic rabbits in natural scenes is an important foundation for subsequent multi-target tracking and behavior recognition. To address the problems of the high complexity of existing algorithms, low detection accuracy, and susceptibility to interference by open cattle farm environment and rabbit occlusion, this study proposes an EYOLOv5 algorithm, based on the efficient target detection network YOLOv5, which achieves high-dimensional feature screening and global information aggregation by introducing SA attention module and MSHA module; introduces PANet + to achieve rich scale semantic feature interaction; CIOU loss and weighted DIOU-NMS are used to achieve accurate screening of overlapping target regression frames, to achieve accurate and rapid detection of domestic rabbit targets in natural scenarios and lay the foundation for improving the modernization of domestic rabbit farming [18]. The screening of preselected boxes using weighted DIOU-NMS allows additional attention to the location information of the center points of the bounding boxes for regression, which facilitates the identification of targets with overlapping occlusions. Different degrees of occlusion between rabbit targets are common, and the introduction of LCIoU and weighted DIOU_NMS is more suitable for nonstructured farm rabbit target detection. To achieve accurate and fast detection of rabbit targets in unstructured farm environments, this study proposes a multiobjective detection framework for rabbits in natural scenes based on EYOLOv5, the specific structure of which is shown in Figure 1. The CSPNet architecture is used as the backbone to extract the features of rabbit targets at three different scales, and the PANet+ architecture is used to generate a feature pyramid in the neck part to aggregate the features, which is finally sent to the head module to obtain the output results with confidence and box coordinate information using anchored boxes on the feature map.

Its specific modules are described as follows:

(1) Backbone module: based on the CSPNet architecture of YOLOv5, the SA attention module and MSHA module are introduced to achieve the filtering and global information aggregation of high-dimensional features.

(2) Neck module: introduction of PANet + module to achieve rich scale semantic features based on the introduction of bridging to achieve mid-level scale feature compensation.

(3) Head module: regression of target frames of rabbits of different sizes using 1×1 convolution on 3 different scale features of the output, which is the same as YOLOv3 and YOLOv4.

(4) Loss module: YOLOv5 uses GIOU Loss as the loss function of the regression frame, and for the unstructured farm environment, this study uses CIOU Loss and weighted DIOU-NMS to achieve accurate screening of overlapping target regression frames.

The main modules of the EYOLOv5 framework include four modules, Conv, Focus, BottleneckCSP, and SPP. Both YOLOv5 and YOLOv4 use CSPNet structure as the backbone to extract rich information features from the input images. The introduction of the CSPNet paradigm module in the backbone network can increase the learning capability of the CNN while reducing computational bottlenecks and memory costs, which help the model maintain accuracy while being lightweight. Specific details of the backbone feature extraction network for EYOLOv5 are as follows:

(1) In the case that the original input image is $3 \times 640 \times 640$, the focus structure is entered first, and the four adjacent positions of the image are stacked to focus the W and H dimensional information to the C-channel space, which is conducive to improving the perceptual field of each point and reducing the loss of original information, and the final output feature map is $32 \times 320 \times 320$.

(2) Enter the superposition module of Conv + BottleneckCSP, each Conv module (Conv + BN + Relu) has a convolution kernel of size 3×3 and stride of 2, which plays the role of downsampling. The BottleneckCSP module implements the extraction of depth features with the same input and output, because the backbone part has 4 CSP modules, the input image is 640×640, so the pattern of feature map change is 640-320-160-80-40-20.

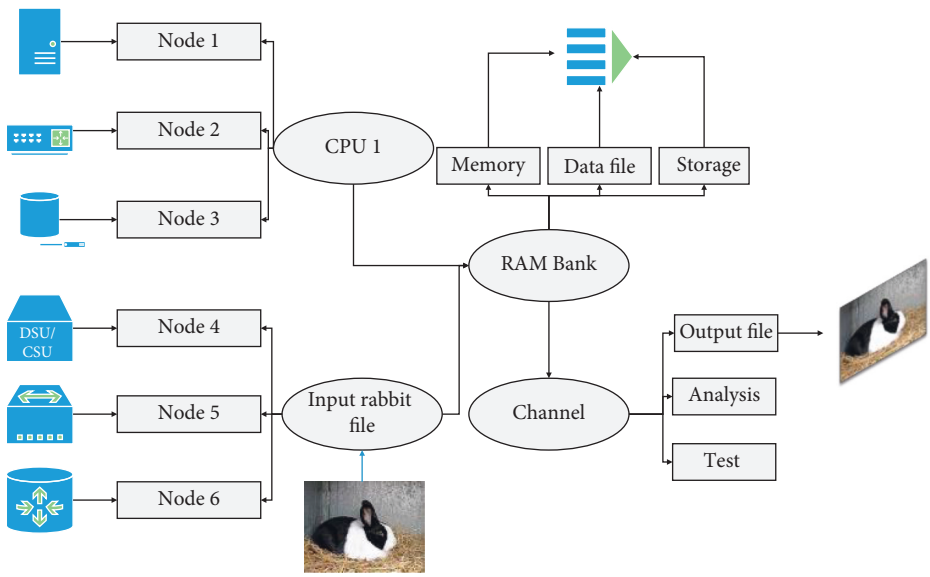


Figure 1: Biological neural network-based multi-objective detection framework for rabbits.

(3) The SA module is introduced after the last down-sampling of Conv to realize the filtering of channel features and spatial features, and then the SPP module is fed to realize the fusion of multi-scale features with different sensory fields.

(4) The multi-headed self-attentive module is introduced in BottleneckCSP-T (Transformer) to realize the integration of high-dimensional global features. The final feature maps of layers 5, 8, and 12 ($128 \times 80 \times 80$, $256 \times 40 \times 40$, $512 \times 20 \times 20$) are obtained to feed into the subsequent modules to achieve accurate detection of rabbit targets at different scales.

The biological neural network model essentially abstracts the neuron as an RC circuit and views the cell membrane as a capacitance C_m, which is charged by an input current I. Since ions leak slowly through the cell membrane in a nonideal state, the EYOLOv5 model introduces a resistance R_m into the circuit to reflect the ion diffusion phenomenon in the cell [19]. This segmented nonlinear decreasing strategy improves the global search ability in the early evolutionary stage, and at the same time with the linear adjustment strategy of the acceleration factor, it avoids the premature convergence of the algorithm in the late stage, through which the global search ability and the local search ability can be better balanced, and the convergence speed and accuracy of the algorithm can be improved. Normally, when there is no current input, the membrane potential $V(t)$ is resting V_{est}, and only when the output from other neurons is received does the membrane potential $V(t)$ change. When the voltage at both ends of C_m exceeds the neuronal voltage threshold V_{th}, the neuron generates an action potential, followed by a discharge phenomenon of the capacitor, and the neuronal membrane potential returns to the resting state and remains in an absolute resting state for some time. During this time, the membrane potential is not affected by the input current, and the period is called the absolute resting period t_{ref}. This process graphically simulates the activation of a biological neuron and the delivery of pulses. It is shown in the following equation:

$$C_m = \begin{cases} \dfrac{I(t) \cdot R}{V(t) - V_{rest}}, V_{rest} < 10 \\ \\ \dfrac{dV(t)}{I \cdot dt}, V_{rest} \geq 10. \end{cases} \tag{1}$$

Since the membrane time, constant indicating the voltage delay time is introduced in the EYOLOv5 model

$$Ut_{ref} = \frac{1}{N} \sum_{j=1}^{N} Ut_i. \tag{2}$$

Thus, the membrane potential $V(t)$ of the neuron can be given by the following equation.

$$V(t) = \frac{C_m}{\sum_{i=1}^{M} Ut_{ref}} + I(t)R. \tag{3}$$

The external input current $I(t)$ is a weighted sum of the presynaptic neuronal pulse signals. The LIF model simulates the role of excitatory/inhibitory synaptic currents and leakage currents in the neuron by simplifying the various ion channels inside the neuron. The advantage is that it is computationally efficient, requiring only about 5 floating-point operations per millisecond for a single neuron simulation, and the dynamics of the synaptic excitatory/inhibitory currents are well defined, as shown in the following equation.

$$R_k = \sum_{i=1}^{k} f_k(w_k). \tag{4}$$

Because synaptic connections primarily carry the process of information transfer between neurons, many researchers

have attempted to find biological methods of classification by mimicking the structure of biological neural networks and the mechanisms of synaptic plasticity learning. Studies have shown that synapses can be strengthened or weakened over time, a process known as synaptic plasticity. Synaptic plasticity is believed to be the basis of learning and memory in biological nervous systems. If the neurons at both ends of a synapse are activated at the same time, the connection between them will be enhanced; conversely, if both neurons are not always activated at the same time, the connection between them will be weakened.

In addition, this study utilizes weighted DIOU-NMS for preselected box screening, which can additionally focus on the location information of the center point of the bounding box for regression and facilitate the identification of targets with overlapping occlusions [20]. Different degrees of occlusion between rabbit targets are common, and the introduction of LCIoU and weighted DIOU_NMS is more suitable for nonstructured farm rabbit target detection.

3.2. Network Architecture Design for Intelligent Identification of Environmental Fault Monitoring in Rabbit Farms. Many factors such as temperature and humidity, ammonia, and other harmful gas concentration in the livestock growing environment will directly affect the growth condition of livestock. Due to the concentration of livestock, livestock barns are closed, poor air circulation, light is restricted, and toxic and harmful gases in the livestock barn are difficult to be eliminated quickly, all of which will cause adverse effects on livestock growth. In short, the five main factors that affect livestock growth are as follows:

(1) *Temperature Influence*. The temperature required by livestock at different stages of growth is different, and too high or too low a temperature can lead to livestock morbidity and even cause the spread of infectious diseases, resulting in the death of livestock. Especially young livestock need to be able to adjust their temperature flexibly because their thermoregulatory mechanisms are incomplete. For example, the optimal temperature range for pregnant rabbits is 18–21°C, for lactating rabbits is 20–22°C, and for lactating rabbits is 29–33°C. When the temperature is at the critical value, it will affect growth and production. Therefore, the temperature must be strictly controlled.

(2) *Influence of Humidity*. Air humidity and temperature together affect the growth of livestock. When the temperature in the livestock house is in the normal range, livestock room humidity on livestock growth has no significant impact; but in the high-temperature environment, humidity on livestock will grow adverse effects; when the livestock room humidity is low, livestock will be resistant to decline or even cause disease. The relative humidity of 60%–70%, more suitable for the growth of livestock, in practice can be relaxed to 50%–80%. In addition to livestock houses, silage also needs to strictly control humidity. The moisture content of 60% is appropriate. The leguminous forage moisture content of 60–70% is appropriate; the coarse and hard raw material moisture content of 78–80% is good.

(3) *Light Intensity*. Light intensity plays an important role in the growth and production of livestock. The impact on livestock production and growth performance is mainly reflected in the light intensity and light time. Proper light helps promote metabolism, which is beneficial to the healthy reproduction of livestock and good for growth and development. In addition, natural light in ultraviolet light also has the effect of sterilization, the right amount of ultraviolet light can also increase the immunity of livestock, but too much light intensity will reduce their metabolism. Therefore, the intensity of light is one of the factors that must be controlled during the growth of livestock.

(4) *Harmful Gases*. The air quality in the farm environment plays an important role in the growth and production process of livestock. Ammonia has a high solubility among harmful gases and is easily dissolved in the mucous membranes of livestock. The body tissues of livestock will be strongly stimulated as a result, which not only makes them susceptible to respiratory diseases but also easily causes necrosis of animal tissues and a series of diseases such as paralysis of the central nervous system.

(5) *PH Value*. In the process of livestock word feed, especially silage storage, in addition to temperature and humidity, PH value is also a very important influencing factor. Good-quality silage PH value should be below 4 (3.8–4.2).

This system takes a rabbit farm as the research object, and each rabbit hutch of the farm is used as an independent control area [21]. The system consists of five parts: ZigBee data acquisition part, PLC control part, centralized control part of the upper computer, remote control, and solar power supply part. The general framework diagram of this control system is shown in Figure 2.

The data collection part consists of coordinator nodes, router nodes, and terminal sensor nodes. Each livestock house has a coordinator node to upload information and send commands; router nodes are used to increase the network coverage, and the network structure in this system is a tree structure for safe and reliable data transmission. An unsuitable environment can lead to problems such as low immunity and resistance, slowed growth, and reduced reproduction rate of rabbits. It may even lead to dry and cracked skin and mucous membranes, causing various skin and respiratory diseases, etc., increasing the disease rate and mortality rate of rabbits and greatly reducing the output rate of farmed products. The terminal node is connected to the sensor, responsible for completing the data collection, using an RF antenna to send the convergence to the coordinator node through the wireless network, and then transmitting

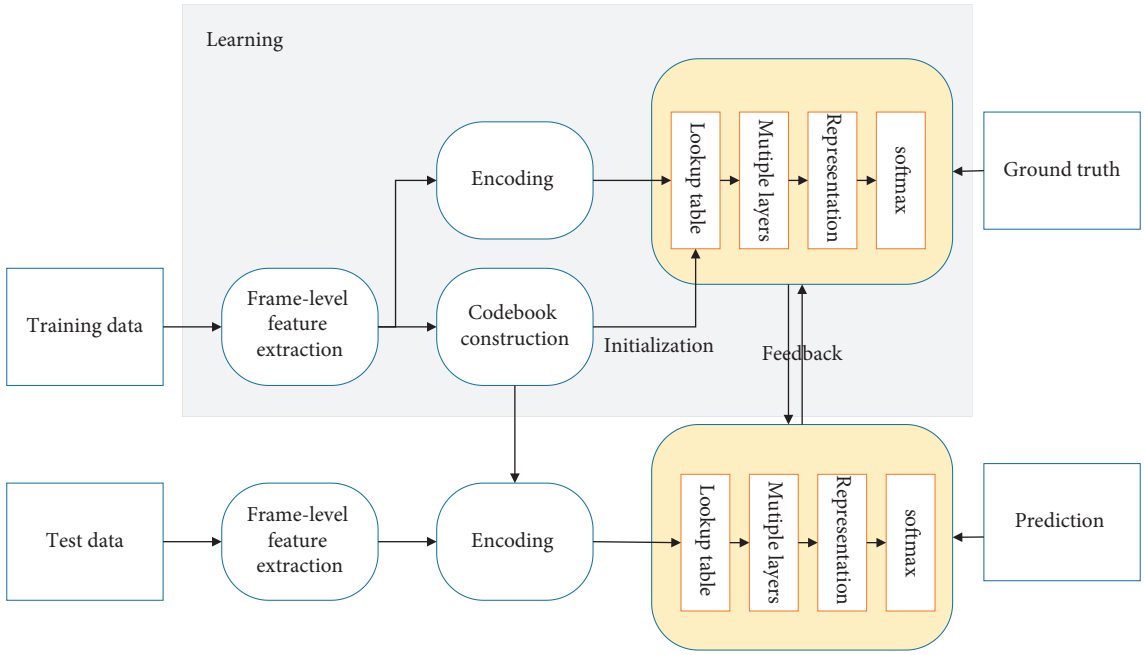


Figure 2: The general structure of the system.

the data to the upper computer for centralized display and control through the serial port. After the software analysis and processing of the upper computer, the data can also be transmitted to the PLC control center to further control the action status of each actuator.

The centralized control part of the upper computer mainly adopts the configuration software Configuration King to display and control the status of the system centrally through the configuration king screen, and monitor the safety of the farm and the real-time growth condition of livestock at any time. The main functions of this part include displaying the environmental parameters of the livestock house, the safety condition of the farm, the working status of the system and the setting of important parameters, report generation, and printing, as well as the historical data query function.

On the one hand, the PLC control part receives the problem data sent from the host computer and analyzes the data through the program, then controls the corresponding actuators to make the problem data normal; on the other hand, the PLC itself measures the temperature and humidity, PH value, and other data of the silage warehouse; selects a suitable control strategy; and controls the actuators such as fans and filling pumps through the inverter.

The cell phone remote control part is realized by using the cell phone and TC35 SMS module. GSM SMS module in hardware only needs to connect the upper computer with serial line RS232 and configure the equipment and parameters on Configuration King; then, it can communicate with serial data for remote control. When the farmer needs to control the farm's equipment remotely, he or she only needs to send an SMS to the SIM card number in the SMS module from his or her cell phone [22]. Similarly, when the farmer wants to know the situation on the farm, he or she only needs to use his or her cell phone to send a query SMS to the SIM card number in the GSM SMS module, and the situation on the farm will be sent to the user's cell phone in the form of an SMS.

The design uses an independent solar power supply system, which consists of solar panels, a battery bank, a charge and discharge controller, and an inverter. When there is a cloudy day or night when the solar energy supply is insufficient, the power supply will be switched to the battery pack, and when the battery supply has problems, it will be automatically switched to the general power grid.

4. Optimized Deployment of Sensor Nodes for Rabbit Colony Monitoring

Rabbits are generally flexible and mobile in rabbit hutches, making it difficult to monitor their precise activities, which can lead to a large amount of wasted resources if the nodes are not optimally deployed. In this paper, we propose a sensor node optimization algorithm for rabbit hutch distribution, which can significantly improve the application efficiency of sensors and reduce the waste of resources. A wireless sensor network is composed of a large number of sensor nodes deployed in the monitoring area, and its main task is to send the local information data monitored by the sensor nodes to the base station [23]. The optimal coverage of the network is an important issue in the wireless sensor network. Since each sensor node is usually randomly placed in the target area. Reasonably deploy a limited number of sensor nodes in the monitoring area so that their monitoring range covers the monitored target area as much as possible, reduce the overlapping coverage of the target area, and ensure that the network collects information in a timely and comprehensive manner, which is of great significance to

reduce the waste of network resources and improve the network performance.

In general, the sensing model of a sensor node is reduced to a model. The commonly used sensing model in the two-dimensional plane is the sensing disc model, i.e., the sensing range of a sensor node is a circular area with the node as the center and the radius, and the value of the sensing radius B_κ is determined by the physical characteristics of the node's sensing unit.

$$UB_\kappa = \max_\kappa(P) + \eta\frac{(\max_\kappa(P) - \min_\kappa(P))}{dev}. \quad (5)$$

In early coverage studies of wireless sensor networks, the perception model is more often used, but it ignores the influence of the sensors themselves and external physical environment factors and assumes that the sensor nodes are deterministic about the monitoring of events, simplifying the coverage problem. However, in practical applications, due to the influence of noise interference and the node itself, the sensing range of sensor nodes is irregularly circular, and the closer the node is to the node, the better the node's sensing capability for the target. Therefore, the monitoring capability of the sensor node shows uncertainty, and its monitoring model is probabilistically distributed with certain characteristics, and the mathematical expression of the probabilistic perception model is:

$$LB_\kappa = \min_\kappa(P) - \eta\frac{(\max_\kappa(P) - \min_\kappa(P))}{dev}, \quad (6)$$

where η is the fluctuation value of sensing radius; dev is a measure of monitoring uncertainty of sensor nodes to represent the dynamic change of sensing radius of sensor nodes; P is used to describe the monitoring probability of nodes for events occurring at the target. The main task of a wireless sensor network is to accomplish the monitoring and information acquisition of the target; therefore, while maximizing the network coverage, it is important to take into account the limitations of network energy, number of nodes, and other resources to ensure the provision of reliable monitoring services, and to adjust the location of sensor nodes to enhance the network coverage according to the actual application requirements. For wireless sensor detection of rabbit colonies, the fence coverage method is generally used. The fence coverage is concerned with the monitoring capability of the moving target. When the moving target crosses the monitoring area along a certain trajectory, the sensing range of the sensor nodes should be able to cover the whole moving trajectory of the moving target, which differs from the target (point) coverage problem in that the location of the covered target is uncertain. The curve as shown in Figure 3 is the target moving path.

The advantage of the fence-covering algorithm is that it is simple and easy to implement and there are not many parameters to be adjusted. In the iterative process of the algorithm, it is always hoped that the particles can search the whole space and locate the approximate range of the optimal solution in the early stage, and can improve the convergence speed and search for the global optimal solution in the later stage. The linear decreasing strategy of inertia weights has been proposed. If the neurons at both ends of the synapse are activated at the same time, the connection between them will be enhanced; conversely, if the two neurons are not always activated at the same time, the connection between them will be weakened. This strategy makes the searchability of the algorithm stronger at the beginning, but if the optimal particle is not searched, then with the enhancement of the local search ability at the later stage of the algorithm, the algorithm is easy fall into local extremes and difficult to jump out, which will lead to premature aging of the algorithm [24]. To overcome this drawback and obtain a better optimization effect, this paper improves the algorithm by using a strategy of dynamic adjustment of inertia weights in segments and linear adjustment of acceleration factors in the following equation:

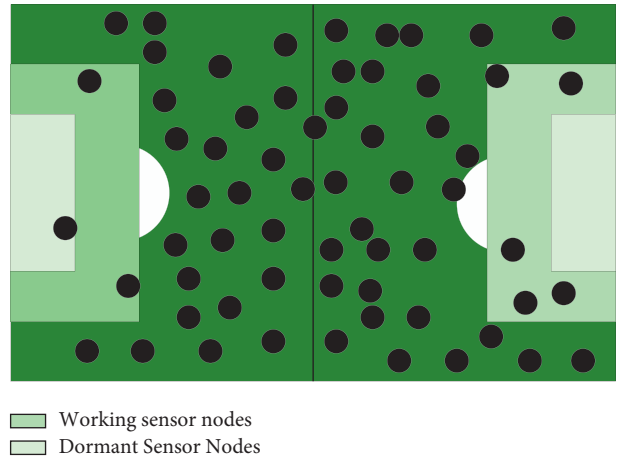


Figure 3: Fence overlay form.

$$\nabla_{b^t} J(w, b; x_i^*, y) = \delta^{l+1}\frac{\partial J(w, b)}{\partial b_{ij}^l} + \lambda\alpha_{ij}^l. \quad (7)$$

The basic idea of this improvement strategy is: by establishing a nonlinear function to dynamically adjust the value of inertia weights $J(w, b)$ so that the pre-evolutionary stage slowly decreases, the particle explores the whole feasible solution space as much as possible with a larger speed to discover a better region; in the post-evolutionary stage, the decreasing trend accelerates, the particle speed x_i^* becomes slower and starts to conduct a fine local search. Meanwhile, a linear strategy is used to adjust the value of the acceleration factorλ so that the algorithm has a large at the early stage of the search, and the motion of the particle mainly refers to its historical information, and at the later stage of the search by increasing to pay more attention to the population information, so that the particle ensures a certain search speed and avoids premature convergence. The design flow of the optimal deployment scheme of sensor nodes based on rabbit colony monitoring is shown in Figure 4.

This segmented nonlinear decreasing strategy improves the global search ability in the early evolutionary stage, while with the linear adjustment strategy of the acceleration factor, the premature convergence of the algorithm in the late stage

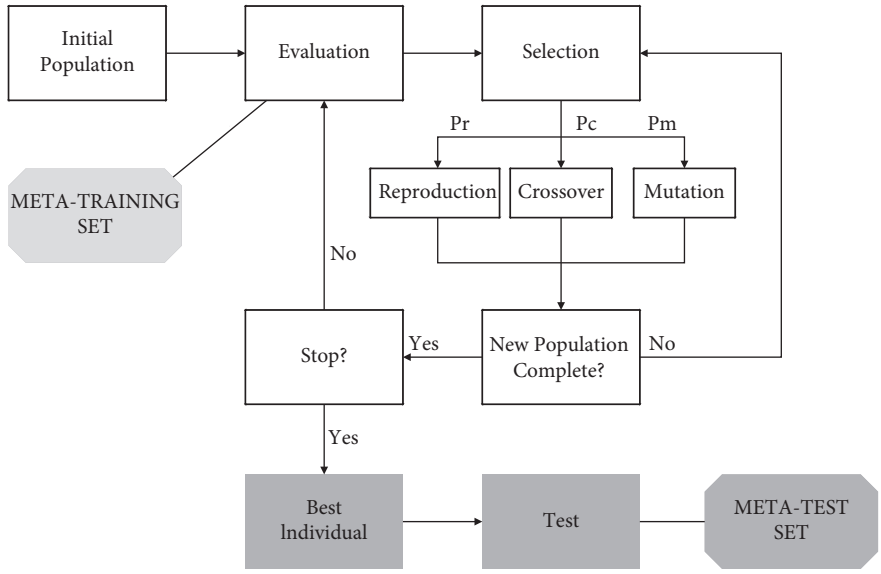


Figure 4: Design flow of the optimal deployment scheme based on rabbit colony monitoring sensor nodes.

is avoided, through which the global search ability and the local search ability can be better balanced, and the convergence speed and accuracy of the algorithm are improved.

5. Experimental Verification and Conclusion

5.1. Monitoring Sensor Node Optimization Deployment Validation. Figure 5 shows the distribution of sensor nodes after the standard algorithm and the algorithm optimization, respectively. By comparison, it can be seen that the distribution of sensor nodes after the standard algorithm optimization has been improved, but there are still some duplicate coverage areas, while the distribution of nodes after the algorithm optimization is more uniform, and there is no duplicate coverage.

As shown in Figure 6, the initial coverage of the sensor network is 85%; using the fastest convergence speed of the algorithm, the algorithm has converged to the global optimum at 200 iterations, and finally, the effective coverage of the optimized network can reach 94.3%; at 64 iterations of the algorithm, the optimized effective coverage is higher than the original sensor deployment method. This is because the particles traverse the entire search space at a larger search speed as the blank range slowly decreases in the early stage of the search, which improves the probability of searching for the optimal solution. In contrast, the standard algorithm is used to gradually converge to the optimal solution only when iterating to generations, and the network coverage after 126 iterations is 95.7%. This is because the nonlinear decreasing strategy of inertia weights is used in the algorithm, which makes the global search ability at the beginning of the iteration larger, and the particles can traverse the whole search space at a larger speed to determine the range of the optimal solution; the blank range decreases rapidly at the late iteration, and the particles begin to gradually converge to the optimal region for fine local search. At the same time, with the introduced linear adjustment strategy of the acceleration factor, the particles can still maintain a certain search speed in the late stage of the search, which avoids the premature convergence of the algorithm. Through the simulation results, we can get that the final network coverage of 99.8% is obtained with only 23 sensor nodes deployed, and it is calculated by Equation 6 that at least 41 sensor nodes need to be deployed to achieve this coverage. Compared with the standard algorithm and algorithm, the improved algorithm can optimize the deployment of sensor nodes more effectively, improve the network coverage, and reduce the deployment cost of the network.

5.2. Validation of a Biological Neural Network Algorithm for Rabbit Colony Recognition. Different from the training method of traditional artificial neuron networks, when training the YOLOv5 network, it is necessary to set a fixed simulation time to deliver the pulse information. In this paper, we set the simulation time of each rabbit activity as 250 ms, of which 200 ms is the learning time and 50 ms is the resting time. During the resting time, the parameters such as membrane potential and synaptic conductance of neurons will decay to the initial value with time to prepare for the learning of the next rabbit activity. Since each sensor node is usually randomly placed in the target area, it is important to reduce the waste of network resources and improve the network performance by how to reasonably deploy a limited number of sensor nodes in the monitored area so that their monitoring range can cover the monitored target area as much as possible and reduce the overlapping coverage of the target area to ensure that the network can collect timely and comprehensive information in the target area. The first set of experiments investigates the effect of network size of SNN on the classification of rabbit population activity dataset when the YOLOv5 algorithm is used. This subsection uses

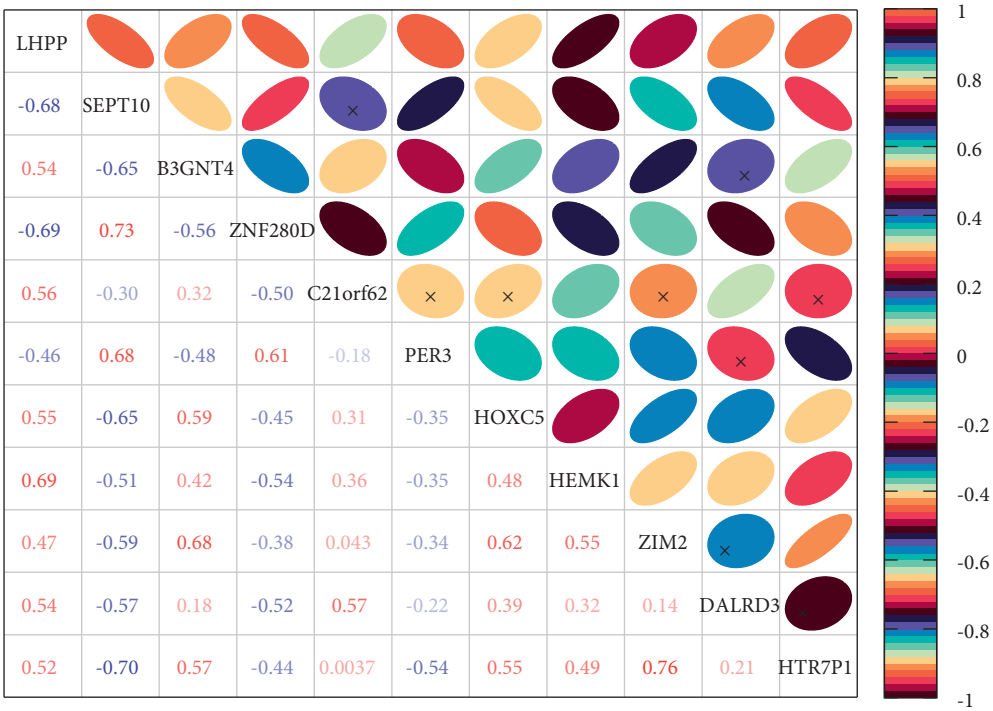


Figure 5: Distribution of nodes after algorithm optimization.

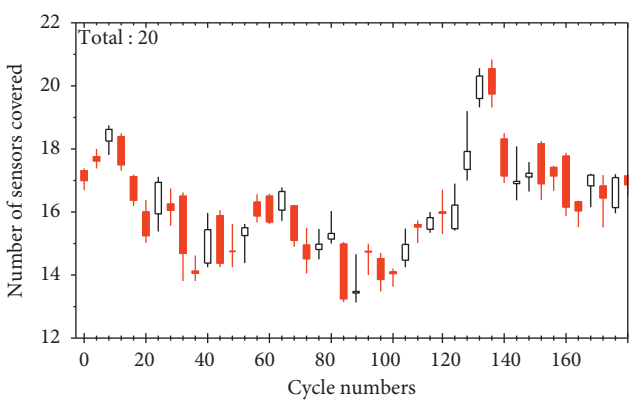


Figure 6: Optimized network coverage graphs.

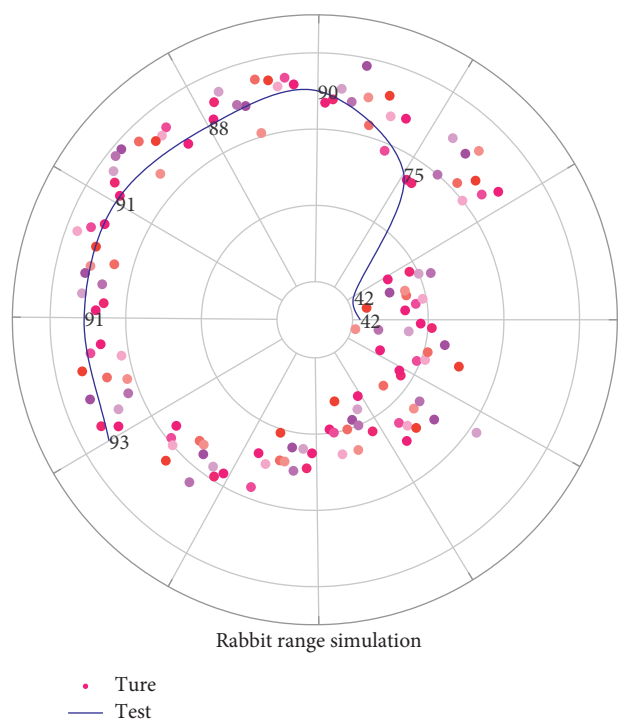


Figure 7: Variation of rabbit group activity recognition accuracy with the number of neurons.

100, 225, 400, and 625 excitatory neurons to train 60,000 samples in the MNIST training set, 10,000 samples each time, and the whole training set is trained five times to record the variation of the classification accuracy of the network with the number of neurons, as shown in Figure 7.

After 30 iterations, the classification accuracy of the network increases as the number of neurons in the network increases. The classification accuracy is 86.13% for the network with 100 neurons, 88.92% for the network with 225 neurons, 91.01% for the network with 400 neurons, and 91.48% for the network with 625 neurons. Although the classification accuracy of the network is improving with the increase of the number of neurons, the training time of the network is seriously increased because the number of training weights increases from 490000 to 705600 while the number of neurons is increased from 625 to 900, which increases 215600 parameters, so only a network of 625 neurons is trained in this paper. To visualize the features learned by the

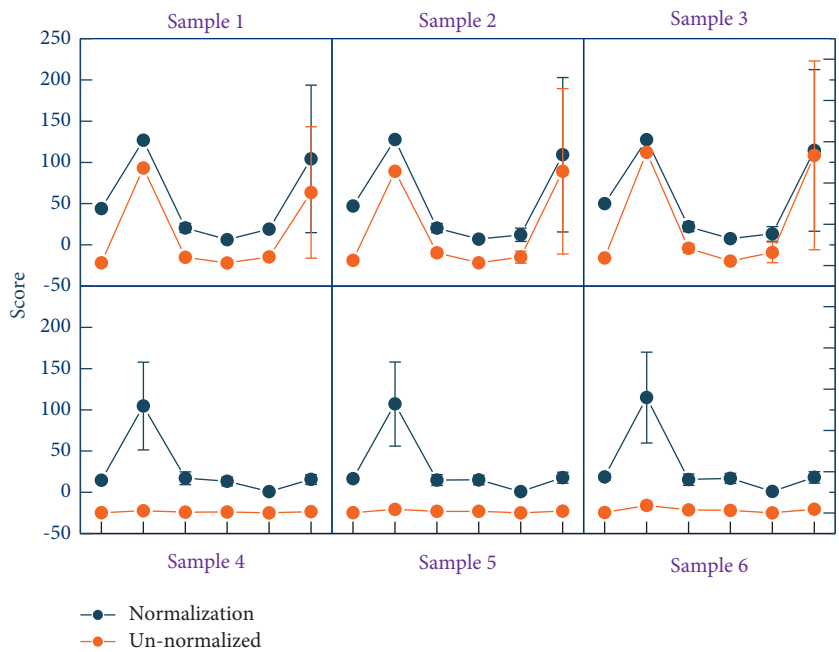


Figure 8: Effect of normalization on the excitation frequency of neurons.

excitatory neurons, for each excitatory neuron in the 625 scales, the 784-dimensional vector of weights connected to the encoding layer is rearranged into 28×28 matrices. The initial weights are random numbers between [0.01, and 0.2]. As the simulation proceeds, the distribution of weights of excitatory neurons gradually converges to their preferred input information.

The final MAP of the original network can be seen in Figure 8 as 50.55%, which will be used as a baseline value for subsequent comparison with the calculated value of SN. In the same way, the original mAP of the YOLOv3 network can be obtained as 65.36%. Next, the parameters are extracted from the converted ".h5" format file, the parsed weight file is obtained similarly to the target identification experiment, and the parameters are normalized for the detection experiment. As in the original ANN network, the membrane potential of the final output layer was decoded to obtain specific values, and the prediction frame was drawn on the input test map using the coordinate conversion formula and the confidence calculation formula in the paper. When conducting the first two experiments, the correlation between the activation values and the converted pulse frequency values was high, but if a deeper network was used for the detection task, the whole network would fail because all the neurons in the back layer were not activated and using different normalization methods, the neurons in the back layer of the network had different activation profiles. Without normalization, most of the neurons are underactivated, so they cannot transmit information at all. With only 23 sensor nodes deployed, the final network coverage of 99.8% is obtained, and it is calculated by equation (6) that at least 41 sensor nodes need to be deployed to achieve this coverage. Compared with the standard algorithm and algorithm, the improved algorithm can optimize the deployment of sensor nodes more effectively, improve the coverage of the network, and reduce the deployment cost of the network. After normalizing the parameters using the normalization factor p, although some of the neurons are saturated with activation, this is more beneficial to the final result compared to the most inactive state; after normalizing the channels, more neurons are activated, so more information can be transmitted. It helps a lot to improve the accuracy of the results.

6. Conclusion

The quality of the environment in livestock farms is an important factor affecting the healthy growth and product quality of livestock and poultry. Scientific and reasonable monitoring and control of the farm environment can not only create a good growth environment for livestock and poultry but is also significant for human health and improving modern production levels. This paper designs a technology-based farm environment intelligent monitoring system, which effectively solves the problems of the existing wired monitoring system, takes rabbits as an example to study the main environmental factors affecting their growth, and selects temperature, humidity, light, and ammonia as the environmental parameters to be monitored by the system. An optimized deployment method of sensor nodes based on an improved algorithm is also proposed to maximize the coverage of the wireless sensor network, which can effectively reduce the blind areas of network coverage and propose an effective optimization solution for the practical application of sensor node deployment in the future.

The system is easy to arrange, stable in transmission, and simple in operation. It can realize 24-hour remote monitoring and recording data of livestock and poultry breeding, save time and energy for breeders, improve job satisfaction and efficiency, reasonably manipulate control equipment according to reliable data, and ensure the health and quality safety of livestock and poultry products. The system has been tested in the field, and the test results show that after adding router nodes, the sensor network can effectively increase the communication distance while ensuring reliability, and the system can realize the collection, storage, and display of environmental information and can be alarmed and controlled according to the set value. The system meets the design requirements and can meet the needs of modern farm environmental monitoring intelligence.

Considering the development cost, this system only collects and monitors four types of environmental information, but many environmental factors affect the growth of livestock and poultry, and the monitoring of other environmental information can be added in the future according to the different types of livestock raised in the farm and the actual needs of users, to monitor the environmental quality in the farm more accurately.

Data Availability

The data used to support the findings of this study can be obtained from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work was supported by Anshun University School of Mathematics and Computer Science.

References

[1] M. Choi, "A study on the efficient Implementation method of cloud-based smart farm control system[J]," *Journal of digital convergence*, vol. 18, no. 3, pp. 171–177, 2020.

[2] J. Wang, M. Chen, J. Zhou, and P. Li, "Data communication mechanism for greenhouse environment monitoring and control: an agent-based IoT system," *Information Processing in Agriculture*, vol. 7, no. 3, pp. 444–455, 2020.

[3] Ö Köksal and B. Tekinerdogan, "Architecture design approach for IoT-based farm management information systems," *Precision Agriculture*, vol. 20, no. 5, pp. 926–958, 2019.

[4] X. Du, S. C. Bayliss, E. J. Feil et al., "Real time monitoring of Aeromonas salmonicida evolution in response to successive antibiotic therapies in a commercial fish farm," *Environmental Microbiology*, vol. 21, no. 3, pp. 1113–1123, 2019.

[5] A. Vij, S. Vijendra, A. Jain, S. Bajaj, A. Bassi, and A. Sharma, "IoT and machine learning approaches for automation of farm irrigation system," *Procedia Computer Science*, vol. 167, pp. 1250–1257, 2020.

[6] M. Liberloo, J. Bijttebier, E. Lacour, D. Stilmant, and F. Marchand, "Reflexive monitoring in action: setting up a monitoring system for learning effectiveness in agroecological farm innovation," *The Journal of Agricultural Education and Extension*, vol. 28, no. 2, pp. 231–250, 2022.

[7] N. Bumanis, I. Arhipova, L. Paura, G. Vitols, and L. Jankovska, "Data conceptual model for smart poultry farm management system," *Procedia Computer Science*, vol. 200, pp. 517–526, 2022.

[8] Z. N. Faysal and G. J. Mohammed, "Remote farm monitoring and irrigation system," *AL-Rafidain Journal of Computer Sciences and Mathematics*, vol. 15, no. 2, pp. 123–138, 2021.

[9] W. Guolong, D. Youzhi, and Z. Peng, "Inversion research in VOCs source emission of naphtha tank farm in a petrochemical enterprise in North China[J]," *Environmental Chemistry*, vol. 40, no. 6, pp. 1877–1884, 2021.

[10] J. Long, J. Mou, L. Zhang, S. Zhang, and C Li, "Attitude data-based deep hybrid learning architecture for intelligent fault diagnosis of multi-joint industrial robots," *Journal of Manufacturing Systems*, vol. 61, pp. 736–745, 2021.

[11] B. Yang, R. Liu, and E. Zio, "Remaining useful life prediction based on a double-convolutional neural network architecture," *IEEE Transactions on Industrial Electronics*, vol. 66, no. 12, pp. 9521–9530, 2019.

[12] B. Ji, X. Zhang, S. Mumtaz et al., "Survey on the internet of vehicles: network architectures and applications," *IEEE Communications Standards Magazine*, vol. 4, no. 1, pp. 34–41, 2020.

[13] X. C. Cao, B. Q. Chen, B. Yao, and W. P. He, "Combining translation-invariant wavelet frames and convolutional neural network for intelligent tool wear state identification," *Computers in Industry*, vol. 106, pp. 71–84, 2019.

[14] J. Chen and A. Yang, "Intelligent agriculture and its key technologies based on internet of things architecture," *IEEE Access*, vol. 7, pp. 77134–77141, 2019.

[15] W. Chen, "Intelligent manufacturing production line data monitoring system for industrial internet of things," *Computer Communications*, vol. 151, pp. 31–41, 2020.

[16] L. Weng, Y. He, J. Peng, J. Zheng, and X. Li, "Deep cascading network architecture for robust automatic modulation classification," *Neurocomputing*, vol. 455, pp. 308–324, 2021.

[17] Y. Zuo, Y. Wu, G. Min, C. Huang, and K. Pei, "An intelligent anomaly detection scheme for micro-services architectures with temporal and spatial data analysis," *IEEE Transactions on Cognitive Communications and Networking*, vol. 6, no. 2, pp. 548–561, 2020.

[18] Y. Han, B. Tang, and L. Deng, "An enhanced convolutional neural network with enlarged receptive fields for fault diagnosis of planetary gearboxes," *Computers in Industry*, vol. 107, pp. 50–58, 2019.

[19] L. Hu, Y. Miao, G. Wu, M. M. Hassan, and I. Humar, "iRobot-Factory: an intelligent robot factory based on cognitive manufacturing and edge computing," *Future Generation Computer Systems*, vol. 90, pp. 569–577, 2019.

[20] M. M. Alam and M. T. Islam, "Machine learning approach of automatic identification and counting of blood cells," *Healthcare technology letters*, vol. 6, no. 4, pp. 103–108, 2019.

[21] W. Qiu, Q. Tang, J. Liu, and W. Yao, "An automatic identification framework for complex power quality disturbances based on multifusion convolutional neural network," *IEEE Transactions on Industrial Informatics*, vol. 16, no. 5, pp. 3233–3241, 2020.

[22] X. Yan, Y. Liu, and M. Jia, "Health condition identification for rolling bearing using a multi-domain indicator-based

optimized stacked denoising autoencoder," *Structural Health Monitoring*, vol. 19, no. 5, pp. 1602–1626, 2020.

[23] Z. Zhang, Y. Xiao, Z. Ma et al., "6G wireless networks: vision, requirements, architecture, and key technologies," *IEEE Vehicular Technology Magazine*, vol. 14, no. 3, pp. 28–41, 2019.

[24] J. Ashraf, A. D. Bakhshi, N. Moustafa, H. Khurshid, A. Javed, and A. Beheshti, "Novel deep learning-enabled LSTM autoencoder architecture for discovering anomalous events from intelligent transportation systems," *IEEE Transactions on Intelligent Transportation Systems*, vol. 22, no. 7, pp. 4507–4518, 2021.

Hindawi
Computational Intelligence and Neuroscience
Volume 2022, Article ID 4816549, 9 pages
https://doi.org/10.1155/2022/4816549

Research Article

Lite-3DCNN Combined with Attention Mechanism for Complex Human Movement Recognition

Maochang Zhu, Sheng Bin, and Gengxin Sun

College of Computer Science & Technology, Qingdao University, Qingdao 266071, China

Correspondence should be addressed to Gengxin Sun; sungengxin@qdu.edu.cn

Received 12 July 2022; Revised 5 August 2022; Accepted 9 August 2022; Published 9 September 2022

Academic Editor: Ning Cao

Three-dimensional convolutional network (3DCNN) is an essential field of motion recognition research. The research work of this paper optimizes the traditional three-dimensional convolution network, introduces the self-attention mechanism, and proposes a new network model to analyze and process complex human motion videos. In this study, the average frame skipping sampling and scaling and the one-hot encoding are used for data pre-processing to retain more features in the limited data. The experimental results show that this paper innovatively designs a lightweight three-dimensional convolutional network combined with an attention mechanism framework, and the number of parameters of the model is reduced by more than 90% to only about 1.7 million. This study compared the performance of different models in different classifications and found that the model proposed in this study performed well in complex human motion video classification. Its recognition rate increased by 1%–8% compared with the C3D model.

1. Introduction

In recent years, with the rapid development of deep learning, computer vision has made rapid progress, and human action recognition has become a research field that has attracted much attention. Despite the continuous improvement of research in this field, there are still many challenges for complex human action recognition in videos.

A 3D convolution network (3DCNN) [1] is widely used in human motion recognition. It is improved based on 2D-CNN [2] and modeling time information through 3D convolution and 3D pooling operation to extract spatio-temporal details in videos. However, the video of complex human movement has complex semantics and a lot of redundant information, such as background clutter, occlusion, and high dimensional data, which bring a lot of difficulties to motion recognition. At the same time, the existing neural network based on 3D convolution has a colossal structure, which requires a lot of computing space and time due to its high requirements on hardware devices. Due to their complex network structures, these deep learning models are incompatible with devices with limited computing and storage space, such as smartphones, tablets, and PCS. Therefore, designing an efficient and lightweight motion recognition algorithm is very important.

In this study, the traditional three-dimensional convolutional neural network framework is improved to reduce the number of convolution kernels and the number of convolution operations. Meanwhile, $3 \times 3 \times 3$ pooling kernels are used for pooling operations, and all zero filling is not used. The self-attention mechanism is added in the final feature extraction stage to establish the connection between spatial pixels. Finally, the softmax layer is used to classify complex human movements. Experimental results show the effectiveness of the proposed algorithm. This method does not use complicated and computationally expensive networks, such as C3D-bidirectional LSTM Net [3] or PWCNet [4], to extract time features from test videos. Instead, the simplified C3D Net was used to extract spatial and temporal features through adequate data pre-processing and then integrated with the attention mechanism [5] to extract global features as much as possible with limited parameters and computation. The trainable parameters of the Lite-3DCNN network structure proposed in this study are reduced to

about one-thirtieth of the original C3D network. The main contributions of this work are as follows:

(1) Simplify the C3D network structure and propose a lightweight 3DCNN architecture for complex human motion classifiers.

(2) The self-attention mechanism is integrated to enhance the learning of dependent features and global features of video frame sequences.

(3) Compared with the traditional C3D network, the recognition accuracy of this method is slightly improved, and the number of parameters is significantly reduced.

2. Related Work

In recent years, human motion recognition based on video has become one of the most popular research fields in computer vision and pattern recognition [6]. It has various applications, such as surveillance, robotics, healthcare, video search, and human-computer interaction. Human motion recognition in the video involves many challenges, such as cluttered backgrounds, occlusion, viewpoint changes, execution rate, and camera movement. Over the decades, several technologies have been proposed to address these challenges.

The framework for action recognition can be divided into two types. One is to create a single network and combine two-dimensional CNN with an RNN. In literature [7], the author first uses a convolutional neural network to extract spatial features. The convolutional layer is followed by RNN (recursive recurrent neural network), which allows time information to flow in time steps. Then, time pooling is used to aggregate the features of all time steps to form video sequence features.

The other is the framework based on 3DCNN [8], which uses three-dimensional convolution to extract spatial features. For example, in reference [9], the author extended the convolutional neural network to 3D to automatically learn spatio-temporal features. Then, a recurrent neural network is trained to classify each sequence considering the time evolution of each time step's learning features. The authors of [10] proposed a method of deep learning to recognize human actions based on motion sequence information in RGB-D video. A new representation emphasizes the critical postures associated with each step. Features obtained from motion in RGB and deep video streams are input to the convolutional neural network to learn distinguishing features. Similarly, Wang and Dantcheva [11] trained and fine-tuned 3D ResNet [12] on the well-known FaceForensic++ dataset, which is an excellent motion recognition network [13]. In addition, generative adversarial networks (GAEL Net [14]) have also been used to design robust facial manipulation detectors. Therefore, researchers began designing more complex architectures to achieve higher detection accuracy. A method of combining 3DCNN with ConvLSTM was proposed in [15] and applied to human action recognition. The 3DCNN model proposed in [16] addresses a complex scene classification problem. It uses the spatial and temporal features of the video to classify scenes as helping or non-helping in natural disasters. The authors of [11] proposed to use exponential linear unit-3D convolutional neural networks to extract deep features of moving videos to represent videos. The ability of state-of-the-art video CNNs (including 3D ResNet, 3D ResNet, and I3D) to detect tampered videos is investigated in work [17]. The authors of [18] proposed a method for anomaly detection in crowd scenes. They offered a 3DCNN architecture and a 3D GAN for domain adaptation to reduce the domain gap. The authors of [19] proposed a method to extract kinematic pose features from 3D joint positions. It is used to classify Support Vector Machines (SVM) and Convolutional Recurrent Neural Networks (CRNN). Vehicle behavior recognition is performed using 3DCNN in the article [20].

These high-precision motion classifiers have huge network scale and complexity. When experiments are carried out on the Utd-MHAD dataset, both the decision level and feature level fusion methods produce higher identification accuracy than those using each sensor mode alone. The highest accuracy of the decision level fusion method [21] is 95.6%. However, it consists of about 27 M trainable parameters. Similarly, networks based on pre-trained VGG-16 [22], ResNet [23], 3D ResNet [12], and optical flow-based methods [24] are networks with high computational costs. Due to their large size and computing power, these efficient networks are incompatible with limited computing and space devices such as smartphones, personal laptops, and tablets. However, lightweight deep learning models are easier to train and less expensive to update when deployed on smartphones, personal laptops, and tablets.

Considering many real-life application scenarios, the deep learning action classification model has been widely used in PCs and personal laptops. Laptop computers carry out many human motion recognition scenarios, and the current configuration of laptop computers is often unable to achieve training and use a large structure of deep learning models. As a result, real-world applications place high demands on lightweight models. Therefore, this paper proposes a light 3D convolutional neural network (Lite-3DCNN) for complex human motion classification.

3. Proposed Method

The 3D convolutional network is an extension based on the 2D convolution, which adds the time dimension to the 2D convolution to fully use the timing information in the video, as shown in Figure 1. It is widely used in video classification and retrieval.

However, the traditional 3D convolutional network framework has huge parameters and requires high machine performance. These models' training and prediction stages consume a lot of time and computing power. At the same time, sports videos often contain high-level semantic information and a large amount of redundant data, and videos of different modes interfere with each other, making the model unable to accurately capture essential features in complex videos. This study optimizes based on the C3D framework, first reducing the number of convolution

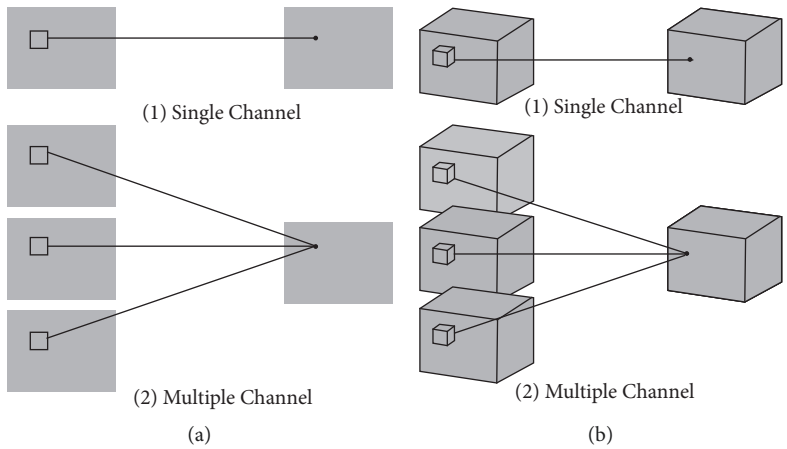


Figure 1: 2D Convolution (a) and 3D convolution (b) diagram.

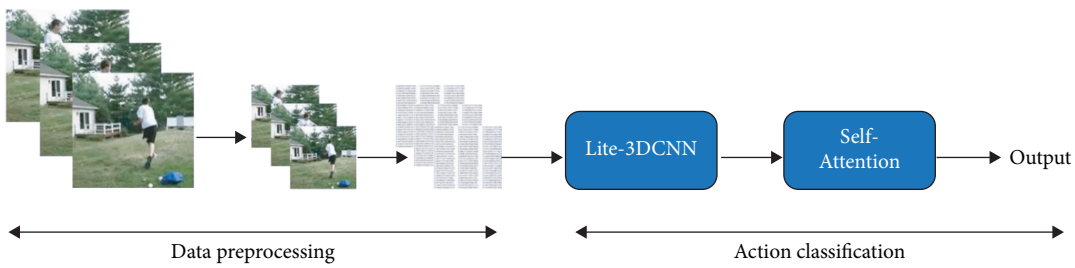


Figure 2: The overall process of this method.

operations and increasing the size of the pooling kernel, which significantly reduces the computational complexity. Then, a PreLU activation function with learnable parameters is used to avoid the problem of vanishing gradients. A self-attention mechanism is introduced into the model to extract long-distance interdependent features in complex videos, dramatically enhancing the feature extraction capability of lightweight 3D convolutional networks. In this study, data pre-processing is performed using average frame skip sampling and scaling. One-hot encoding is performed on the data, which enriches the training data features and makes the calculation between the elements more reasonable. The experimental results in Section 4 demonstrate that the method proposed in this study is suitable for complex human motion classification and is an efficient model that is convenient for training and prediction.

The proposed approach is divided into two phases, as shown in Figure 2. The first stage is the pre-processing stage. The video is clipped and scaled to continuous video frames, and then the four-dimensional matrix with depth is transformed. The four-dimensional matrix comprises the video frame's length, width, channel number, and depth. The transformed four-dimensional matrix plus batch_size is used as the input of the 3D convolution operation in the classification stage. In the second stage, the five-dimensional matrix calculated in the pre-processing stage is used as the input of Lite-3DCNN in the detection stage. The input five-dimensional matrix consists of the batch size, the width, and height of the video frame, and the depth and channel number, respectively, i.e., [batch_size, width, height, depth, channel]. Lite-3DCNN processes the input and learns the key features. Finally, an autonomous attention mechanism is added to enhance the learning of long-term dependent features.

3.1. Data Preprocessing. First, OpenCV was used to clip the video 20mmc20 times. Since the video data length is inconsistent with intercepting the image samples as evenly as possible, the clipping method uses average skipping frame sampling. The depth of a video frame is the number of times a video is clipped. $frames_{total}$ represents the total number of frames for a video, $de\ pth$ indicates the number of frames you want to intercept, and $output_{frames}$ means the video frame set after clipping, as shown in formula (1) and Figure 3,

$$output_{frames} = \sum_{i=0}^{depth} \frac{i * frames_{total}}{depth}. \quad (1)$$

Then, resize it to 32×32 and store all the processed images, including the video frame's width and height and height of the video frame and the number of channels. Finally, all the videos in each category are traversed, and then the four-dimensional array obtained after each video processing is combined to form a five-dimensional X.

As the input_shape format of Conv3d required, the data dimensions were adjusted to suitable inputs using the transpose method. Finally, the input data is X, the label is Y, and the label Y is processed by one-hot encoding [25], which makes the feature calculation among features more reasonable and improves the computing speed. The calculation method is shown in Figure 4.

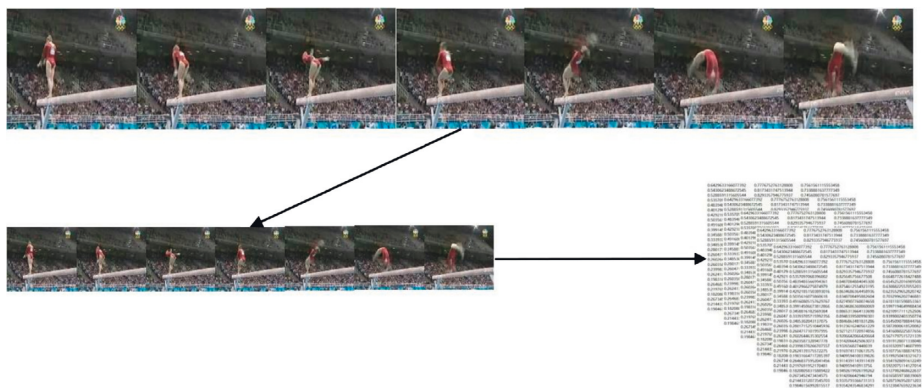


Figure 3: Data pre-processing.

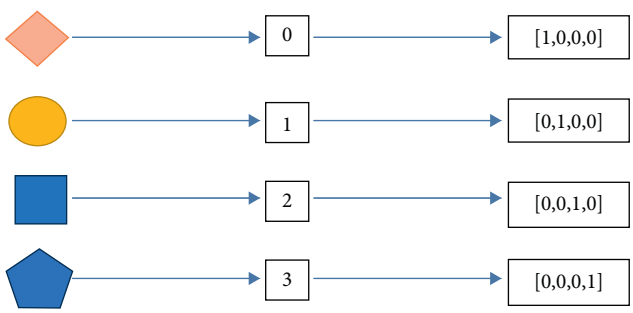


Figure 4: One-hot encoding.

3.2. Action Classification Model. The original C3D [8] network consists of five pool layers, eight 3D convolutional layers, and then two fully connected layers; the last one is the softmax layer for action prediction, in which the number of convolution cores in each convolution layer is 64, 128, 256, 256, 512, and 512. All pooling layers are maximum pooling, with the first pooling layer having a kernel size of 1×2 x 2 (in order not to merge time signals prematurely). The size of the remaining convolution kernel is 2×2 x 2, and the maximum pooling step is 1, which means that the size of the output signal is reduced by eight times compared with the input signal. The C3D network structure is shown in Figure 5.

The Lite-3DCNN network structure proposed in this study contains only four convolution operations, and the size of the convolution kernel at each layer is reduced to 32, 32, 64, and 64. At the same time, the ReLU activation function in the original network was abandoned in this paper. PreLU and softmax activation functions were used alternately after each convolution layer for activation operation.

According to Figure 6 and formula (2), the gradient of the ReLU activation function is 0 when $x < 0$. Hence, the rise of this neuron and subsequent neurons is always zero, which is gradient disappearance. In formula (2), in PReLU, the slope a_i of the negative part is not defined in advance but is constantly updated through backpropagation, as shown in formula (3). In this way, the problem of gradient disappearance can be solved, and the classification accuracy can be improved only by adding a few parameters,

$$\begin{cases} ReLU\,(\mathrm{x}) = \begin{cases} x, if\, x < 0 \\ 0, if\, x \geq 0 \end{cases} \\ PReLU\,(\mathrm{x}) = \begin{cases} x_i, if\, x_i < 0 \\ a_i x_i, if\, x_i \geq 0 \end{cases} \end{cases}, \tag{2}$$

$$\Delta a_i = \mathrm{U}\Delta a_i + \epsilon \frac{\partial \mathrm{y}}{\partial a_i}, \tag{3}$$

where **U** represents the momentum, ε represents the learning rate, and the initial $\mathbf{a_i}$ is 0.25. Experiments show that the PreLU function can accelerate model convergence and improve classification accuracy.

The proposed method only uses maximum pooling twice, changing the size of the pooling kernel to $3 \times 3 \times 3$ and further reducing the number of parameters on the premise of sacrificing a few features. To extract the most valuable elements from the limited number of features, the self-attention mechanism with 512 output dimensions was connected to the full connection layer with the same number of units before the softmax classification layer at the end of the model, and then normalized and finally sent to the output layer. The preprocessed data is input into the Lite-3DCNN network, and high-level semantic information is collected at a higher level of the deep convolutional network. Next, the Self_Attention network identifies long-term motion correlations from features extracted by 3D convolution. Therefore, the Lite-3DCNN coupled Self_Attention architecture proposed by us can better extract the spatio-temporal features of data while minimizing the time and space complexity. The complete network architecture information is shown in Figure 7.

Self-attention is borrowed from natural language processing, so it retains names like query, key, and value. The input convolution feature maps are feature maps extracted by backbone CNN. The structure of self-attention is divided into three branches from top to bottom: query key and value. Figure 8 is the basic structure of self-attention, and the calculation formula is shown in formula (5),

$$Attention\,(Q, K, V) = softmax\left(\frac{QK^T}{\sqrt{d_k}}\right)V, \tag{4}$$

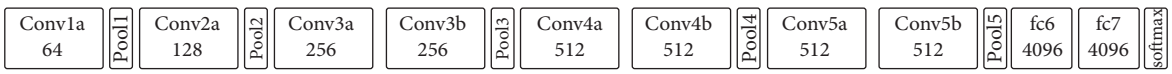


Figure 5: C3D network structure.

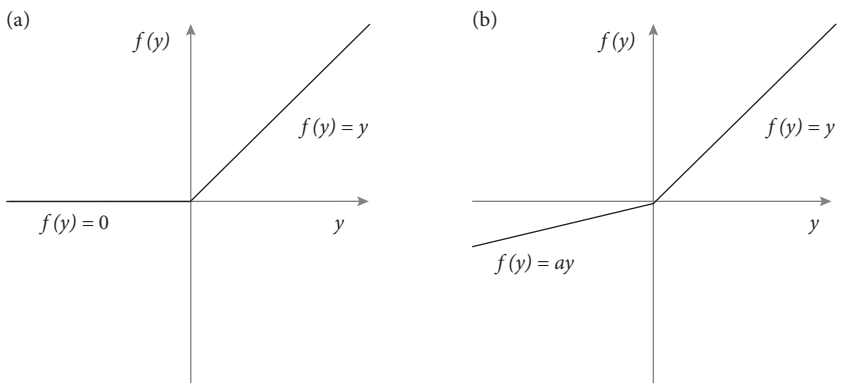


Figure 6: Activation function comparison. (a) ReLU (x) and (b) PReLU (x).

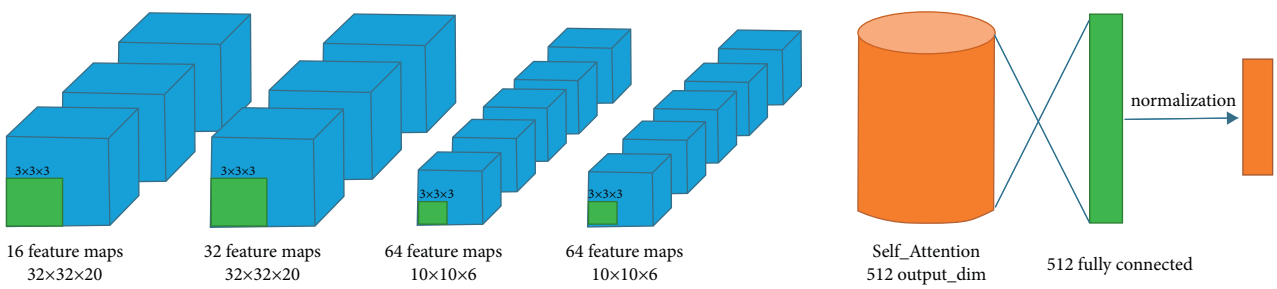


Figure 7: Action recognition architecture of this study.

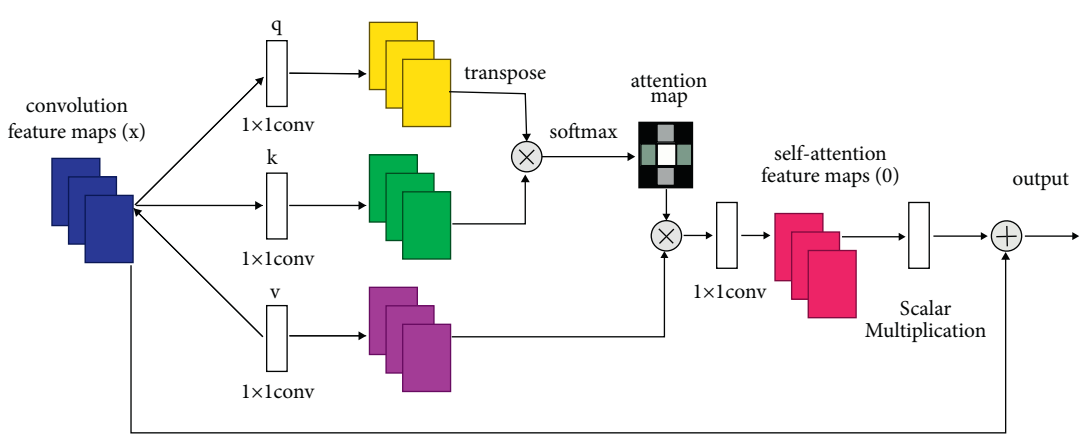


Figure 8: Self-attention mechanism unit.

where $Attention(Q, K, V)$ refers to the value of attention obtained and Q, K, and V are the query vector matrix, key vector matrix, and value vector matrix, respectively. Each row in these three matrices represents a corresponding vector. Q, K, and V are typically obtained by multiplying the input sequence X by three matrices, W^q, W^k, W^v.

First, for each sample, we have a d_q-dimensional query vector, forming an $N \times d_q n$-dimensional query vector matrix Q. You can think of the query vector as the characteristic of the model.

Then, for each piece of information (vector) in our "information base," there is a d_q-dimensional key vector and a d_v-dimensional value vector, forming a key-value pair. Suppose there are n_v pieces of information, then they constitute the key vector matrix K of n_v by d_v dimension and the value vector-matrix V of n_v by d_v dimension, respectively. You can think of key vectors as features of information and value vectors as the information content.

$QK^T_{n * n_v}$ represents the similarity between n query vectors (sample features) and n_v key vectors (information

Table 1: The network structure and parameters of this framework (20 class).

Layers	Output shape	Parameters
Input layer	32, 32, 20, 3	0
conv3d	32, 32, 20, 32	2624
activation	32, 32, 20, 32	655360
conv3d_1	32, 32, 20, 32	27680
activation_1	32, 32, 20, 32	0
max_pooling3d	10, 10, 6, 32	0
Dropout	10, 10, 6, 32	0
conv3d_2	10, 10, 6, 64	55360
activation_2	10, 10, 6, 64	0
conv3d_3	10, 10, 6, 64	110656
activation_3	10, 10, 6, 64	0
max_pooling3d_1	3, 3, 2, 64	0
dropout_1	3, 3, 2, 64	0
time_distributed (flatter)	3, 384	0
self__attention	3, 512	589824
Dense	3, 512	262656
batch_normalization	3, 512	2048
dropout_2	3, 512	0
global_average_pooling1d	512	0
dense_1	20	10260

features). For example, if we assume $n = 2$ and $n_v = 3$, then the first behavior [2, 3, 5] represents that the similarity between the first sample and the first, second, and third information is 2, 5, and 3 respectively.

We then apply an activation function $\omega\ (\cdot)$, typically softmax $(\cdot/)$, to obtain the correlation or similarity distribution $\omega\ (\mathrm{QK}^{\mathrm{T}}_{\mathrm{n} * \mathbf{n}_{\mathrm{v}}})$ between samples and pieces of information. For the previous example, we simply call $\omega\ (x) = [x_i / \sum x_i]$. The result is [0.2, 0.5, 0.3], which means that the correlation or similarity between the first sample and the first, second, and third information is 20%, 50%, and 30%, respectively.

Finally, multiply $\omega\ (\mathrm{QK}^{\mathrm{T}}_{\mathrm{n} * \mathbf{n}_{\mathrm{v}}})$ and $V_{\mathbf{n}_{\mathrm{v}} * \mathbf{d}_{\mathrm{v}}}$, and get $Attention(Q, K, V)$, that is, the weighted sum of the value vector (information); the weight is the distribution of correlation or similarity between each sample and each piece of information, and this is the final result of self-attention. The network structure and parameters of this framework are shown in below Table 1.

4. Experimental Discussion

4.1. The Dataset. The method was trained and tested on the UCF-101 dataset [24]. The dataset contains videos of different types of sports, such as handstand walking, canoeing, horse racing, etc. The UCF-101 dataset was generated from a collection of YouTube videos, with videos in 101 action categories divided into 25 groups, each of which can be composed of 4–7 action videos. Videos from the same group may have some standard features, such as similar backgrounds, similar viewpoints, etc. They are shown in Figure 9.

At the same time, it offers the most incredible variety in motion, with wide variations in camera movement, object appearance and posture, object proportions, viewpoint, cluttered backgrounds, lighting conditions, and so on, making it the most challenging dataset to date. The original dataset contains 13320 original videos and 50 related sports videos, of which 30 sports videos are randomly selected in this study.

4.2. Contrast Experiment. According to the nature of the deep learning model in this study, the video is first processed as video frames. To minimize clipping and retain relatively complete video features, the average structure hopping sampling method is adopted in the data pre-processing stage, and then the video frames are scaled. This paper extracts 20 RGB video frames from each video clip. Each video clip is fed individually into a Lite-3DCNN network stream with a frame size of $20 \times 32 \times 32$.

In the experiment in this paper, the initial learning rate of model training was set at 0.001, the PreLU activation function was used to accelerate model convergence, and the adaptive moment estimation (Adam) optimizer [26] was used during training, which combined the advantages of AdaGrad and RMSProp optimization algorithms. The update step size is calculated using the first moment estimation and second moment estimation.

In formula (5), β_1 is the exponential decay rate, controlling the weight distribution (momentum and current gradient), and β_2 is the exponential decay rate, maintaining the influence of the previous gradient square. t is a time step, initialized to 0. g_t is the gradient when the time step is t. θ is the parameter to be updated, and $f(\theta)$ is the random objective function of parameters. m_t is the first-order moment estimation of the gradient, and u_t is the second-moment estimation of the slope. $\check{m}_t$, $\check{u}_t$ is the correction of m_t and u_t, respectively. σ is the learning rate, and ε is a constant to maintain numerical stability.

The specific update rule is as shown in formula (5): initialize $\beta_1 = 0.9$, $\beta_2 = 0.999$, $\epsilon = 10\mathrm{e} - 8$, and $\sigma = 0.001$. The minimum batch of training is 32 samples for data training,

$$\begin{cases} t = t + 1 \\ g_t = \nabla_\theta f_t(\theta_{t-1}) \\ m_t = \beta_1 m_{t-1} + (1 - \beta_1) g_t \\ u_t = \beta_2 u_{t-1} + (1 - \beta_2) g_t^2 \\ \check{m}_t = \dfrac{m_t}{1 - \beta_1^t} \\ \check{u}_t = \dfrac{u_t}{1 - \beta_2^t} \\ \theta_{t+1} = \theta_t - \dfrac{\sigma}{\sqrt{\check{u}_t} + \epsilon} \check{m}_t \end{cases} . \tag{5}$$

Figure 9: Sports video dataset.

Figure 10: C3D model accuracy.

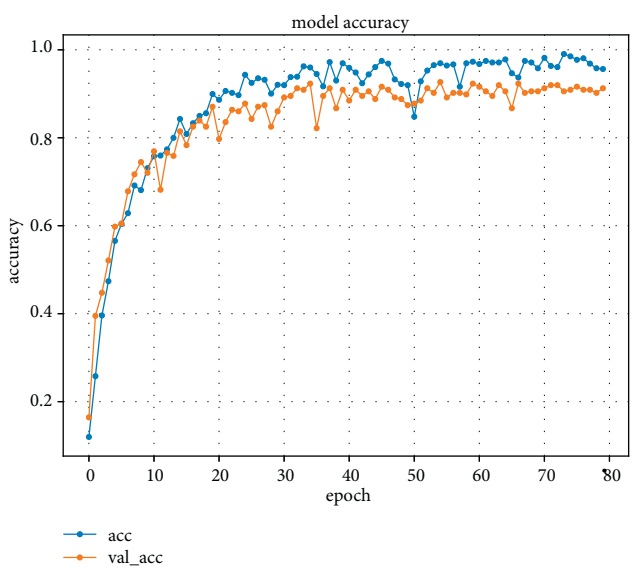


Figure 11: Proposed method accuracy.

Table 2: The validation accuracy of the proposed method for the complex human movement of the UCF-101 dataset.

Model	10 class (%)	20 class (%)	30 class (%)
C3D	82.2	84.7	83.5
Lite-3DCNN	85.3	80.2	70.6
Lite-3DCNN-LSTM	81.1	83.5	75.2
Lite-3DCNN-BiLSTM	84.5	85.3	79.5
Proposed method	**91.6**	**88.5**	**84.8**

Figures 10 and 11 show the change in the prediction accuracy of the C3D model and the model in this paper, respectively. The experiment carried out 80 rounds of training under 10 classifications and finally obtained the accuracy of the training set and the test set. It can be seen from the figure that the convergence speed of the C3D framework training is slow, and the curve rises erratically, resulting in oscillations. The overall trend of the method in this paper is rising and stable, the convergence speed is fast, and the final accuracy rate is about 9% higher than that of the traditional C3D model. The learning rate of these two methods is the same, so this may be because the C3D model cannot accurately capture the long-distance interdependence characteristics of complex actions, and the C3D model lacks normalization processing, resulting in singular values in the training process, which affect the speed and final accuracy of model learning.

In the comparative experiment, this study combined different network structures and verified the effectiveness of complex human motion classification on the UCF-101 dataset. The combination of lightweight 3D convolution and long short-term memory network is added in the experiment because, considering the reduced ability of the simplified C3D model to extract time series features, the LSTM network can well extract the context of video frames.

According to Table 2, the lightweight 3D convolutional network performs well on the 10-class classification problem, but the accuracy rate is significantly reduced with increasing the number of classifications. This may be because the depth of the lightweight 3D convolutional network is not enough. When faced with multi-classification tasks of complex motion, the lightweight model cannot extract

Table 3: The trainable parameters (in millions) of the proposed method and other methods for the UCF-101 dataset.

Model type	10 class (M)	20 class (M)	30 class (M)
C3D	52.87	61.30	61.34
Lite-3DCNN	1.609	1.616	1.621
Lite-3DCNN–LSTM (512)	3.120	3.122	3.135
Lite-3DCNN–BiLSTM (512)	5.219	5.224	5.229
Proposed method	1.712	1.716	1.884

richer features to distinguish different categories of videos. The performance of the classic C3D model is relatively stable, indicating that even if faced with more classification tasks, the C3D network architecture can still maintain a sure accuracy. Still, it needs to train more than 50 million parameters.

The combination of the Bi LSTM network and Lite-3DCNN has produced a specific result. Table 2 shows that the classification accuracy of the Lite-3dcnn combined with the LSTM framework is lower than that of the Lite-3DCNN model in the ten classification tasks. This is because when the number of classifications is small, the performance of the Lite-3dcnn model is good enough, and the advantages of the LSTM unit do not play a role. Still, the bidirectional LSTM unit extracts the information below the video frame to a certain extent, so it performs better than the C3D model. However, at 20 and 30 categories, the LSTM unit plays an advantage, making up for the simple structure of the Lite-3DCNN network. Even so, there is still no superior performance of the C3D model because the significant trainable parameters of the C3D model improve the ability of multi-classification tasks.

Experimental results show that the performance of the lightweight 3D convolutional architecture deteriorates with the increase in the number of classes. Although the method proposed in this paper also offers such a trend, the results are still better than the C3D framework on 30 classification tasks. The fundamental reason for this result is that the three-dimensional convolutional neural network can extract the spatio-temporal features of video data to a certain extent. The self-attention mechanism focuses on the global key features, increasing the receptive field with almost no increase in computational cost. Compared with the LSTM network, the self-attention mechanism and lightweight three-dimensional convolution network are better integrated, and more accurate prediction results are obtained.

4.3. Parameter Quantity Comparison. According to Table 3, the parameter amount of the method in this paper is only one-thirtieth of the C3D model. The accuracy is improved by about 4% in the 10–30 classification task. At the same time, in the case of adding a small number of parameters, the classification accuracy of the method in this paper is improved by about 9% on average compared with the lightweight 3D convolutional network, and the accuracy rate is still slightly higher than that of the C3D model when completing 30 classification tasks. This shows that the introduced self-attention method has played an important role, effectively making up for the deficiency of lightweight 3D convolution feature extraction capability. Although the combination of bidirectional LSTM and Lite-3DCNN produces some effect, the number of parameters is still about three times that of our model. Even under thirty categories, the trainable parameters of our model are only 1.884 M.

5. Conclusions

Complex human motion videos usually contain high-level semantic information and a large amount of redundant information. Although the classification framework based on the traditional three-dimensional convolution network can better complete the classification task, such a framework has many parameters. It requires a lot of time and computing power. This research introduces an efficient and lightweight human motion recognition framework, combining the lightweight C3D model and self-attention mechanism. The self-attention mechanism is used to capture critical global features. The receptive field is increased with only a few parameters, which makes up for the lightweight three-dimensional convolution network shortage. In the data processing stage, this study uses the average frame skipping sampling to reduce the data size as much as possible while retaining more complete features and uses the method of hot coding to enrich the data features and minimize interference. The experimental results show that, based on the ucf-101 dataset, the accuracy of the proposed method in the task of 10–30 classification is between 91.6% and 84.8%, which is about 5% and 10% higher than other models on average, and the parameter quantity is only one-thirtieth of that of the C3d model. However, the classification accuracy of the method proposed in this paper decreases slowly with the increase of categories, and no more classification experiments have been carried out in the study. In future research, we will consider combining the two-stream method and retraining in a more extensive dataset to improve the framework of this study further.

Data Availability

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

Conflicts of Interest

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

Acknowledgments

This work was supported by a grant from Shandong Natural Science Foundation of China (No. ZR2021MG006).

References

[1] K. Kamnitsas, C. Ledig, V. F. J. Newcombe et al., "Efficient multi-scale 3D CNN with fully connected CRF for accurate brain lesion segmentation," *Medical Image Analysis*, vol. 36, pp. 61–78, 2017.

[2] V. D. Hoang, D. H. Hoang, and C. L. Hieu, "Action recognition based on sequential 2D-CNN for surveillance systems," in *Proceedings of the IECON 2018-44th Annual Conference of the IEEE Industrial Electronics Society*, pp. 3225–3230, IEEE, Washington, DC, USA, 21-23 October 2018.

[3] S. Arif, J. Wang, A. A. Siddiqui, R. Hussain, and F Hussain, "Bidirectional LSTM with saliency-aware 3D-CNN features for human action recognition," *Journal of Engineering Research*, vol. 9, no. 3A, p. 2873, 2021.

[4] D. Sun, X. Yang, and M. Y. Liu, "Pwc-net: cnns for optical flow using pyramid, warping, and cost volume," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 8934–8943, Salt Lake City, UT, USA, 18-23 June 2018.

[5] H. Zhao, J. Jia, and V. Koltun, "Exploring self-attention for image recognition," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 10076–10085, Seattle, WA, USA, 13-19 June 2020.

[6] D. Wu, N. Sharma, and M. Blumenstein, "Recent advances in video-based human action recognition using deep learning: a review," in *Proceedings of the 2017 International Joint Conference on Neural Networks (IJCNN)*, pp. 2865–2872, IEEE, Anchorage, AK, USA, 14-19 May 2017.

[7] N. McLaughlin, J. M. Del Rincon, and P. Miller, "Recurrent convolutional network for video-based person re-identification," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 1325–1334, Las Vegas, NV, USA, 27-30 June 2016.

[8] D. Tran, L. Bourdev, and R. Fergus, "Learning spatiotemporal features with 3d convolutional networks," in *Proceedings of the IEEE international conference on computer vision*, pp. 4489–4497, Cambridge, MA, USA, 20-23 June 1995.

[9] M. Baccouche, F. Mamalet, and C. Wolf, *Sequential Deep Learning for Human Action recognition[C]//International Workshop on Human Behavior Understanding*, pp. 29–39, Springer, Berlin, Heidelberg, 2011.

[10] E. P. Ijjina and K. M. Chalavadi, "Human action recognition in RGB-D videos using motion sequence information and deep learning," *Pattern Recognition*, vol. 72, pp. 504–516, 2017.

[11] Y. Wang and A. Dantcheva, "A video is worth more than 1000 lies. Comparing 3DCNN approaches for detecting deepfakes," in *Proceedings of the 2020 15th IEEE International Conference on Automatic Face and Gesture Recognition (FG 2020)*, pp. 515–519, IEEE, Buenos Aires, Argentina, 16-20 November 2020.

[12] K. Hara, H. Kataoka, and Y. Satoh, "Can spatiotemporal 3d cnns retrace the history of 2d cnns and imagenet?" in *Proceedings of the IEEE conference on Computer Vision and Pattern Recognition*, pp. 6546–6555, San Juan, PR, USA, 17-19 June 1997.

[13] A. Rossler, D. Cozzolino, and L. Verdoliva, "Learning to detect manipulated facial images," 2014, http://arxiv.org/abs/1901/08971.

[14] J. Y. Baek, Y. S. Yoo, and S. H. Bae, "Generative adversarial ensemble learning for face forensics," *IEEE Access*, vol. 8, pp. 45421–45431, 2020.

[15] T. Wang, J. Li, M. Zhang, A. Zhu, H. Snoussi, and C. Choi, "An enhanced 3DCNN-ConvLSTM for spatiotemporal multimedia data analysis," *Concurrency and Computation: Practice and Experience*, vol. 33, no. 2, p. e5302, 2021.

[16] B. Mishra, D. Garg, and P. Narang, "A hybrid approach for search and rescue using 3DCNN and PSO," *Neural Computing & Applications*, vol. 33, no. 17, pp. 10813–10827, 2021.

[17] M. Al-Hammadi, G. Muhammad, W. Abdul, M. Alsulaiman, M. A. Bencherif, and M. A. Mekhtiche, "Hand gesture recognition for sign language using 3DCNN," *IEEE Access*, vol. 8, pp. 79491–79509, 2020.

[18] W. Lin, J. Gao, Q. Wang, and X. Li, "Learning to detect anomaly events in crowd scenes from synthetic data," *Neurocomputing*, vol. 436, pp. 248–259, 2021.

[19] M. A. R. Ahad, M. Ahmed, A. Das Antar, Y. Makihara, and Y. Yagi, "Action recognition using kinematics posture feature on 3D skeleton joint locations," *Pattern Recognition Letters*, vol. 145, pp. 216–224, 2021.

[20] H. Hou, Y. Li, and C. Zhang, "Vehicle behavior recognition using multi-stream 3D convolutional neural network[C]// 2021 36th youth academic annual conference of Chinese association of automation (YAC)," *IEEE*, pp. 355–360, 2021.

[21] H. Wei, R. Jafari, and N. Kehtarnavaz, "Fusion of video and inertial sensing for deep learning–based human action recognition," *Sensors*, vol. 19, no. 17, p. 3680, 2019.

[22] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," in *Proceedings of the International conference on learning representations*, IEEE, April 25 to 29 2022.

[23] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the 2016 IEEE conference on computer vision and pattern recognition (CVPR)*, pp. 770–778, 27-30 June 2016.

[24] I. Amerini, L. Galteri, R. Caldelli, and A. Del Bimbo, "Deepfake video detection through optical flow based cnn," in *Proceedings of the 2019 IEEE/CVF International conference on computer vision workshop (ICCVW)*, pp. 1205–1207, 27-28 Oct. 2019.

[25] A. C. H. Choong and N. K. Lee, "Evaluation of convolutionary neural networks modeling of DNA sequences using ordinal versus one-hot encoding method," in *Proceedings of the 2017 International Conference on Computer and Drone Applications (IConDA)*, pp. 60–65, IEEE, Kuching, Malaysia, 09-11 November 2017.

[26] D. P. Kingma and J. Ba, "Adam: a method for stochastic optimization," 2014, http://arxiv.org/abs/1412.6980.

Hindawi
Computational Intelligence and Neuroscience
Volume 2022, Article ID 9521107, 11 pages
https://doi.org/10.1155/2022/9521107

Research Article

A Method for Evaluating the Green Economic Efficiency of Resource-Based Cities Based on Neural Network Improved DEA Model

Zhifeng Shen,[1] **Ning Liu**,[1] **Xialing Li**,[1] **and Zhengguang Kang**[2]

[1]*Jiangsu University of Technology, Changzhou, Jiangsu 213001, China*
[2]*Jiangsu Academy of Science and Technology for Development, Nanjing, Jiangsu 210042, China*

Correspondence should be addressed to Zhifeng Shen; shenzhifeng@jsut.edu.cn

Received 1 July 2022; Revised 22 August 2022; Accepted 24 August 2022; Published 8 September 2022

Academic Editor: Ning Cao

In this study, we use BP neural network to improve the DEA model to conduct in-depth research and analysis on the method of green economic efficiency evaluation of resource-based cities. The traditional DEA cannot make ranking and analysis of effective units, which affects the accuracy of empirical analysis. Accordingly, the BP-DEA model is introduced to further conduct a comparative eco-efficiency analysis of relatively effective provinces. In this study, the optimal inputs and outputs are calculated by DEA, and further, the BP neural network is used to fit the functional relationship between the optimal inputs and outputs, and by adding variables, the trained neural network can be used for the prediction of the optimal outputs. In this study, the BP-DEA model is used to empirically investigate the temporal evolution trend, spatial differences, and efficiency differences in eco-efficiency. Meanwhile, breaking through the limitation that DEA can only calculate regional efficiency values, this study combines the Malmquist index to compare and decompose the eco-efficiency of different provinces to analyze the sources of total factor productivity changes. The results show that the method can clarify the gap between the actual operation of each indicator and the reference point; it can identify how much room for improvement still needs to be made for each indicator, and it can also determine whether each city should be rewarded or penalized and its specific amount. Finally, based on the evaluation of eco-efficiency and the main constraints, corresponding policy recommendations are proposed. Finally, based on the evaluation results of the BP-DEA method, this study analyzes the overall efficiency improvement of cities in the two study areas in three dimensions: urbanization construction, ecology, and economic development put forward seven types of urban efficiency improvement and propose targeted urban development suggestions according to regional characteristics.

1. Introduction

In the era of big data, how to use big data for effective analysis has become the focus of attention in various industries. Due to the diversity of sources, quantities, structural forms, and other characteristics of big data, it covers many fields, but has a low value density. The existence of data noise and data redundancy in big data sets can have an incalculable negative impact on data analysis. In addition, the functional relationships between multiple variables are also covered in big data sets, which may produce a certain bias in the data analysis results [1]. Therefore, before using big data for analysis and research, data preprocessing should be performed first to eliminate redundant and invalid data. However, traditional big data preprocessing methods do not consider the functional relationships between variables. Data noise and data redundancy in large data sets can have immeasurable negative effects on data analysis. In addition, the large data set also covers the functional relationship between multiple variables, which may cause some deviations in the data analysis results. Therefore, before using big data for analysis and research, it is necessary to preprocess big data to eliminate redundant and invalid data. Data envelopment analysis (DEA) can effectively deal with the bias problem caused by the functional relationship between variables [2]. In the process of data preprocessing by DEA,

the most effective data are obtained by filtering the efficiency values, eliminating the outliers and redundant values, and reducing the quantity of data without changing the quality of data. There is no need to predict the functional relationship between input and output variables, and there is no need to set weights in advance, and the most effective data can be obtained by filtering the obtained efficiency values. There is no need to remove outliers and redundant values, and reducing the amount of data without changing the quality of the data is an effective way of data preprocessing that can be applied to machine learning. With urbanization, cities are facing many challenges such as ecological land reduction and environmental pollution, and the contradiction between economic development and the ecological environment is becoming increasingly prominent [3]. Eco-efficiency proposes to maximize value while minimizing resource consumption and environmental pollution, which is an inherent requirement for sustainable urban development. Since the twenty-first century, with the increasing intensity of human activities, the water crisis and ecological deficit have become increasingly serious. At present, the world is facing the common challenge of the water crisis, which not only makes the prospect of economic development a moot point but also may threaten the stability of society and the survival of human beings [4]. In sustainable development, resources and the environment are not only endogenous variables for economic development but also rigid constraints on the scale and speed of economic development. For a long time in the past, governments around the world have focused more on economic growth, placing too much emphasis on GDP growth, and often neglecting issues such as resources and the environment. The problems of resource depletion, environmental degradation, and unsustainability brought about by economic growth have intensified. Clarifying the current situation, spatial differences, and constraints of eco-efficiency have become the primary issue facing sustainable development. Among the existing urban efficiency evaluation methods, DEA has been widely used due to its unique advantages. DEA is a nonparametric technical efficiency analysis method, and it has been more than 40 years since the DEA method was proposed. However, due to the increasing number of elements absorbed by cities in the process of development, they have increasingly evolved into a complex and multilevel system, and with the development of regional integration, the city network system has gradually formed and strengthened, and the competitive effect and complementary effect between cities have emerged. Therefore, it is necessary to improve the relevant urban efficiency evaluation methods. Economic growth has brought about problems such as resource depletion, environmental degradation, and unsustainability, which have intensified. The world is facing the common challenge of water crisis, which will not only make the prospects of economic development impossible but also threaten the stability of society and the survival of human beings.

The way of urban development must shift from focusing on speed and scale in the past to efficiency and quality, i.e., to achieve high-quality urban development [5]. Urban efficiency is a comprehensive reflection of high-quality urban development. At present, reconstructing the urban input-output index system and evaluating urban efficiency are of great significance to promote high-quality urban development with efficiency change. Evaluation of urban efficiency is a hot issue in urban development research, and urban efficiency evaluation has achieved fruitful research results in terms of evaluation methods and evaluation perspectives. DEA is a nonparametric analysis method used to evaluate the relative efficiency of a group of decision-making units (DMUs) with multiple inputs and multiple outputs [6]. With the continuous research on the DEA model, many scholars have made various improvements based on the traditional DEA model. Although the defects such as incomplete ranking and distortion of weights in the traditional model have been compensated, no spatial improvement has been made to the traditional DEA model for cities, which are regional-type evaluation units with spatial attributes. The study of how to improve the DEA model spatially will improve the targeting of urban efficiency evaluation and will be more conducive to improving the rationality of urban resource allocation. Based on the above problems, this study introduces the concept of spatial interaction among cities based on the traditional DEA model, uses the geographic detector tool, combines the modified gravity model, reconstructs the evaluation set in the traditional DEA model by the degree of spatial connection among cities, and realizes the spatial improvement of the DEA model; with the help of BP (back propagation) neural network model, a BP-DEA efficiency evaluation method is proposed.

2. Related Works

With the rapid development of information technology, the speed of data analysis and processing has greatly increased, and creating real-time data streams has become a popular trend. To improve their competitiveness, companies need to know not only how to create data quickly but also how to process, analyze, and return it to users quickly to meet their real-time needs [7]. In terms of diversity, the reason for the diversity of data formation is the wide variety of data sources, including various search engines, browsing traces, and social networks; in addition to this, the data are classified in a wide variety of formats. In terms of value, we can conclude that although the value of big data is high, its value density is low; often, only a few hundred or even a few dozen data of thousands of data have value [8]. Big data are not only a revolution of data but also a revolution of thinking, and it has been widely used in government, health care, education, finance, food safety, and other fields, changing our way of life and even our way of thinking. Suocheng et al. [9] used DEA as a tool to preprocess data sets. DEA can identify data that do not match a certain attribute and remove it to get more accurate predictions [9]. Cook et al. [10] demonstrated the high growth rate of DEA-based research. Preprocessing the data based on the DEA method leads to fewer records and less computation, reducing the burden of the study; at the same time, outliers are eliminated to make the subsequent training data more general [10]. The Cobb–Douglas production function model and the Solow

economic growth equation are used to calculate the contribution rate of capital, labor, and technological progress to economic growth, and then, the DEA-Malmquist index method is used to calculate the total factor productivity of each province [11]. The goal of removing outliers, preserving universality, and reducing the size of the database used for ANN (artificial neural network) training is achieved. This new data analysis method can improve the prediction accuracy while making the training data set much smaller.

Environment and resources are the general terms for natural factors such as natural existence and a manufactured creation, which lay down the survival, development, and progress of human society; they are characterized by wholeness, scarcity, and value and are the unity of static and dynamic. Resource and environment index system often consists of resource consumption index system and environmental pollution index system [12]. As people pay increased attention to the resources and environment, they gradually realize that the role of resources and environment on economic growth is also becoming increasingly obvious. Peykani et al. [13] analyzed input-output models for 10 sectors to reveal the complex interrelationship between energy, environment, and economic welfare, proposed pollution emission factors and European sulfur deposition carriers, and studied the relationship between energy use, and environmental impact, employment, and economic development [13] Kononets et al. [14] analyzed and evaluated the environmental carrying capacity of Anhui Province by establishing the entropy TOPSIS model and came up with four key factors affecting the environmental carrying capacity of Anhui Province [14]. Tan et al. [15] argued that the current research on the bearing capacity of resources and environment has not yet formed a unified systematic theory, and should strengthen the basic theory and bearing mechanism, technical standards and norms, evaluation, and system integration research so that the research on the bearing capacity of resources and environment can be more standardized and systematic for the better benefit of human beings [15].

In terms of research on the application area of eco-efficiency, foreign countries have more research on enterprises and product systems. The research organically combines factors such as product development and design, system structure, and identification of key issues with eco-efficiency research and applies the results to the process of healthy and sustainable development of enterprises. China is more inclined toward regional studies, such as the Western provinces, six central provinces, Beijing-Tianjin-Hebei city cluster, and the middle reaches of the Yangtze River city cluster. By comparing the differences in eco-efficiency among cities, targeted improvement strategies are proposed in terms of policy, management, and technology. Secondly, in terms of evaluation methods, the single ratio analysis method, indicator system analysis method, and model analysis method described above have their advantages and disadvantages, and appropriate evaluation methods need to be selected according to the purpose of the study and the characteristics of the research object. With the advancement of urbanization, cities are faced with many challenges such as the reduction in ecological land and environmental pollution, and the contradiction between economic development and ecological environment has become increasingly prominent. Overseas research on the theory and application of model analysis and single ratio analysis has made great progress. In terms of the use of the DEA model, foreign studies have improved the traditional DEA model to different degrees and aspects based on the characteristics of the evaluation unit [16]. Related applied studies have shown an increasing trend in recent years. However, since most of the studies are based on the application of existing models, scholars have not further explored and compared the principles, applicability, and accuracy of the models, resulting in certain discrepancies in the results of eco-efficiency assessment, which can directly influence the government's decision-making and thus may lead to further deterioration of resources and environment. Based on this, this study discusses the characteristics of the base model of DEA and improves the DEA model with the help of the BP neural network model to meet the research of eco-efficiency in this study.

3. Methodology

3.1. Data Envelope Analysis. Charnes first proposed data envelopment analysis (DEA) in 1978, in which a set of decision-making units (DMUs) of the same type is selected, and the relative efficiency of the DMUs is calculated by analyzing data on input and output indicators through a mathematical planning model and then determining the relative effectiveness of the DMUs and whether scale efficiency and technical efficiency are achieved. DEA is a nonparametric statistical method that can obtain DEA models from input and output data and perform economic analysis. Therefore, DEA is highly valued by the international academic community and has become the most common method in the field of efficiency evaluation. For a given complex large data set, even if there are many kinds of variables in the data set and the relationship between variables is complicated, DEA preprocessing can perfectly avoid the negative factors with its own unique mathematical planning advantages and complete the efficiency evaluation of variable indexes. When evaluating the efficiency of DMU based on the DEA data preprocessing method, it is not necessary to know the attribute relationships embedded inside the complex data set in advance, and it is not necessary to determine the variable weights to obtain the effective DMU and provide the improvement direction and adjustment values for the invalid DMU, which has certain advantages in this study.

BP neural network is a multilayer feedback network trained by the error back propagation algorithm proposed by a group of scientists led by Rumelhart and McCelland in 1986 as an artificial intelligence information processing system capable of learning nonlinear functional relationships between variables and underlying patterns. This neural network model is characterized by the fact that it does not need to assume the mapping relationship between the input and output in advance, but uses the learning

adaptation ability among the network nodes to map the fitting result with the minimum distance from the desired output value through continuous training [17]. Therefore, it is more common in the research of artificial neural network model applications. The whole computational process of this BP algorithm includes the positive propagation of the signal and the inverse propagation of the error. That is, the error output is from the output point to the input point, while the adjustment weights and thresholds are from the input point to the output point. In the positive propagation process, the input signal is transmitted to the implicit layer through each node of the input layer and then transmitted to the output layer through the nonlinear processing of the implicit layer to complete the positive transmission of the signal. If the actual output does not match the expected output, the back propagation of the output error occurs. During the back propagation, the output error is transferred to the upper layer, and the weights and thresholds of that layer are corrected in the direction of gradient descent [18]. The network training keeps reducing the error along the gradient direction until the termination condition is met, and the termination condition is usually set to the minimum error or the maximum number of training sessions to obtain the weights and thresholds under that training expectation. In this study, we take a common single hidden layer BP neural network as an example to illustrate the self-learning process of BP networks, i.e., the BP algorithm. By correcting the error and adjusting the connection weights and thresholds, the training is repeated until the error is less than the set minimum error, and then, the learning process is finished, and the trained network can map the functional relationship between the variables and the underlying patterns more accurately, to realize the prediction function of the network.

$$\begin{aligned} Q_j &= f\left(\sum_{i,j=1} W_{ij}\right) \times f\left(x_i - q_j\right), \\ Y_i &= f\left(\sum_{j,k=1} T_{jk}\right) - f\left(\mathrm{O}_j \times \mathrm{q}_k\right). \end{aligned} \tag{1}$$

Among them, f is the action function; q is the neuron threshold.

DEA is an efficiency measurement method based on linear programming ideas with a strict mathematical derivation process, so the model can provide target values and improvement amounts for invalid decision units, but this also leads to certain limitations in the application of the method; i.e., it is more sensitive to evaluation index data and only suitable for evaluating decision units with objective values (nonpredicted values) and lacks certain predictive ability. In contrast, BP neural network as an artificial intelligence information processing system can learn the nonlinear functional relationships between variables and the underlying patterns, and its advantage lies in the predictive capability. The general idea of modeling in this study is to use the input-output index values from 2015 to 2021 as the BP neural network input and the comprehensive efficiency values measured by the CCR model as the BP neural network output, in which a total of 70 samples (70%) of the Yangtze River Economic Zone data from 2015 to 2018 are selected as the network training samples, and a total of 30 samples (30%) from 2019 to 2021 are selected as the test samples of the network. Through the learning and training of the BP neural network, the underlying relationship between the input-output index and the comprehensive efficiency is fitted, and when the trained efficiency prediction model passes the test, it shows that the model has a strong generalization ability.

According to the definition of the efficiency evaluation index, the efficiency evaluation of a province can be seen as the integration of all the input and output indicators into a relatively representative index, and then, the ratio of a single input to a single output is calculated. Assuming that the calculated province is denoted as DMU_{jo} ($jo = 1, 2, \ldots, \mathrm{m}$), the optimal weight vectors q and p are chosen to satisfy the maximum value of the evaluation index k_{jo} for the province, if none of the provincial efficiency indices is greater than 1. The evaluation index k_{jo} for a province refers to the eco-efficiency value of the j_0 evaluated provincial area, which is the following $0 < k_{jo} < 1$. Its economic significance is the extent to which the input index of the evaluated j_0 can be scaled down if the economic output of the evaluated can be replaced by any other linear combination of regional outputs, and the magnitude of the scaling down ratio is k_{jo}. If the eco-efficiency of the region is valid, $k_{j_0} = 1$; when $k_{jo} < 1$, $1 - k_{jo}$ represents the maximum proportion of input reduction. DEA linear programming model is mainly used to discriminate whether the output of each province is valid or not, and the model to discriminate the validity of the input of the study object ($L^{\cdot}$) is obtained by doing the Charnes–Cooper transformation on the inverse form of the objective function of the model ($\bar{L}$).

$$\begin{aligned} k_{jo} &= \min \frac{p^t y_{jo}}{q^T x_{jo}}, \\ s.t &\begin{cases} \varpi^T + q^T y_{jo} \le 0 \\ \varpi \ge 0 \\ (j = 1, 2, \ldots, n) \end{cases} . \end{aligned} \tag{2}$$

3.2. BP-DEA Model Structure Design. The DEA method has its advantages in dealing with multiple input and multiple output problems, which largely avoids the influence of human subjective factors, and its evaluation results are more scientific and accurate. However, the DEA method also has its defects, which cannot be further simulated and predicted and cannot provide clear solution ideas for decision-makers. Since BP neural network requires training samples and corresponding tutor values, this process still cannot avoid the need to use other methods to assist in the evaluation, and

the more common is the use of the expert scoring AHP method. This inevitably introduces the human subjective factor, and the evaluation period will be prolonged by the time of collecting expert experience scores [19]. Because of the respective advantages and shortcomings of the above two methods, this study organically combines the two methods for science and technology achievement transformation evaluation. Firstly, the original evaluation index data are input into the DEA model to calculate the comprehensive evaluation value of the transformation of scientific and technological achievements based on the DEA model, and then, the BP neural network is trained with this evaluation value as the tutor value. The BP-DEA neural network comprehensive evaluation model constructed in this study is shown in Figure 1. On the other hand, its relative efficiency evaluation value can be used as a piece of objective tutor information to replace the subjective expert rating used in the traditional BP network, to guide the BP. On the other hand, its relative efficiency evaluation value can be used as objective tutor information to replace the subjective expert rating used in traditional BP networks, thus guiding the BP neural network to self-learning and training, and forming a stable network structure for specific evaluation and prediction.

To make the information provided by the DEA method more accurate and scientific, it is necessary to first clarify the objective of the evaluation, analyze the decision unit around this objective, including the main objective, subobjectives, and factors affecting these objectives, and establish a hierarchical structure. Then, the nature of various factors is clarified, distinguishing whether the factors are variable, controllable, and primary and secondary relationships. Then, the quantitative and qualitative relationships among the factors are considered. Sometimes, it is also necessary to distinguish whether the decision unit has boundaries or not and to clarify the structure and hierarchy of the decision unit [20]. When the evaluation problem is determined, it is necessary to determine the indicators that can reflect the evaluation objectives and reflect the qualitative relationships among the indicators into the weight constraints, to construct the evaluation indicator system. Then, the decision unit is selected; i.e., the reference set is determined, and the selected reference set must have the same objectives, tasks, external environment, and input-output indicators, and the decision unit should be representative to a certain extent. Then, the index data are collected and organized, and finally, the appropriate DEA model is selected for calculation and analysis according to the purpose of validity analysis and the actual problem context. According to the calculation results obtained in Stage 2, we analyze and compare them, find out the reasons for invalid units, and give improvement suggestions.

The BCC model is obtained by widening the premise assumption of constant payoffs of scale in the CCR model by adding a condition $\sum_{j=0}\lambda_j = 1$ to obtain the BCC model under variable payoffs of scale. Since the model considers different forms of payoffs to scale, the efficiency measured by the BCC model is purely technical. In case of negative values of input-output indicators, the BCC model can be used to process a linear transformation of the data. In addition to the three formulas that can be consulted to roughly determine the number of neurons in the hidden layer, there is another method that can be used to determine the number of neurons in the hidden layer. The first step is to make the number of implicit neurons variable or to put in enough implicit neurons to eliminate those that do not work by learning until they are not contractible. Similarly, it is possible to start with a relatively small number of neurons, learn a certain number of times, and then increase the number of hidden layer neurons if unsuccessful, until a reasonable number of hidden units is reached. These two methods are combined to determine the number of hidden units [21]. First, the reference formula shows that the approximate number of hidden units is around 5. For this reason, it is possible to set the loop and check when the network is more efficient and the output error is smaller when the number of hidden units is taken between 4 and 6. The errors and effects during training are compared to determine a reasonable number of hidden units. As shown in Figure 2, when the hidden layer neurons are in 4–6, both train twice to reach the target error, but the error message is different for different hidden layer neurons. When the hidden layer neurons are 4, the validation set error is 0.0034; when the hidden layer neurons are 5, the validation set error is 0.0087; and when the hidden layer neurons are 6, the validation set error is 0.0064, so the number of hidden layer neurons is 4.

4. Design of Evaluation Indicators for Green Economic Efficiency of Resource City-Type Cities

System theory is to treat the specific object under study as an overall system and sorts out the interactions among the constituent elements of the system with the basic idea of starting from a holistic perspective and grasping its internal structure and dynamic changes to achieve the goal of optimizing the whole. The basic attributes of the system include wholeness, hierarchy, correlation, dynamics, and purpose. In constructing the theoretical analysis framework of the factors influencing the efficiency of construction land use, not all factors act independently on the efficiency of construction land use, but directly or indirectly influence the efficiency of construction land use through interconnected pathways; i.e., most of the influencing factors have a strong correlation with each other, and all, in other words, most of the influencing factors have a strong correlation with each other, and all the influencing factors can be regarded as a whole system.

From 2003 to 2005, with the rapid development of industrialization and urbanization, energy consumption intensity and major pollutant emissions showed an upward trend. The energy consumption per unit of GDP rose by 9.8%, and the total emissions of sulfur dioxide and chemical demand increased, respectively, by 32.3% and 3.5%. At the

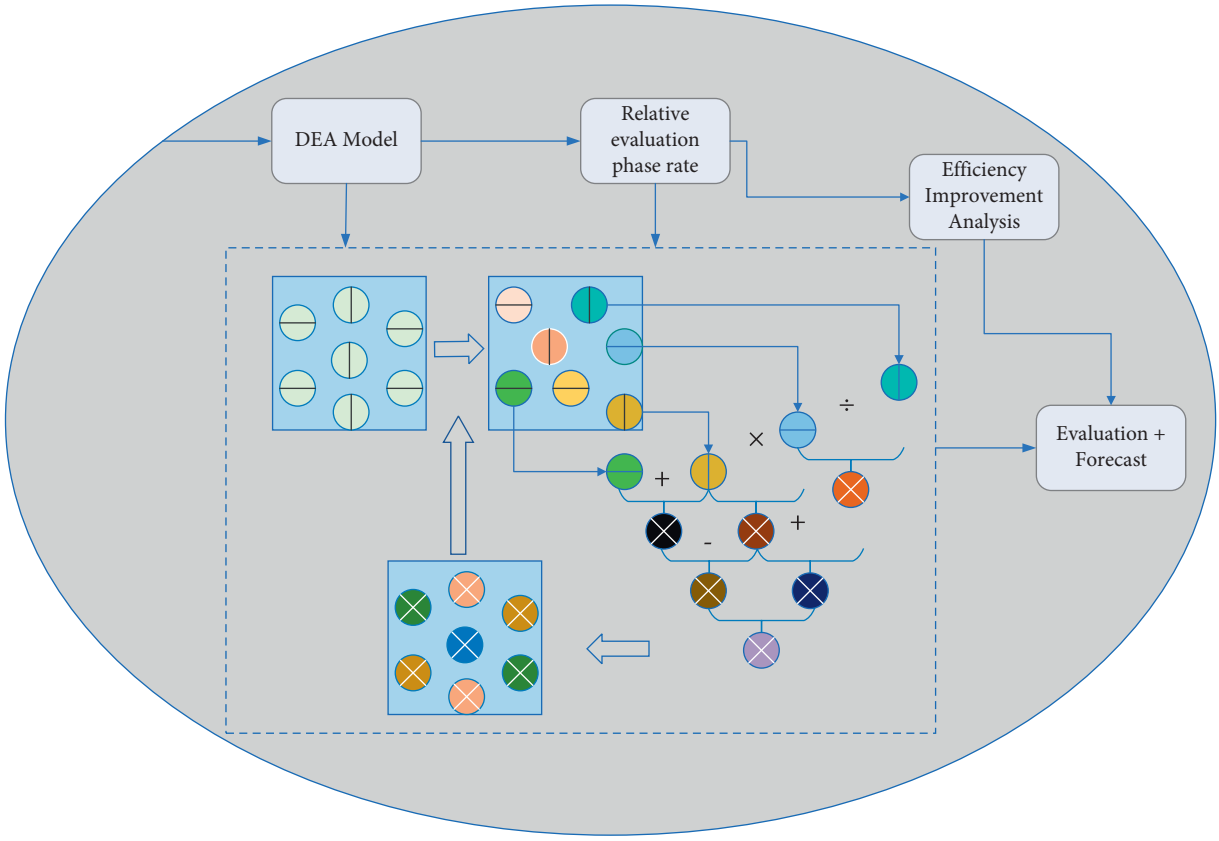


Figure 1: BP-DEA network comprehensive evaluation model.

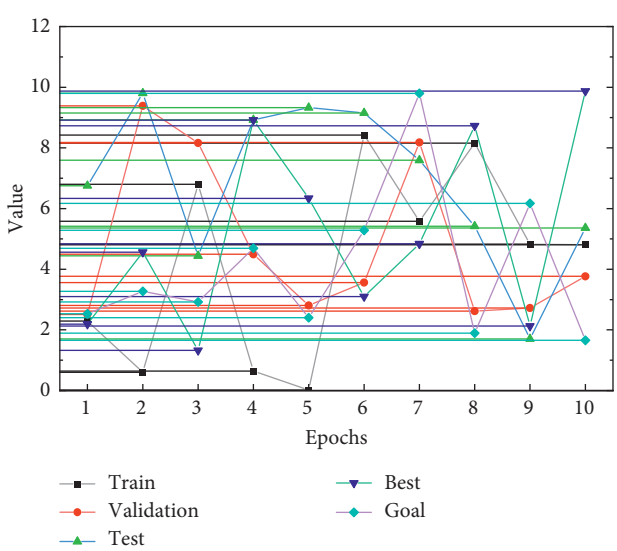


Figure 2: Relationship between the number of neurons in the hidden layer and the number of trainings.

same time, these few years are an important time period for the accumulation of capital investment. During this period, a large amount of capital flows to the secondary industry, which leads to the ineffectiveness of the scale and management of this industry to a certain extent. At the same time, because of not giving up the use of the original backward mechanical equipment, there is even a regression in technology. In terms of environmental governance and protection, the discharge of many environmental pollutants has not reached the predetermined target. During this period, China was highly dependent on coal energy, but due to backward equipment, the utilization efficiency of coal resources was relatively low, and many equipment using coal resources did not have "desulfurization" devices, which made the pollutant emissions in some areas very high serious.

From 2006 to 2010, the significant improvement of regional total factor energy efficiency was mainly due to the reversal of the rising trend of energy consumption intensity and major pollutant emissions during the period of the government's "Eleventh Five-Year Plan" during the rapid development stage of industrialization and urbanization. During the "Eleventh Five-Year Plan" period, the energy consumption per unit of GDP has changed from an increase of 9.8% in the three years after the "Tenth Five-Year Plan" to a decrease of 19.1%. The increase of 32.3% and 3.5% turned into a decrease of 14.29% and 12.45%, and the energy efficiency level and environmental quality were significantly improved. At the same time, the adjustment of industrial structure and the further improvement of infrastructure have also promoted the improvement of total factor energy efficiency to a certain extent.

The law of diminishing marginal efficiency, also known as the law of diminishing marginal utility, is one of the basic

principles in the field of economics [22]. Increasing the input of this resource may lead to a decrease in the output volume, i.e., the output volume of the increased unit resource is decreasing, as shown in Figure 3. When the input of construction land is too small compared with the input of other factors, increasing the input of construction land will not only lead to an increase in total output but also lead to an increase in the marginal output of a new unit of construction land, but once the appropriate ratio between factors is crossed, increasing the input of construction land makes the increase in total output. However, once the ratio between factors is exceeded, the increase in total output is not obvious and the marginal output per unit of new construction land decreases gradually.

To address the shortcomings of this study, this study incorporates natural ecosystems into the evaluation of urban eco-efficiency and selects ecosystem service values and landscape diversity indicators as the characterization indicators of natural ecosystems, taking 34 prefecture-level cities in *X* Province as examples for the specific measurement of ecosystem indicators. Natural ecosystems provide the material basis for human survival and development. Humans can obtain the food, energy, and water they need from natural ecosystems, and they can also obtain various benefits such as disaster prevention and carbon sequestration, and rest, all of which reflect the welfare that nature gives to human survival development. Therefore, the value of final products and services provided by natural ecosystems to human beings should also be used as an important indicator to measure the sustainable development of a country. Based on the above analysis, this study incorporates indicators characterizing natural ecosystems into urban eco-efficiency assessment and selects the value of ecosystem services and landscape diversity indicators as indicators characterizing natural ecosystems [23]. The value of ecosystem services and the landscape diversity index characterize the value and structure of natural ecosystems, and it is difficult for social statistics to show their changes. Due to the specialized nature of data acquisition and data processing, these indicators are rarely used in the social science field, and it is difficult to obtain data on ecosystem structure indicators. Natural ecosystem data usually require remote sensing and GIS spatial technology to identify the changes in ecosystem structure and reflect the differences between regions. In this study, we use ESA global land cover data for 34 cities in *X* Province, select 2012, 2015, 2018, and 2021 as the study years with a 3-year interval, and use GIS technology to spatially process the data to measure the natural ecosystem index data of 34 cities in the region during 2012–2021, to further use the model to measure regional ecosystem indicators. The data were spatially processed using GIS technology to measure the natural ecosystem indicators of 34 cities in the region during 2012–2021.

In this study, we firstly processed the acquired spatial data, and the main processing method was ArcGIS spatial technology method. First, based on the national resource and environmental database classification standard and land cover remote sensing classification system, combined

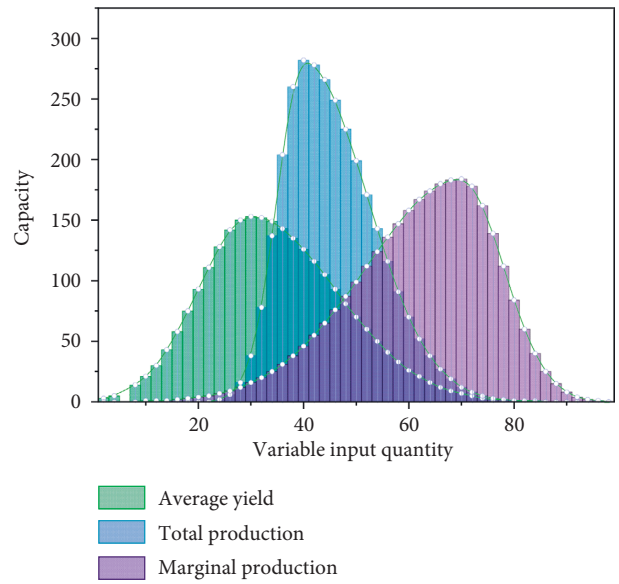


Figure 3: Variable factor inputs versus yield.

with the ecological environment of *X* Province and other information, we identified and classified the ecosystem types in Northeast China. Since ecosystem services involve many aspects, and the connection between ecological processes and economic values is complex, there are still many uncertainties in the understanding of ecosystems, and there is a certain roughness in the value assessment of ecosystems. In the actual measurement, the formula for calculating the value of ecosystem services is as follows. Based on the diverse characteristics of big data sources, quantities, different structural forms, and real-time nature, the value it covers is very high, but its value density is very low. Data noise and data redundancy in large data sets can have immeasurable negative effects on data analysis.

$$\mathrm{ESV} = \sum_{k=1}^{n} A_k \times VC_{k.} \tag{3}$$

The total value of ecosystem services ($) in the study area is represented by k, the number of ecosystem types, VC_k is the coefficient of total ecological service value per unit area of ecosystem type k ($/hm^2), and A_k is the area of ecosystem type k in the study area (hm^2).

In recent years, the application of diverse tools such as remote sensing and geographic information system (GIS) in landscape ecology has enabled the quantitative measurement of landscape diversity. The SHDI can characterize the changes in landscape structure and function over time and reflect the complexity of the landscape. The SHDI usually takes the value range of SHDI ≥ 0, and the larger the SHDI is, the more complex the landscape is. The SHDI index can not only calculate and analyze the changes in diversity and heterogeneity of different landscapes in the same period but also compare the changes in the same landscape in different periods.

5. Analysis of Results

5.1. BP-DEA Model Results. Put in a relatively small number of neurons at the beginning, after learning a certain number of times, if unsuccessful, the number of neurons in the hidden layer is increased until a reasonable number of hidden units are reached. In the evaluation results of traditional standard DEA, there are often cases where the efficiency of multiple DMUs is all 1, and both more DMUs are relatively efficient, and at this time, the standard model cannot distinguish these DMUs that are also efficient. In the radial DEA model, the measurement of the degree of inefficiency only includes the proportion of all inputs and outputs that are reduced or increased in equal proportion. For the inefficient DMU, the gap between its current state and the strongly efficient target value includes a slack improvement component in addition to the equiproportional improvement component, which is not captured in the efficiency value measurement.

To avoid singular data, which in turn affects the analysis results, this study needs to adopt normalization operations for the sample data. When each index component of the input/output vectors of the model does not have the same dimension, each index component needs to be normalized in the respective value domain [24]. When each index component of the input/output vectors has the same magnitude and the same size, it needs to be normalized in the whole sample data value domain. Considering the output range of the BP neural network transfer function and the type of each evaluation index, the maximum-minimum normalization method is chosen to process the sample data in this study.

$$X^* = \frac{x - \min A}{\max A + \min A}, \tag{4}$$

where x is a raw value of attribute A, $\min A$ and $\max A$ are the minimum and maximum values of attribute A, respectively, and X^* is the standard value after the maximum-minimum normalization method, and the range is $0 \le X^* \le 1$.

BP Neural Network Training: where the number of training steps is set to 1000, the initial learning rate is 0.05, the error value at the end of training is 0.001, and the learning rate is 0.05. When the training is completed, the training results need to be checked, and the error between the actual output value of the training set and the desired output value is analyzed to see whether the error between the two meets the accuracy requirements. If the accuracy requirements are met, it means that the training of the neural network has been completed and the test set can be tested, and if the accuracy requirements are not met, it is responsible for the need to adjust the parameters and retrain. The maximum relative error in the training sample is 3.43%, which is less than 5% of the accuracy requirement, indicating that the training is qualified and the BP neural network model has been trained. The results of comparing the real value of the training set with the predicted value are shown in Figure 4. The way of urban development must shift from focusing on speed and scale in the past to efficiency and quality, that is, to achieve high-quality urban development.

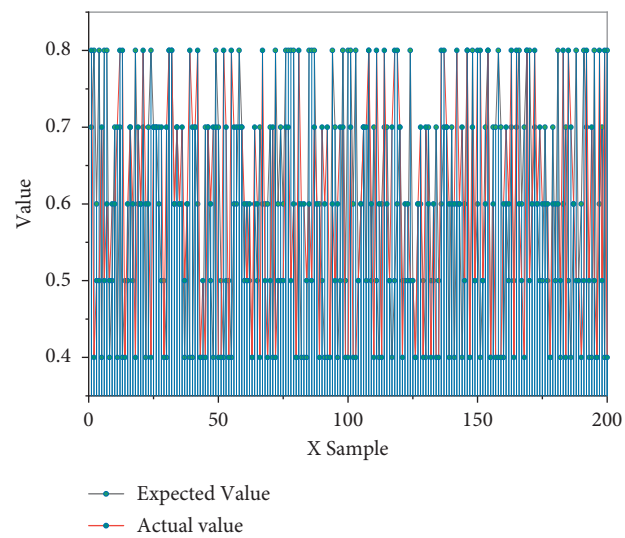


Figure 4: Comparison of the true and predicted values of the training set.

As can be seen from Figure 4, the expected output of the BP network model prediction sample is consistent with the actual output, indicating that the network model can better reflect the mapping relationship between the input and output values, and we can evaluate the efficiency of the terminal through this model. Using the AHP-based BP neural network evaluation model to evaluate the operation condition of X Province and analyze the current efficiency level of X Province is important for the enterprises in X Province to make and implement a series of targeted development strategies based on the actual operation condition, and then, the current operating situation of the enterprise is analyzed and judged whether the enterprise needs to reallocate the existing resources based on the results of the efficiency evaluation, to enhance the efficiency of the enterprise and improve the operational efficiency. In this stage, the most important step is to conduct an efficiency analysis, explore the factors affecting the efficiency of the enterprise, analyze the reasons affecting the efficiency of the enterprise in X Province, and then give corresponding improvement suggestions based on the specific reasons to improve the efficiency of the enterprise and increase the core competitiveness of the enterprise. If the input and output indicators have negative values, the BCC model can be used to linearly transform the data.

The learning rate is related to the speed of weight change at each step of training, and if the weights of the model change quickly, it will cause high-frequency oscillation of the loss function, and if the weights change too slowly, the harder the model will be to converge, and then, it will take longer to learn the training, but usually, the smaller the learning rate, the better the model learns, and the better it is for BP neural network training to approach the minimum. Therefore, a smaller value is preferred for tuning, so a learning rate in the range of [0.01, 0.08] is usually tried. By comparing the number of training steps and errors in the table, it can be found that the model performance is optimal

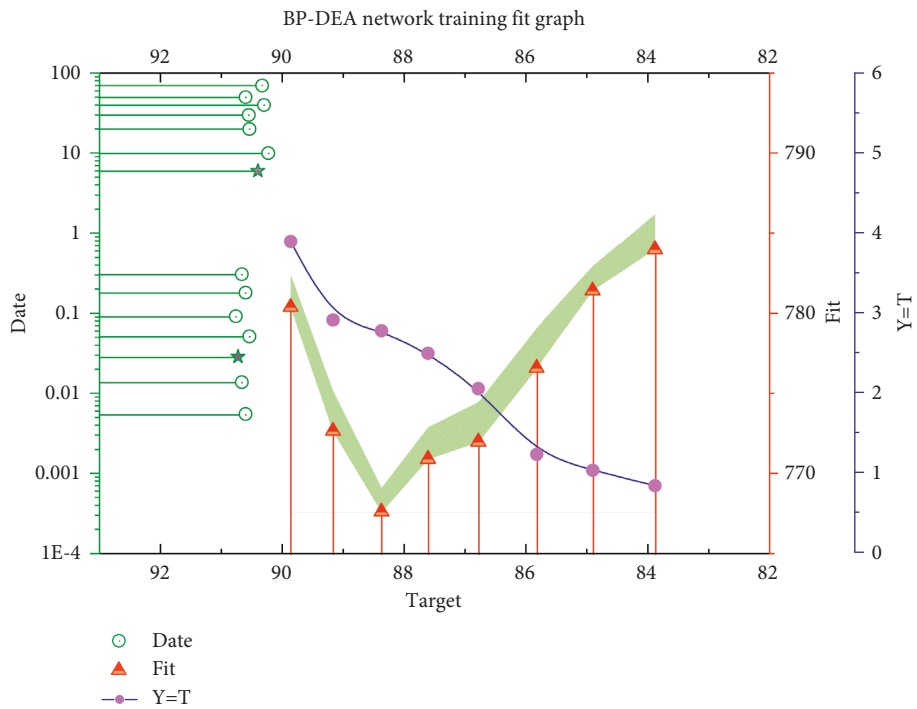


Figure 5: BP-DEA network training fit graph.

when taking 0.01, so the net learning rate of the efficiency prediction model is determined to be 0.01. The network training results are shown in Figure 5.

5.2. Simulation Experiment of Green Economic Efficiency Evaluation Model for Resource City-Type Cities. To explore the efficiency situation of the economic belt of Province *X* in 2022, and to better provide efficiency improvement suggestions for the relevant sector managers, a forecast-based scenario is set in this study. By forecasting the turnover of Province *X* in 2022, the efficiency prediction model constructed in the previous paper is used to measure the efficiency of Province X in 2022 under the projected road transport output when matching the current 2018 inputs, to provide a reference benchmark for the relevant managers to improve efficiency. As can be seen from Figure 6, the trend of turnover in Province X over time from 2010 to2018 is in a nonlinear form, and this study only forecasts for the next two periods, so it is suitable to use the three-time exponential smoothing method, and the Excel tool is selected to build the forecasting model. The termination condition of the BP algorithm is usually set as the minimum error or the maximum training times, to obtain the weights and thresholds under the training expectations.

After the parameters are estimated by the SFA regression model in the second stage, the provinces are adjusted to the same external environment and random error scenario so that each province faces the same business environment and conditions, thus further adjusting the resource environment input variables in the first stage to exclude the external environment factors, and finally substituting the adjusted input variable values with the output values in the first stage into the data envelope model again for the count. Finally, the adjusted input variable values and the output values of the first stage were again substituted into the data envelope model for calculation. Overall, the overall level of adjusted eco-efficiency in China is low, with large internal differences, and the trend of change in the east, middle, and west has also changed, showing approximately an "M"-shaped trend, with two transitions, in 2007 and 2013, respectively. As shown in Figure 7, after excluding external environmental variables and random disturbances, the regional eco-efficiency of each province has changed significantly. Compared with the first stage, the eco-efficiency has decreased significantly, by about 11.6%, which indicates that the external environmental factors and random luck disturbances that cannot be controlled can overestimate the eco-efficiency level of each province region to a certain extent.

In essence, eco-efficiency contains both resource-saving and environment-friendly contents. With the rapid development of the economy, a series of problems such as resource scarcity and environmental pollution have gradually become prominent, which restrict the sustainable development of the economy. Although in recent years, energy has become increasingly scarce and environmental pollution has become more serious, a series of corresponding measures taken by our government has made our overall eco-efficiency still in a slow upward trend, which of course is inseparable from our scientific and technological progress, the improvement of people's awareness of environmental protection, and the improvement of people's cultural literacy. A DMU of the same type refers to a set of DMUs with the same tasks and goals, the same environmental conditions, and the same input and output metrics.

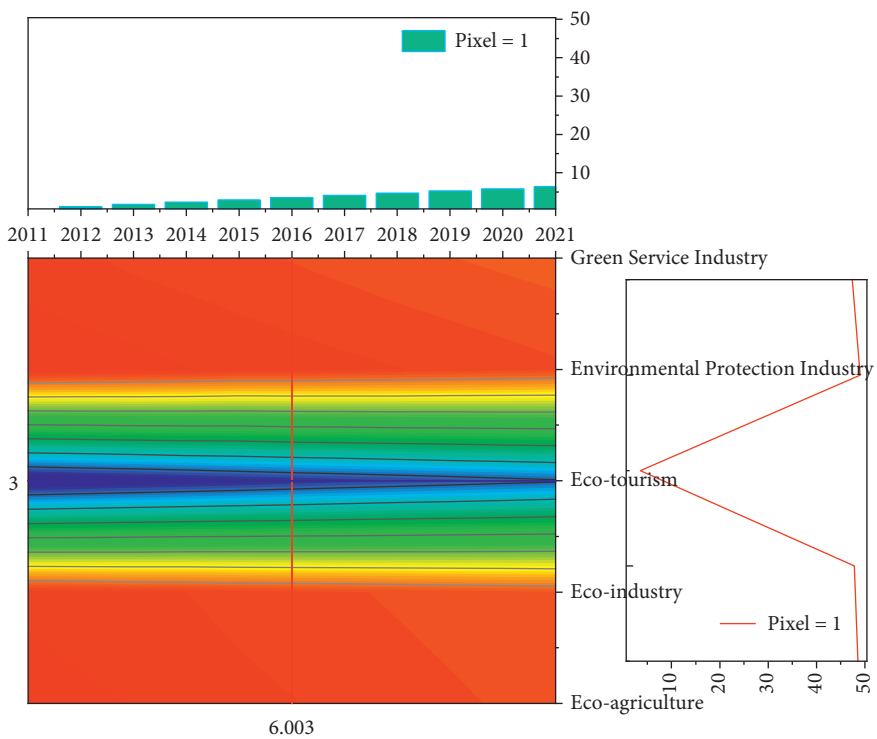


Figure 6: 2011–2021 X Province turnover trends.

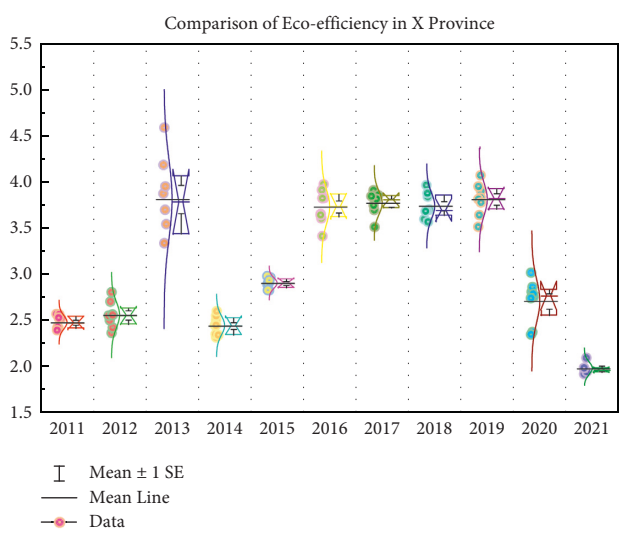


Figure 7: Comparison of eco-efficiency in X Province.

6. Conclusion

This study builds a regional ecological efficiency evaluation model based on BP neural network improved DEA model and evaluates the economic performance under the constraints of resources and environment. In this study, the use of DEA can effectively deal with the problem of deviation caused by the functional relationship between variables. The amount of data without changing the data quality is reduced, and the BP-DEA model is used to study and analyze the resource utilization and its influencing factors. The optimal input and output are calculated through DEA, and the BP neural network is further used to fit the functional relationship between the optimal input and output. By adding variables, the trained neural network can be used to predict the optimal output. Combined with the Malmquist index, the ecological efficiency of different provinces is compared and analyzed, and the gap between the actual operation of each indicator and the reference point is clarified. This research has achieved certain achievements in the evaluation method of green economic efficiency of resource-based cities. However, limited by personal knowledge ability and practical conditions, the study still has some deficiencies. The result measured by the BP-DEA method is relative efficiency, and there are still limitations in the analysis of time series for a certain decision-making unit. The parameter setting of BP neural network is relatively cumbersome, and there is no more direct selection method. Later, the heuristic algorithm can be introduced to simplify the setting process of relevant parameters.

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 or personal relationships that could have appeared to influence the work reported in this study.

Acknowledgments

The work of this study was supported by Philosophy and Social Science Research in Colleges and Universities in Jiangsu Province (2020SJA1180) and Project of Social Science Foundation of Jiangsu Province (21EYC014).

References

[1] Z. Jin, "Retracted article: green city economic efficiency based on cloud computing and machine learning," *Arabian Journal of Geosciences*, vol. 14, no. 11, pp. 1007–1016, 2021.

[2] K. Cui and X. Jing, "Research on prediction model of geotechnical parameters based on BP neural network," *Neural Computing & Applications*, vol. 31, no. 12, pp. 8205–8215, 2019.

[3] Y. G. Zhang, J. Tang, R. P. Liao, M. F.Y. Zhang, and X. M. Y. Wang, "Application of an enhanced BP neural network model with water cycle algorithm on landslide prediction," *Stochastic Environmental Research and Risk Assessment*, vol. 35, no. 6, pp. 1273–1291, 2021.

[4] Y. Deng, H. Xiao, J. Xu, and H. Wang, "Prediction model of PSO-BP neural network on coliform amount in special food," *Saudi Journal of Biological Sciences*, vol. 26, no. 6, pp. 1154–1160, 2019.

[5] 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.

[6] S. Ding, C. Su, and J. Yu, "An optimizing BP neural network algorithm based on genetic algorithm," *Artificial Intelligence Review*, vol. 36, no. 2, pp. 153–162, 2011.

[7] B. H. M. Sadeghi, "A BP-neural network predictor model for plastic injection molding process," *Journal of Materials Processing Technology*, vol. 103, no. 3, pp. 411–416, 2000.

[8] Q. Wei, J. Zhang, and X. Zhang, "An inverse DEA model for inputs/outputs estimate," *European Journal of Operational Research*, vol. 121, no. 1, pp. 151–163, 2000.

[9] D. Suocheng, L. Zehong, L. Bin, and X. Mei, "Problems and strategies of industrial transformation of China's resource-based cities," *China Population, Resources and Environment*, vol. 17, no. 5, pp. 12–17, 2007.

[10] W. D. Cook, L. Liang, Y. Zha, and J. Zhu, "A modified super-efficiency DEA model for infeasibility," *Journal of the Operational Research Society*, vol. 60, no. 2, pp. 276–281, 2009.

[11] S. S. Ganji and A. A. Rassafi, "DEA Malmquist productivity index based on a double-frontier slacks-based model: Iranian road safety assessment," *European Transport Research Review*, vol. 11, no. 1, p. 4, 2019.

[12] S. R. Seyedalizadeh Ganji and A. A. Rassafi, "Measuring the road safety performance of Iranian provinces: a double-frontier DEA model and evidential reasoning approach," *International Journal of Injury Control and Safety Promotion*, vol. 26, no. 2, pp. 156–169, 2019.

[13] P. Peykani, E. Mohammadi, and F. S. Seyed Esmaeili, "Stock evaluation under mixed uncertainties using robust DEA model," *Journal of Quality Engineering and Production Optimization*, vol. 4, no. 1, pp. 73–84, 2019.

[14] N. Kononets, O. Ilchenko, and V. Mokliak, "Future teachers resource-based learning system: experience of higher education institutions in Poltava city, Ukraine," *The Turkish Online Journal of Distance Education*, vol. 21, no. 3, pp. 199–220, 2020.

[15] M. Tan, H. Zhao, G. Li, and J. Qu, "Assessment of potentially toxic pollutants and urban livability in a typical resource-based city, China," *Environmental Science and Pollution Research*, vol. 27, no. 15, pp. 18640–18649, 2020.

[16] Y. Qin, Y. Luo, J. Lu, L. Yin, and X. Yu, "Simulation analysis of resource-based city development based on system dynamics: a case study of Panzhihua," *Applied Mathematics and Nonlinear Sciences*, vol. 3, no. 1, pp. 115–126, 2018.

[17] L. Wang, Y. Wang, Y. Sun, K. Han, and Y. Chen, "Financial inclusion and green economic efficiency: evidence from China," *Journal of Environmental Planning and Management*, vol. 65, no. 2, pp. 240–271, 2022.

[18] M. Zhang and B. Li, "How to design regional characteristics to improve green economic efficiency: a fuzzy-set qualitative comparative analysis approach," *Environmental Science and Pollution Research*, vol. 29, no. 4, pp. 6125–6139, 2022.

[19] S. Naseer, H. Song, M. S. Aslam, D. Abdul, and A. Tanveer, "Assessment of green economic efficiency in China using analytical hierarchical process (AHP)," *Soft Computing*, vol. 26, no. 5, pp. 2489–2499, 2022.

[20] Z. Wang, X. Wang, and L. Liang, "Green economic efficiency in the Yangtze River Delta: spatiotemporal evolution and influencing factors," *Ecosystem Health and Sustainability*, vol. 5, no. 1, pp. 20–35, 2019.

[21] Q. Li, "Regional technological innovation and green economic efficiency based on DEA model and fuzzy evaluation," *Journal of Intelligent and Fuzzy Systems*, vol. 37, no. 5, pp. 6415–6425, 2019.

[22] X. Zheng, H. Yu, and L. Yang, "Technology imports, independent innovation, and China's green economic efficiency: an analysis based on spatial and mediating effect," *Environmental Science and Pollution Research*, vol. 29, no. 24, pp. 36170–36188, 2022.

[23] G. Zeng, C. Geng, and H. Guo, "Spatial s effect of strategic Emerging Industry agglomeration and green Economic Efficiency in China," *Polish Journal of Environmental Studies*, vol. 29, no. 5, pp. 3901–3914, 2020.

[24] S. S. Ganji and A. A. Rassafi, "Road safety evaluation using a novel cross efficiency method based on double frontiers DEA and evidential reasoning approach," *KSCE Journal of Civil Engineering*, vol. 23, no. 2, pp. 850–865, 2019.

Hindawi
Computational Intelligence and Neuroscience
Volume 2022, Article ID 8937314, 10 pages
https://doi.org/10.1155/2022/8937314

Research Article

Construction of a Prediction Model for Distance Education Quality Assessment Based on Convolutional Neural Network

Peizhang Wang

School of Management, Anhui Science and Technology University, Chuzhou, Anhui 233100, China

Correspondence should be addressed to Peizhang Wang; wangpz@ahstu.edu.cn

Received 4 July 2022; Revised 2 August 2022; Accepted 6 August 2022; Published 5 September 2022

Academic Editor: Ning Cao

This paper introduces the principles and operation steps of convolution and pooling of convolutional neural networks in detail. In view of the shortcomings of fixed sampling points and single receptive field in traditional convolution and pooling forms, deformable convolution and deformable pooling are introduced to enhance the network's ability to adapt to image details and large displacement problems. The concepts of warp, loop optimization, and network stack are introduced. In order to improve the optimization performance of the algorithm, three subnetwork structures and stack models are designed, and various methods are used to improve the prediction accuracy of distance education quality assessment. In order to improve the accuracy and timeliness of education quality assessment, this paper proposes a distance education quality assessment model based on mining algorithms. The prediction index is selected by the improved BP neural network. It is required to establish the input layer node as the input vector based on the number of data sources since the input layer is used for data input. The neural network is trained with a quarter of the mining data, and the mining algorithm is further trained with network error trials. A fuzzy relationship matrix is created based on the assessment of teaching quality's hierarchical structure. This leads to the conclusion of the fuzzy thorough evaluation of the effectiveness of distant learning. Experiments show that the proposed model has an average accuracy of 96%, the average teaching quality modeling time is 25.44 ms, and the evaluation speed is fast.

1. Introduction

Distance education is a type of education that transmits courses to one or more students away from the campus using a variety of media for systematic teaching and communication. Nowadays, remote learning refers to learning that is done by computer technology, both real-time and nonreal-time, audio, and video (live or video). Modern distant learning is a brand-new educational format modified for use with contemporary information technology [1]. In May 2020, the Ministry of Education revealed at a press conference that online teaching will become the new normal after resumption of school in colleges and universities. Our country's higher education enterprise has obtained the rapid development, accepted the higher education, and the object to have the significant increase. The number of university teachers has also increased significantly, and the effects of higher education have had a significant impact on the country's economic development [2, 3]. College teachers are both educators and researchers, and the requirements for teachers have reached a new height [4–6].

For neural networks, people often hope that the established network model can focus on the most representative feature information in a certain field, so as to avoid the problem of overfitting and improve the performance of the network model. The scale is getting bigger and bigger, and there are more and more internal parameters, and the problem of information overload caused by it also needs to be solved urgently. By introducing the attention mechanism, the above two types of problems can be solved, focusing on the most useful information for the task among the many information, while reducing the attention to the secondary information and eliminating irrelevant information, which can improve the accuracy of the model and avoid information overload question.

The commonly used attention models include temporal attention and spatial attention, and the specific function of attention can be divided into hard attention mechanism and soft attention mechanism. The hard attention mechanism means that in front of multiple objects or features, only a weight value of 1 is taken, while the others are 0, that is, only one point is concerned; while the soft attention mechanism is to take the weight of multiple objects or features. The value setting range, such as [0, 1], is only divided into primary and secondary instead of partial generalization, so the soft attention mechanism is the most commonly used attention mechanism. Since the advantages of the attention mechanism are applicable to regular data, it is most widely used in the research of image processing and computer vision.

In order to systematize the management, office automation and informationization of teacher evaluation in colleges and universities, it is necessary to conduct research and discussion in all aspects. The content and system of teacher assessment, as well as data processing and analysis of assessment results, need further study. How to make comprehensive use of pedagogy, statistics, information network technology, and computer technology to design a high-efficiency and high-quality system is an important content of teacher management and evaluation. This research suggests a mining algorithm-based methodology for evaluating the quality of distant education in order to increase its accuracy and efficacy.

2. Method

2.1. Data Mining Algorithms. It has the advantage of allowing mining computations to be performed using data stored in the current information system, and to encapsulate complex statistical techniques and mining algorithms through computer application programs. Even without mastering these techniques, the same function can be accomplished, thereby focusing more on the problem to be solved [7, 8]. In light of the aforementioned factors, this work intends to incorporate data mining technology into the field of teaching assessment and investigate a model for evaluating the quality of distant education.

At present, artificial neural networks are a fast growing area of frontier research that have profound effects on cognitive science, information technology, and artificial intelligence in computer science. Additionally, it is crucial for data mining [9, 10]. For a very long time, BP neural networks have been applied in data mining. However, the modified approach has to be improved because it has a problem with local minimization.

The selection of the predictive indicator, which is based on the performance of the students in distance education, is the first stage in building the improved BP neural network model. Next, the vectors for the output and input are created [11, 12]. Input layer nodes must be set as input vectors in accordance with the amount of data sources since enter layer is used to input data. Table 1 outlines the procedures for calculating the number of data sources.

The excitation function of the input layer node is as follows:

$$k(v) = \frac{1}{\exp(-\beta)v^4}. \tag{1}$$

Among them, $k(v)$ represents the activation function of the input layer node. v represents the function argument. β represents the slope control parameter.

The neuron corresponding to the input layer is the input vector. The following formula can be used to determine the output vector:

$$O_i = k(v) \cdot f\left(\sum_{i=1}^{n} W_i \theta_i\right)^2, \tag{2}$$

where O_i represents the output of the output vector. W_i represents the weight of the output node. θ_i represents the threshold of the output layer.

Then, for the improved BP neural network, choose the number of layers or adjust the number of neurons in the hidden layer. Then, choose the corresponding hidden layer nodes, that is, begin to invest in a few hidden layer nodes [13, 14]. Then gradually increase the number of invested nodes until the number of nodes is more manageable. This procedure must be tested. The hidden layer node's corresponding output formula is as follows:

$$y_i = O_i f\left(\sum_{j=1}^{n} w_{i\,j} x_j - \theta_i\right), \tag{3}$$

where y_i represents the corresponding output of the hidden layer node. w_{i_j} represents the corresponding network weight of the hidden layer. x_j stands for input node.

Using a quarter of the mining data, that is, training samples, to test and improve the BP neural network, first put eight neurons, and then increase to fifteen, and conduct tests, respectively [15]. Figure 1 displays the specific training results. Choose the number of neurons that corresponds to the best combination of training time and training error.

According to the above table, when the number of neurons is 15, the error can be kept to a minimum.

On this basis, different training functions are used to detect network errors, and a new BP neural network model is presented. A correction algorithm is used to optimize the design of the BP neural network based on the experimental data [16, 17]. Tests for network failures are shown in Table 2.

Choose lm to improve the BP neural network based on the results in the preceding table. And, based on the results in Tables 2 and 3, a three-layer improved BP neural network model is built. The model's network structure is 4, 15, and 1, with a total of four input nodes [18–20]. There is one output node and fifteen hidden nodes.

$$\begin{cases} f\left(L^{(k+1)}\right) = \min f\left(X^{(k)} + S\left(L^{(k)}\right)\right)^2 \eta^{(k)}, \\ L^{(k+1)} = L^{(k)} + \eta^{(k)} S\left(L^{(k)}\right). \end{cases} \tag{4}$$

Among them, $L^{(k)}$ represents the threshold of the network and the vector of all values. $S(L^{(k)})$ represents the search direction of the space vector formed by each function component in X. $\eta^{(k)}$ stands for the minimum step size.

Table 1: The process of identifying data sources.

Step	Specific contents
Fit state established	Establish fit state in simulation environment
Data denoising	Clear wrong data
Data source exclusion	Exclude unreliable or marginal data sources
Algorithm choice	Select an algorithm for preparing data. Particularly the method that makes up the difference for the data deficiency

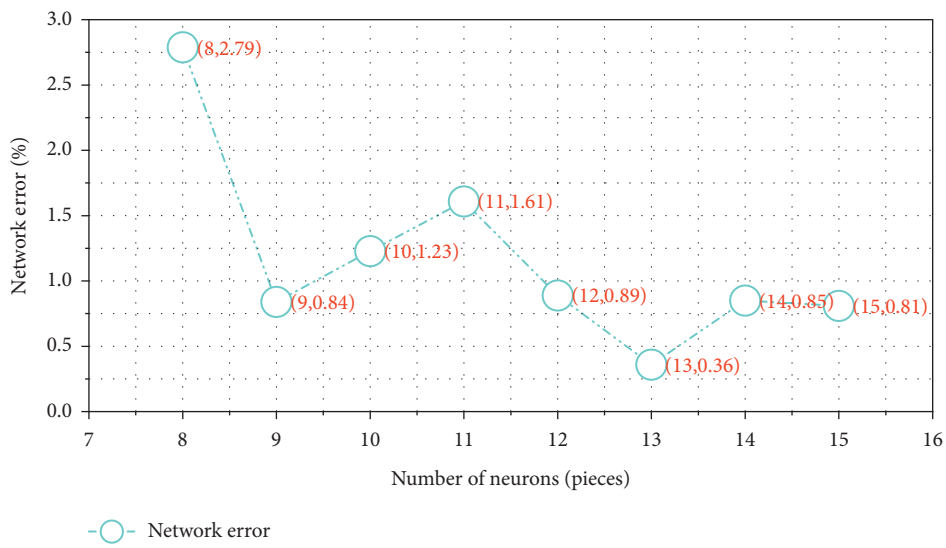


Figure 1: Training results.

Table 2: Network error test results.

Training function name	dm	da	Dx	lm
Average network error	0.0014	0.0109	0.0041	0.0007

Table 3: Teaching quality evaluation index hierarchy.

Target layer	Criterion layer	Indicator layer
Teaching quality	Teacher teaching situation	T11 well prepared before class T12 The main points of the explanation are highlighted T13 link theory with practice T14 teacher-student interaction T15 modern teaching methods T16 focus on ability development T17 caring for students T18 teacher table T19 The overall effect of teaching is good
	Course information	T21 course content T22 course load

2.2. A Model for Evaluating the Quality of Distance Learning. A system of excellent teaching quality evaluation indexes must have the qualities of authenticity, specificity, and convenience. Table 3 summarizes and categorizes the elements of all teaching quality evaluation systems, displaying a set of comprehensive, multifaceted, and multilevel three-dimensional evaluation index systems.

$r_{ij}\,(i = 1, 2, \cdots, m;\ j = 1, 2, \cdots, n)$ describes the result of the ith factor of the teacher being evaluated since the first factor and establishes the following fuzzy evaluation matrix:

$$R = \begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1n} \\ r_{21} & r_{22} & \cdots & r_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ r_{m1} & r_{mw} & \cdots & r_{mn} \end{bmatrix}, \tag{5}$$

$S = (s_1, s_2, \cdots, s_m)$ is a very important set of B. When both R and S are known, both can be blurredly transformed to obtain model $W = (R \times S) = (d_1, d_2, \cdots, d_n)$[18, 21, 22].

Construct a comprehensive evaluation factor set Z.

Many different elements are frequently included in an excellent teaching quality rating method. It will be challenging to analyze, nevertheless, if the criteria to be taken into account are overly extensive. There is established a set of evaluation criteria for thorough consideration: $X =$ "The lesson preparation is thorough, and the teaching style is appropriate." The emphasis in the classroom is on helping students understand; engaging in classroom discussion; and energizing the learning environment. Be tough with yourself and set an example for kids. Teach students according to their aptitude. Focus on ability training. Care for students and be genuinely appreciated by them".

Build an evaluation set Y.

The evaluation set can serve as a reflection of a teacher's instructional ability. The evaluation set commonly divides teachers' teaching levels into three categories: "outstanding," "good," "average," and "bad." As a result, each level contained in Y can only have values inside a specific range during the actual evaluation. The text range can be set as [50, 100] and divided into four decreasing intervals based on the analysis mentioned above. If the obtained score falls within [80, 100], it can be set at 90, designating the teacher as having "outstanding" teaching status. If the obtained score falls within [70, 90], it can be set to 80, indicating that the teacher's performance as a teacher is "good." The teacher's teaching status may only be described as "average" if the obtained score is placed in the range [60, 80], since that score's median value is 70. If the obtained score falls between [50 to 70], the score might be set to 60, indicating that the teacher's teaching condition is "poor," or not good. In reality, though, fractional division is based on the middle value of each fractional interval, and the parameter column vector can be adjusted to $Y = [90, 80, 70, 60]$.

Establishing the evaluation factor set's weight set is crucial since each factor holds a unique key location inside the evaluation factor set, meaning that each factor's weight is different. How to assess the accuracy of evaluation results by judiciously allocating the weight of each component. The Delphi technique is a reliable system for determining weight. The main principle of the Delphi method is to collect anonymous expert opinions, analyze them, and then transmit them to other experts in a similar manner.

Build a fuzzy relationship matrix.

Based on teachers' daily behavior in a variety of criteria, all assessors provide objective evaluations. By using induction and sorting, a fuzzy relationship matrix R is created, and the possibility measure is then derived. If there are 200 people participating in the evaluation of a certain aspect of teacher evaluation, of which 140 were rated as "excellent", 40 were rated as "good", 20 were rated as "general", the probability of the result being "excellent" is $140/200 = 0.7$. The probability of good is $40/200 = 0.2$. The probability of "general" is $20/200 = 0.1$. The probability of "poor" is $0/200 = 0$. Obtain the results of the fuzzy through evaluation.

$$W = (R \times S) = (d_1, d_2, d_3, d_4). \tag{6}$$

The value obtained by using $S = (W \times V) = 90 * d_1 + 80 * d_2 + 70 * d_3 + 60d_4$ can serve as the teacher being evaluated's final assessment score. The range in the assessment set Y allows one to determine the teacher's teaching quality rating.

2.3. Overall Architecture of Convolutional Neural Network. The network model in this paper is optimized and improved based on the basic model of the convolutional neural network, including the feature extraction part, the feature fusion part, and the final optical flow output part. The process is shown in Figure 2.

The function of the feature extraction module is to extract the spatial information features of two adjacent frames in a continuous image sequence. It is mainly composed of a deformable convolutional layer and a traditional square convolutional layer. Its input is the adjacent two frames of RGB images I1 and I2. Two branches with the same structure and independent parameter weights extract the spatial features of the two frames to ensure that the feature dimensions of the output are consistent; the improvement of this part lies in the introduction of a deformable convolution layer, compared with the existing traditional square convolution layer. The composed convolutional neural network feature extraction module can better extract the image detail features of adjacent frames and improve the ability to capture large displacements, so as to solve the problems of large displacements and image detail rendering in optical flow prediction.

The feature fusion part can fuse the spatial features of two deep adjacent frames extracted from continuous image sequences into one feature, and calculate the correlation between the two frame features at the same time, which is mainly composed of a convolution layer based on the channel attention mechanism. The importance of each channel is calculated, and additional weights are added to each channel to reconstruct the features; the improvement of this part is reflected in the feature fusion method using the attention mechanism. Compared with the existing matching-based fusion method, the correlation between two frames can be better calculated. The key to optimizing the occlusion problem lies in the calculation of deep feature correlation. Therefore, the improvement of this part can, in principle, enhance the network in a targeted manner.

In order to improve the prediction effect of the model for the problems of occlusion, large displacement, and image detail rendering between consecutive frame images, and at the same time to ensure its universality, it has excellent accuracy for various images and motion forms. Using the strategy of cyclic optimization, through the improvement of the DANet-S network structure, a total of three subnetworks with different characteristics are designed, and then multiple subnetworks are connected, and the output optical flow of each layer of network and the actual optical flow are calculated through Warp. The loss of the flow enables the lower network to focus on learning this part of the loss, thereby

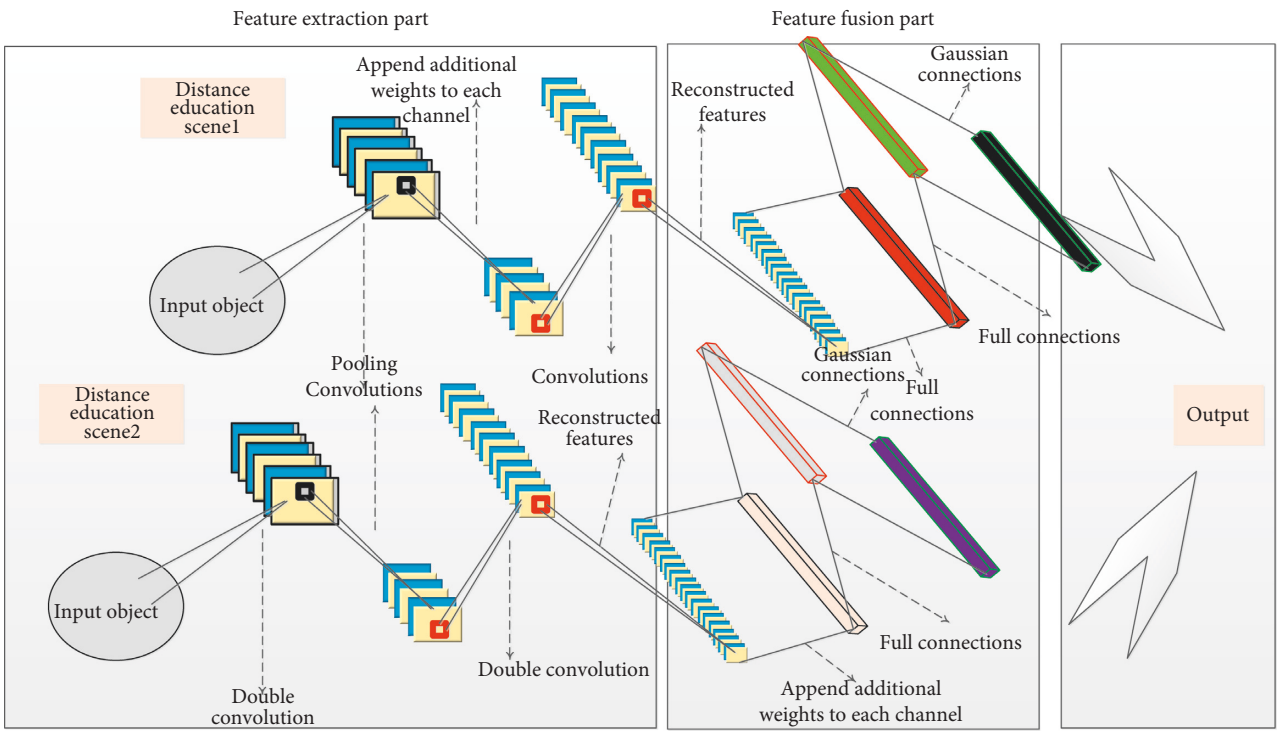


Figure 2: DANet-S structure.

improving the prediction accuracy of the method in this paper.

2.4. Interframe Feature Association Layer Based on Attention Mechanism. The optimization of the feature fusion part of the convolutional neural network for optical flow prediction is the key to solving the occlusion problem. Shallow-level spatial feature extraction cannot accurately characterize the occluded part of the image. It is necessary to calculate the correlation between two adjacent frames in the deep-level feature, so as to find the motion and correspondence of the occluded part in the two images.

The matching mechanism employed by the model is difficult to capture the complex motion between adjacent frames. The attention mechanism is suitable for the convolutional neural network to solve targeted and focused image problems. Therefore, in order to better calculate the correlation between the spatial features of two adjacent frames, we improve the network model's ability to deal with the problem of image object occlusion. In this paper, a feature correlation layer based on the channel attention mechanism is designed. Through this module, the correlation of the features of two adjacent frames is calculated, and the fusion reconstruction is performed to enhance the network's processing ability to the occlusion problem.

The attention mechanism is like when the human eye or a camera observes or shoots an object, its focus must be focused on the object of interest, while ignoring other objects or backgrounds. It is blurred, and for an intelligent human brain, it will even directly filter out irrelevant information such as the background, so as to eliminate interference; similarly, when the target object changes, its attention will also shift and focus on another target. This means that the spatial distribution of attention or attention points is different, and the importance of each object or background is proportional to the distribution of attention; in terms of time, it is impossible for every sentence of a piece of news to have the same amount of attention. Equally important, some segments are the key information of the reported event, while some segments are modifiers that are not related to the event. The distribution of these segments in time is different, which also leads to the listener's attention on the time series.

When processing text, the input sentence is composed of multiple words, and the words have different grammatical status and different importance in understanding the meaning of the sentence. The Decoder framework treats these words equally, extracts semantic information in the same way for each word, then converts it into the output word, and then combines it into a complete sentence. When dealing with image data or image features, this framework also treats each region of the image or feature spatial distribution as "the same", and the next step is processed with the same weight for information in different spatial or channel positions.

Education quality evaluation is based on certain educational quality standards and goals, and relies on scientific educational evaluation methods and methods to make value judgments on the development status and influence of the evaluation object. Evaluation can provide reliable information for human development and education decision-making and assist education decision-makers in choosing the best strategies. The distance education quality assessment

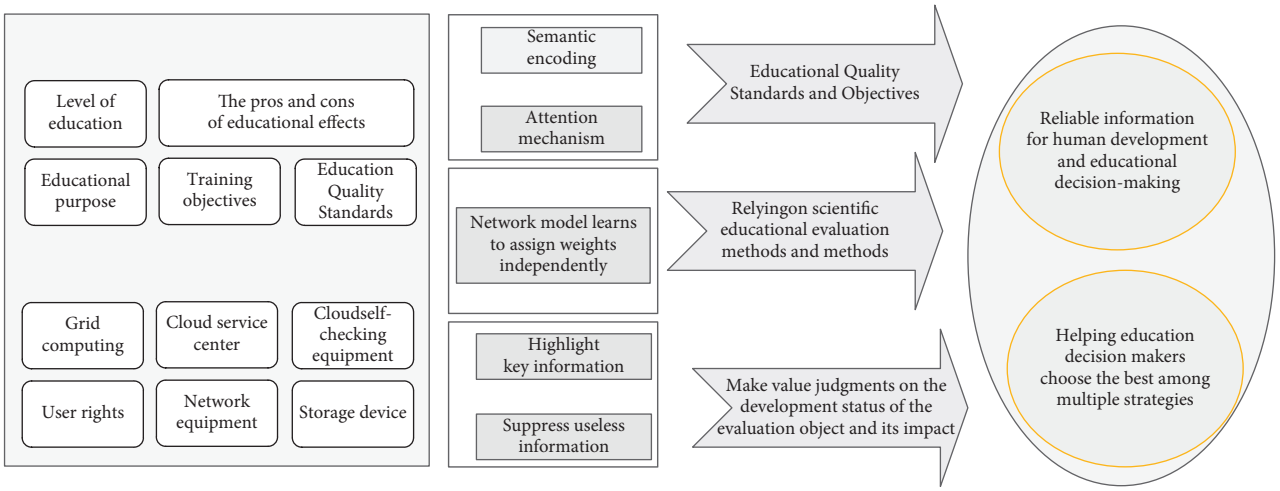


Figure 3: Distance education quality assessment mode based on convolutional neural network.

mode based on convolutional neural network is shown in Figure 3.

3. Experiment and Analysis

The purpose of this experiment is to explore the application effect of data mining methods in the evaluation of college education quality. This paper selects three types of distance online teaching platforms with prominent styles for testing: the first type of distance online teaching platform is characterized by high quality and relatively rich course content, and the instructors are experienced teachers. For example, NetEase Open Course, MOOC., which we call the content class. The second category is characterized by strong interactivity and timely feedback, usually live teaching, such as QQ groups and Tencent conferences. This article calls it the interactive category. The third type of platform is compatible with both and meets a certain degree of interactivity. It can record live courses and has playback functions, such as Tencent Classroom and DingTalk Live. This article calls them compatible classes.

The users of the online teaching platform are relatively fixed. This paper randomly selects 100 teachers and students for evaluation by visiting 10 colleges and universities in a province and sending emails to users of the online teaching platform. 220 evaluation samples were scored by 100 users, and the specific distribution of their identity, gender, education level, and subject area is shown in Table 4. It can be seen from the distribution that the survey sample is representative to a certain extent and can better evaluate the use of the online teaching platform.

This paper uses the methods of Reference [5] and Reference [6] as the control group to verify the three methods to evaluate the quality of distance teaching of teachers in the 10 universities mentioned above.

3.1. The Accuracy of Distance Teaching Quality Assessment. The accuracy of distance teaching quality assessment of each university is counted, and the results are shown in Figure 4.

Table 4: Basic data.

Basic situation	Category	Frequency
Gender	Male	47
	Female	53
Used identity	Teacher	21
	Student	79
Academic area	Natural science	59
	Social science	41
Education level	Bachelor degree and below	57
	Graduate student or above	43

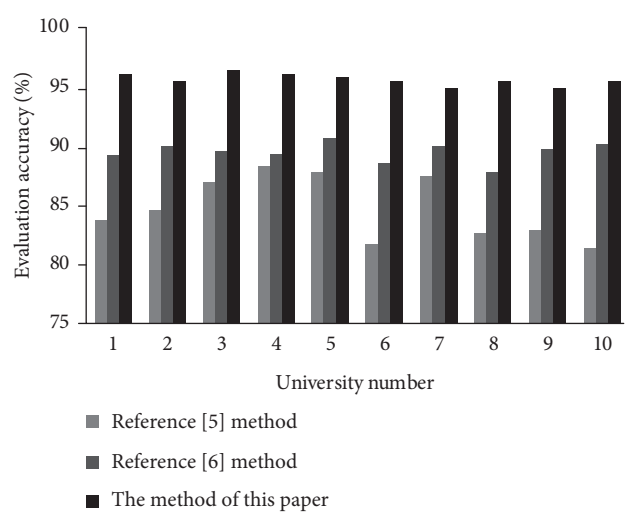


Figure 4: Comparison of the accuracy of distance teaching quality assessment.

We analyze the results of the accuracy rate of teaching quality evaluation in Figure 1:

(1) The teaching quality evaluation mistake is the biggest, and the approach in Reference [5] has an

average evaluation accuracy of 85%. This is due to the fact that its parameters are chosen at random, making it impossible to develop an evaluation model that captures the intricate and ever-changing aspects of teaching quality. As a result, teaching quality assessment results in many overfitting points and has the worst teaching quality assessment effect.

(2) The average accuracy of the distance teaching quality assessment method in Reference [6] is 90%, and the teaching quality assessment error is smaller than the standard BP neural network. This is because the support vector machine is introduced, and a better teaching quality evaluation model than the standard BP neural network is established, which can effectively describe the complex and changing characteristics of teaching quality. Due to the existence of some underfitting points in teaching quality evaluation, the evaluation effect needs to be further improved.

(3) The two competing approaches of distant mathematics teaching quality evaluation accuracy are far from the average teaching quality evaluation accuracy of 96 percent used in this paper. This is because the model in this paper introduces the BP neural network to select the predictor, and uses a quarter of the mining data to train the neural network, and further trains the mining algorithm through the network error test. Therefore, an evaluation model that can adjust the complex characteristics of optimal educational change in an overly precise manner is developed. The model effectively improves the evaluation effect of distance teaching quality.

3.2. Efficiency of Distance Teaching Quality Assessment. In practice, with the continuous increase of data, the evaluation of teaching quality in colleges and universities has also received increasing attention. In order to analyze the validity of university education quality evaluation, the modeling time of university teaching quality under different modes was calculated. The results are shown in Figure 5.

It can be considered from Figure 2 that the suggest time of the approach in Reference [5] is 62 ms, and the suggest time of the educating high-quality modeling in Reference [6] is 39.27 ms. The average teaching quality modeling time of this method is 25.31 ms. The comparison results show that the approach described in this paper's teaching quality modeling time is significantly reduced, and the effectiveness of teaching quality evaluation in remote institutions is enhanced.

3.3. Prediction Results and Analysis of Distance Education Quality Assessment. The model in this paper is mainly built on the public server of the research group. The processor of the deep learning server uses Intel Core I7-5960X (main frequency: 3.0 GHZ), the model of the graphics accelerator is NVIDIA GeForce GTX1080Ti, and the operating system is

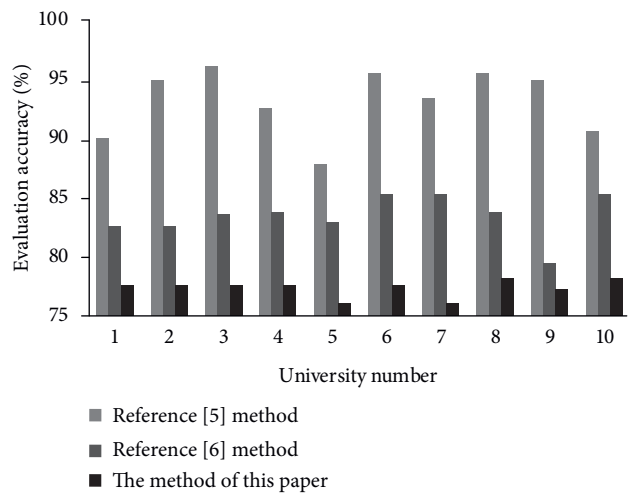


Figure 5: Efficiency comparison of distance teaching quality assessment.

Linux (Ubuntu version is 18.04), the Cuda version is 9.0, the program is built and trained under the Pytorch 0.4.1 framework, and the main programming languages used are python and linux scripting languages.

The network is trained and adjusted in the form of supervised learning, and the average end-point error of the predicted optical flow and the real optical flow is used as the loss function of the network backpropagation, that is, the comparison between the ground truth from the dataset and the image after the predicted optical flow interpolation. The data set and strategy used for training will greatly affect the network performance.

The prediction robustness and $F1$ value of the distance education quality assessment of the three algorithms are shown in Figures 6 and 7, respectively.

3.4. General Analysis of Distance Teaching Quality Assessment. Ten courses from a university were chosen as the test items, and their distance teaching quality assessment accuracy and modeling time were tallied in order to investigate the generality of the data mining algorithm's distance teaching quality assessment model. Table 5 presents the outcomes.

As shown in Table 5, the model in this research not only produces highly accurate findings for the quality evaluation of distance teaching, but also that the evaluation speed is rather quick. The approach can be used to evaluate the effectiveness of teaching in colleges and universities in a way that is practical and applicable.

4. Discussion

In this paper, deformed convolution, pooling, and channel attention mechanisms are introduced into the convolutional neural network model. In principle, the feasibility of deformed convolution and pooling to solve complex motion problems is analyzed, and a feature extraction part based on deformed convolution is established. Furthermore, the superiority of the attention mechanism in the field of optical

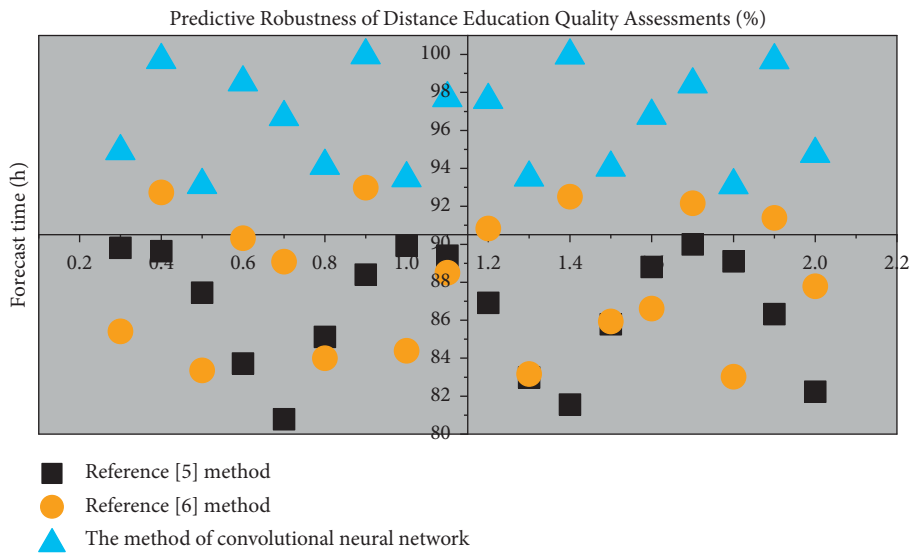


Figure 6: Prediction robustness of distance education quality assessment for three algorithms.

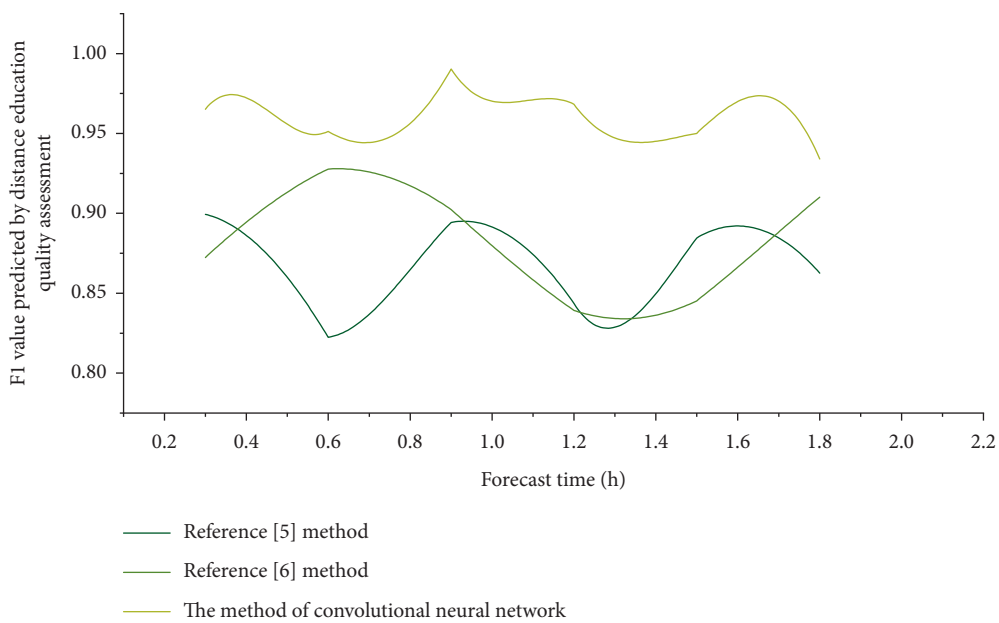


Figure 7: $F1$ value predicted by distance education quality assessment of three algorithms.

Table 5: The generality of distance teaching quality assessment model.

Course title	Evaluation accuracy/ %	Modeling time/ s
University English	95.23	22.61
Communication principle	94.62	23.53
Linear algebra	94.84	23.86
Engineering mechanics	93.65	22.47
University Chinese	95.13	26.95
Basic computer science	96.35	24.67
Machine learning	95.29	23.29

flow prediction is clarified, which realizes the feature association layer based on the channel attention mechanism.

The initial parameters of the network model in this paper are uniformly set through the same initialization method. The network level and cascade method are all based on experience and experiments, and lack theoretical support. This is also the current deep learning in the network structure. A series of related theories should be established to solve common problems in parameter optimization, and there should be corresponding standards for solving problems of different scales and types.

The deformable convolution and channel attention mechanisms used in the model in this paper have targeted optimization for issues such as occlusion, large displacement, and image detail presentation in the field of optical flow prediction, but other related issues are not covered. For the prediction of single-frame image optical flow, multi-target motion and weak texture, no effective solutions have been proposed. These issues should also be the focus of future research.

Because the convolutional neural network has a certain randomness in the training convergence process, especially when the data set is not sufficient, it is easy to lead to local convergence, and the global optimal solution cannot be obtained, which leads to poor stability of network training. The proposed optical flow prediction network model introducing deformable convolution and channel attention mechanism also suffers from similar problems. Therefore, it is necessary to conduct more in-depth research and exploration on the training and convergence methods of the network, and improve the stability and accuracy of network training from the aspects of network structure, propagation mode, basic composition, and optimization training strategy. This has important practical significance for the research and application of optical flow prediction algorithms, and has more practical significance for the development of neural networks, deep learning and artificial intelligence, and is the next development direction of these fields.

Teachers' teaching and research work is the guarantee of school education quality. In the research work, teachers evaluate the teachers who have listened to the class, point out their shortcomings, and praise their strengths. The teachers who evaluate the class not only reflect on their own shortcomings in the process of pointing out the teachers who are listening to the class but also can improve their own teaching. Teachers express their opinions in teaching and research activities, and different teachers have different understanding levels of the same content. An exchange between different teaching methods, discussing different courses using different teaching methods can improve teachers' teaching ability, thereby improving their teaching ability.

5. Conclusion

The basic concepts and principles of the attention mechanism are introduced, the structural flow of the Encoder-Decoder framework is analyzed in detail, and the application form of the attention mechanism in the convolutional neural network is expounded. This paper proposes a feature correlation layer based on the channel attention mechanism, which improves the network adjustment and adaptation ability and correlation calculation ability without increasing too much computational cost. The cause of teaching reform in colleges and universities must be advanced day by day due to the ongoing advancement of the times. How to raise teaching standards is the most crucial challenge in the reforming process. We can only introduce exceptional potential to society by reforming colleges and institutions. In order to develop a mathematical model for assessing the effectiveness of distant education, this research makes use of data mining technology. After the number of samples is determined, the BP neural network based on 1/4 mining is used for training, and the network error experiment is carried out. On this basis, the quality of distance education is evaluated by fuzzy correlation matrix. According to experimental data, the model's average assessment accuracy is 96 percent, the average teaching quality modeling time is 25.44 milliseconds, and the evaluation speed is higher. The model developed in this research has certain advantages over the other models in terms of evaluating the quality of remote education.

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.

References

[1] S. Aziz, M. Mahmood, and Z. Rehman, "Implementation of CIPP model for quality evaluation at school level: a case study," *Journal of Education and Educational Development*, vol. 5, no. 1, pp. 189–206, 2018.

[2] L. Shen, J. Yang, and X. Jin, "Based on Delphi method and analytic hierarchy process to construct the evaluation index system of nursing simulation teaching quality," *Nurse Education Today*, vol. 79, pp. 67–73, 2019.

[3] M. Daumiller, R. Rinas, J. Hein, S. Janke, O. Dickhauser, and M. Dresel, "Shifting from face-to-face to online teaching during COVID-19: the role of university faculty achievement goals for attitudes towards this sudden change, and their relevance for burnout/engagement and student evaluations of teaching quality," *Computers in Human Behavior*, vol. 118, Article ID 106677, 2021.

[4] X. Zhang and K. Wu, "The construction of evaluation model of Chinese traditional culture multimedia teaching resources allocation in big data environment," in *Proceedings of the 2018 International Conference on Intelligent Transportation, Big Data & Smart City (ICITBS)*, pp. 133–136, IEEE, Xiamen, China, January 2018.

[5] W. Huang, "Simulation of English teaching quality evaluation model based on Gaussian process machine learning," *Journal of Intelligent and Fuzzy Systems*, vol. 40, no. 2, pp. 2373–2383, 2021.

[6] S. Qianna, "Evaluation model of classroom teaching quality based on improved RVM algorithm and knowledge recommendation," *Journal of Intelligent and Fuzzy Systems*, vol. 40, no. 2, pp. 2457–2467, 2021.

[7] A. Cervero, A. Castro-Lopez, L. Álvarez-Blanco, M. Esteban, and A. Bernardo, "Evaluation of educational quality performance on virtual campuses using fuzzy inference systems," *PLoS One*, vol. 15, no. 5, Article ID e0232802, 2020.

[8] C. Guo and Y. Liu, "BP neural network-based evaluation on university teachers' teaching quality," in *Proceedings of the 2020 3rd International Conference on E-Business*, pp. 285–289, Wuhan China, December 2020.

[9] L. Wenwen, "Modeling and simulation of teaching quality in colleges based on BP neural network and training function,"

Journal of Intelligent and Fuzzy Systems, vol. 37, no. 5, pp. 6349–6361, 2019.

[10] Y. Jiang, J. Zhang, and C. Chen, "Research on a new teaching quality evaluation method based on improved fuzzy neural network for college English," *International Journal of Continuing Engineering Education and Life Long Learning*, vol. 28, no. 3/4, pp. 293–309, 2018.

[11] L. Yuan, Z. Xiaofei, and Q. Yiyu, "Evaluation model of art internal auxiliary teaching quality based on artificial intelligence under the influence of COVID-19," *Journal of Intelligent and Fuzzy Systems*, vol. 39, no. 6, pp. 8713–8721, 2020.

[12] C. Lu, B. He, and R. Zhang, "Evaluation of English interpretation teaching quality based on GA optimized RBF neural network," *Journal of Intelligent and Fuzzy Systems*, vol. 40, no. 2, pp. 3185–3192, 2021.

[13] C. Liu, Y. Feng, and Y. Wang, "An innovative evaluation method for undergraduate education: an approach based on *BP* neural network and stress testing," *Studies in Higher Education*, vol. 47, no. 1, pp. 212–228, 2022.

[14] P. Fan, "Application of deep learning and cloud data platform in college teaching quality evaluation," *Journal of Intelligent and Fuzzy Systems*, vol. 39, no. 4, pp. 5547–5558, 2020.

[15] X. Peng and J. Dai, "Research on the assessment of classroom teaching quality with q -rung orthopair fuzzy information based on multiparametric similarity measure and combinative distance-based assessment," *International Journal of Intelligent Systems*, vol. 34, no. 7, pp. 1588–1630, 2019.

[16] A. K. Shukla, P. Singh, and M. Vardhan, "An adaptive inertia weight teaching-learning-based optimization algorithm and its applications," *Applied Mathematical Modelling*, vol. 77, pp. 309–326, 2020.

[17] C. Troussas, A. Krouska, C. Sgouropoulou, and I. Voyiatzis, "Ensemble learning using fuzzy weights to improve learning style identification for adapted instructional routines," *Entropy*, vol. 22, no. 7, 735 pages, 2020.

[18] Z. Sun, M. Anbarasan, and D. Praveen Kumar, "Design of online intelligent English teaching platform based on artificial intelligence techniques," *Computational Intelligence*, vol. 37, no. 3, pp. 1166–1180, 2021.

[19] J. W. Gong, H. C. Liu, X. Y. You, and L. Yin, "An integrated multi-criteria decision making approach with linguistic hesitant fuzzy sets for E-learning website evaluation and selection," *Applied Soft Computing*, vol. 102, Article ID 107118, 2021.

[20] B. Hu, "The evaluation method of English teaching efficiency based on language recognition technology," *International Journal of Continuing Engineering Education and Life Long Learning*, vol. 30, no. 4, pp. 445–459, 2020.

[21] L. Lin, "Smart teaching evaluation model using weighted naive bayes algorithm," *Journal of Intelligent and Fuzzy Systems*, vol. 40, no. 2, pp. 2791–2801, 2021.

[22] E. T. Lau, L. Sun, and Q. Yang, "Modelling, prediction and classification of student academic performance using artificial neural networks," *SN Applied Sciences*, vol. 1, no. 9, pp. 1–10, 2019.

Hindawi
Computational Intelligence and Neuroscience
Volume 2023, Article ID 9843832, 1 page
https://doi.org/10.1155/2023/9843832

Retraction

Retracted: Application of CNN-Based Machine Learning in the Study of Motor Fault Diagnosis

Computational Intelligence and Neuroscience

Received 17 October 2023; Accepted 17 October 2023; Published 18 October 2023

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

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

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

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

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

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

References

[1] X. Peng, L. Wei, and W. Gao, "Application of CNN-Based Machine Learning in the Study of Motor Fault Diagnosis," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 9635251, 9 pages, 2022.

Hindawi
Computational Intelligence and Neuroscience
Volume 2022, Article ID 9635251, 9 pages
https://doi.org/10.1155/2022/9635251

Research Article

Application of CNN-Based Machine Learning in the Study of Motor Fault Diagnosis

Xiuyan Peng, Lunpan Wei, and Wei Gao

College of Intelligent Systems Science and Engineering, Harbin Engineering University, Harbin, Heilongjiang 150001, China

Correspondence should be addressed to Lunpan Wei; weilunpan@hrbeu.edu.cn

Received 13 June 2022; Revised 26 July 2022; Accepted 5 August 2022; Published 5 September 2022

Academic Editor: Ning Cao

With the development of science and technology, the rapid development of social economy, the motor as a new type of transmission equipment, in the production and life of people occupies a pivotal position. Under the rapid development of computer and electronic technology, manufacturing equipment is becoming larger, faster, more continuous, and more automated. This has resulted in complex, expensive, accident-damaging, and high-impact equipment for electric motors; even routine maintenance requires significant equipment maintenance and maintenance costs. If a fault occurs, it will cause serious damage to the entire equipment and can even have a major impact on the entire production process, leading to a serious economic and social life. In this paper, a CNN-based machine learning fault diagnosis method is proposed to address the problem of high incidence of motor faults and difficulty in identifying fault types. A fault reproduction test is constructed by machine learning techniques to extract vibration time domain data for normal operating conditions, rotor eccentricity, stator short circuit, and bearing inner ring fault; divide the data segment into 15 speed segments, extract 13 typical time domain features for each speed segment; and perform mathematical statistics for fault diagnosis. Compared with the traditional algorithm, the method has more comprehensive feature information extraction, higher diagnostic accuracy, and faster diagnostic speed, with a fault diagnosis accuracy of 98.7%.

1. Introduction

As a kind of mechanical equipment to realize energy conversion, electric motor is commonly used in practical production with its simple structure and high operating efficiency, and it has become the most widely used power equipment at present. Due to the specificity of the working principle of electric motor, it is often complicated in structure and harsh in working condition. Usually, its structure is complex and the working conditions are uncontrollable, so the failure rate is high. However, once the failure occurs [1], the normal production and economic efficiency will bring serious impact, and even to the safety of the operator brings a direct threat. Therefore, the implementation of effective fault diagnosis of the motor is essential for its normal operation.

At present, most enterprises still adopt the method of scheduled maintenance for large motors, and the time point of maintenance depends on the accumulated running time of the motor. When the abnormality occurs, the maintenance and repair will be carried out, and this method relies excessively on the experience of engineers [2]. Using any of the abovementioned methods for motor maintenance can no longer meet the current requirements, and the phenomenon of motor overhaul and disrepair still exists widely. Therefore, condition monitoring should be carried out to effectively improve the accuracy of fault diagnosis and condition prediction [3]. The promotion of condition maintenance has become a realistic requirement for enterprise production and an inevitable development trend for equipment management and maintenance work.

According to the report of EPRI, 53% of the motor failures originate from mechanical causes, such as bearing failure, unbalance, and looseness; 47% from electrical causes, of which 10% originate from the rotor, such as casting defects resulting in unbalanced breath, and broken strips; 37% from the stator winding (see Figure 1, external factors

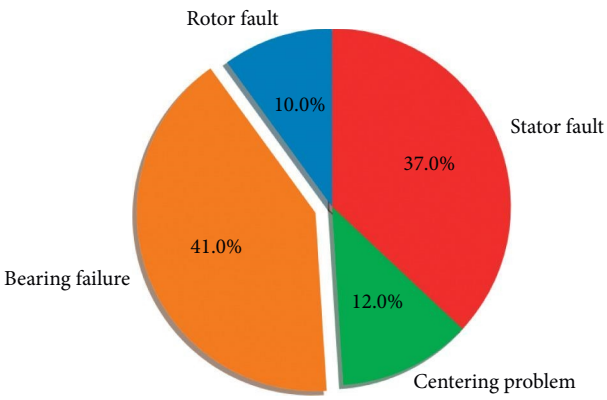


Figure 1: Motor failure source.

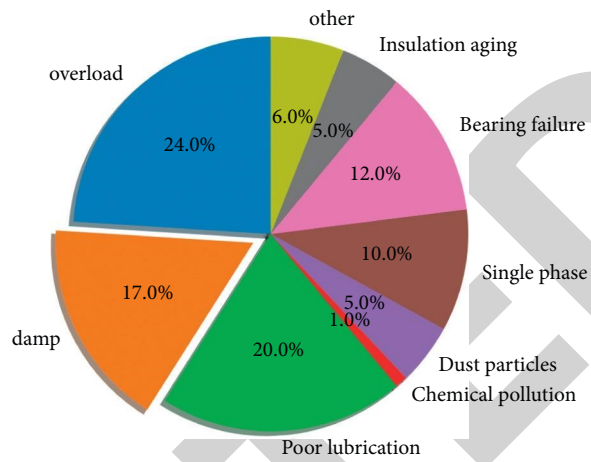


Figure 2: External factors causing failure.

causing the failure include overload, humidity, poor lubrication, and chemical pollution (see Figure 2.

The source of motor failure is mainly divided into four parts, including rotor fault, which accounts for 10%; stator fault, which accounts for 37%; bearing failure, which accounts for 41%; centering problem, which accounts for 12%.

In this paper, by constructing a fault reproduction test, the vibration time domain data of normal operating conditions, rotor eccentricity, stator short circuit, and bearing inner ring fault are extracted, and the data segment is divided into 15 speed cycles, each segment contains 13 typical time domain feature extraction and mathematical statistics for troubleshooting.

2. Related Work

2.1. Motor Research Review. A large number of studies have been conducted by domestic and foreign researchers on electric motors. The algorithm of RBF kernel function is chosen to diagnose the motor by characterization and parameter optimization in [4]. A parameter estimation algorithm is used to obtain four characteristic parameters of motor faults through the study in [5]. A method to achieve effective fault diagnosis in [6], for the problem that it is difficult to effectively measure and process the fault signal through the research equipment, the method of using wavelet changes to extract the fault feature signal is proposed, which can effectively characterize the real-time health status of the equipment and effectively realize the online health monitoring of the processing equipment.

According to the optimization process of maintenance strategy, motor fault repair can be divided into three types: ex post maintenance, preventive maintenance, and predictive maintenance [7]. Ex post facto maintenance means overhauling after a failure occurs. This strategy has the advantage of saving investment in condition monitoring and avoiding excessive maintenance and is generally used for a small range of noncore equipment [8]. The method has the disadvantage of not being able to predict accidental downtime, which raises maintenance costs in actual production, and the secondary damage to the equipment brings disastrous consequences to the enterprise, resulting in endless equipment disassembly and maintenance, which eventually leads to production losses and management loss [9–11]. Preventive maintenance is, according to a certain time for the interval, called regular maintenance. The strategy has good controllability, and the maintenance time is determined according to the need. In addition, it prevents accident surprises and reduces the possibility of catastrophic accidents, so the number of spare parts can be controlled, thus reducing cost capital [12]. However, failure-free equipment is also frequently overhauled, and the phenomenon of overmaintenance cannot be avoided, bringing sometimes even less benefit than the cumulative damage caused by maintenance, and the strategy does not completely eliminate unplanned downtime, and there is no targeted life analysis and optimization between different equipment. Predictive maintenance means that maintenance is based on equipment condition monitoring [13]. This strategy reduces unplanned equipment downtime by purchasing and using only the spare parts needed for the equipment when needed and repairing only when appropriate, thus achieving no failures without maintenance [14]. However, this method places high demands on monitoring instruments and system construction, and the high costs incurred for monitoring services as well as personnel inputs do not extend the life of the equipment.

About motor fault diagnosis principle and technology, mainly including, diagnosis principle: motor fault is generally based on vibration, temperature, noise, and other change conditions for diagnosis [15]. In the early stage of motor failure, engineers use sensors to conduct temperature tests on motor take points to initially confirm the type of fault generation, then determine the location of fault generation through noise, and analyze electrical signals through vibration testers to collect data from the motor to determine the location and cause of the fault. This diagnostic method is suitable for minor faults arising from the use of the motor, such as bearing, gear damage, loose structure, alignment problems, poor lubrication and other mechanical aspects of the problem [16]. For complex and serious motor failures, a spectrum analysis instrument can be used to regularly collect data from the load motor and perform time-domain frequency-domain

analysis based on current variations and the waveforms obtained. This method can analyze the deterioration trend of motor faults over time through the changing waveforms and current and other parameters [17]. The motor can be disassembled and examined using a high-voltage insulation tester, ohmmeter, and insulation resistance meter to determine the cause of the motor failure. In addition to the above methods, staff can apply the appropriate electrical inspection equipment to test and analyze the insulation structure of the motor, the life of the motor insulation structure, and motor performance factors, so as to further diagnose the motor failure [18]. The electrical circuit and the magnetic circuit form the main part of the motor, and the two work together to transform energy. Motor failure has both electrical causes and mechanical elements; therefore, factory quality, production process, incoming inspection, winding deterioration, bearing lubrication and alignment, unstable grid voltage, excessive load, and other conditions cannot make the motor in good running condition; thus, fatigue accumulation causes damage to the motor [19]. In production life, equipment personnel can choose the corresponding diagnostic methods for fault analysis and elimination based on the basic principles of motor failure.

Based on the above research, the authors of this paper propose to use the fusion of wavelet change and machine learning to carry out fault diagnosis research on electric motors. Through the design of fault reproduction experiments, several typical fault states of motors are restored, and their corresponding vibration signals are measured, and the purpose of diagnosing different states of motors is achieved through signal measurement and processing, combined with the high accuracy of machine learning methods in fault classification.

2.2. Introduction to CNN. In the 1960s, Hubel et al. introduced the concept of receptive fields through their studies of cat visual cortex cells, and in the 1980s, Fukushima proposed the concept of a neurocognitive machine based on the concept of receptive fields, which can be seen as the first implementation of a convolutional neural network; a neurocognitive machine that decomposes a visual pattern into many subpatterns (features) and then enters a hierarchical recursively connected feature planes that are processed, and it attempts to model the visual system, so that it can accomplish recognition even when objects are displaced or slightly deformed.

Convolutional neural network (CNN) is a variant of the multilayer perceptron (MLP). It was developed by biologists Huber and Wiesel in their early work on the visual cortex of the cat. A complex architecture exists for the cells of the visual cortex. These cells are sensitive to subregions of visual input space, which we call receptive fields, in such a way that they cover the entire visual field area in a flat manner. These cells can be divided into two basic types: simple cells and complex cells. Simple cells respond maximally to the pattern of edge stimuli from within the receptive field. Complex cells have a much larger receptive field, which is locally invariant to stimuli from an exact location.

In general, neurocognitive machines contain two types of neurons, namely, sampling elements, which are responsible for feature extraction, and convolutional elements, which are resistant to deformation, involving two important parameters, namely the receptive field, which determines the number of input connections, and the threshold parameter, which controls the degree of response to feature subpatterns. Convolutional neural networks can be seen as a generalised form of neurocognitive machines, which are a special case of convolutional neural networks.

In this application, its CNN training steps are as follows:

Step 1: Upload dataset

Step 2: Input layer

Step 3: Convolutional layer

Step 4: Aggregation layer

Step 5: Convolutional and pooling layers

Step 6: Dense layer

Step 7: Logit layer

2.3. Fault Characteristics of Motors. Vibration is a common mechanical fault. Motors, like other mechanical devices, produce a certain amount of vibration when they are in operation. For various models and specifications of electric motors, their vibrations have certain representative and permissible limits. In the event of a motor fault, the amplitude, vibration type, and frequency spectrum will change to a certain extent. Different faults vibrate at different frequencies, so that the working conditions of the motor can be objectively reflected. Through the analysis of the vibration signal, the working condition of the motor can be better understood, thus laying the foundation for the fault diagnosis of the motor. The vibration of an electric motor is caused by a variety of factors, and its occurrence location and characteristics are not the same. The various vibration characteristics of electric motors and the related influencing factors are analysed from both mechanical and electromagnetic aspects.

(1) Mechanical vibration: in the case of unbalanced rotor, abnormal rolling bearing, abnormal sliding bearing, and improper installation and commissioning, it can lead to mechanical vibration. Uneven distribution of the weight of the motor rotor can cause a shift in the center position, and the rotation will produce centrifugal force on one side, resulting in different support forces, thus making the motor not work properly. When the rotor is out of balance, its vibration frequency is equal to the rotor frequency and the amplitude increases as the rotor increases. The unbalance of the rotor is caused by the loss or displacement of rotor parts, displacement of rotor coils due to insulation shrinkage, loosening, unbalanced couplings, and the generation of dirt on the surface of the fan and rotor.

(2) Electromagnetic vibration: the motor uses the mutual coupling of magnetic fields in the air gap for the exchange of energy, that is, the use of magnetic fields to achieve the transformation of electromechanical energy. When the motor is in operation, the rotor rotates around the inner cavity of the stator and the magnetic field between the stator and rotor interacts to form a rotational force that acts on the stator seat, causing its periodic deformation and vibration. In the case of abnormal stator, abnormal rotor conductor, and uneven gap, the magnetic field of the motor will change, which will cause the vibration characteristics of the frame to change. When the stator is abnormal, its electromagnetic vibration is mainly twice the frequency of the power supply and disappears immediately after power failure.

3. Motor Fault Diagnosis Method

A fault diagnosis method using wavelet variation and machine learning is proposed for the problem such as difficult and complex feature extraction and difficult fault identification. And the bench experiment is carried out to extract the vibration time domain data of normal working condition, rotor eccentricity, stator short circuit, and bearing inner ring fault; calculate the variance contribution rate; and classify the faults by machine learning method with the obtained two-dimensional features.

3.1. Motor Fault Feature Extraction. Mechanical vibration exists during motor operation, so the motor with faults is selected for fault reproduction vibration signal measurement. In practice, motor faults are divided into 3 main categories: unbalanced rotor, bent rotor, and loose base, so experiments of motor faults and normal 4 motor states need to be constructed. In the actual operation of the motor, the sources of vibration are more numerous. And any mechanical equipment has inherent mechanical vibration, so it is difficult to effectively extract its corresponding fault feature signal.

3.2. Motor Simulation Test with Different Fault States. A three-phase motor of model HJNl 100L1-4 from Marathon was selected. Different fault reproduction tests were conducted on a university motor test bench. The signals of the four states of the motor were collected using an INV data acquisition system with a sampling frequency of $f(s) = 50Hz$ and an acquisition time of $t = 10\ s$, respectively.

The relevant devices and their performance parameters are as follows: motor HJNl 100L1.4, standard voltage 380 v, standard power 2.2 kw, motor speed l425 r/min, mass 35 kg; acceleration sensor CA-YD-186, frequency range 0.1–6 kHz; signal acquisition system INV303/306; signal adaptor YE3832 IEPE, the output amplitude is (2.5 ± 2.2)V.

The fault simulation test is conducted on the motor. The motors of four state categories were selected for the test. Test plan: the motor end cover and stator outside the installation of acceleration sensors, respectively. All use the contact installation method, using magnets to fix the acceleration sensor on its surface. The signals were acquired by replacing the motors in each of the four states on the fault test bench. A total of 160 $(4 \times 2 \times 20)$ samples were obtained through the INV data system. Thevibration time domaindata for a certain period of time of 1s for normal operating conditions, rotor eccentricity, stator short circuit, and bearing inner ring fault were first randomlyselected as shown in Figures 3–6.

Then, the normal motor and rotor unbalance time domain waveforms for a period of 1000s were randomly selected, as shown in Figure 7 and 8, respectively.

Comparing Figures 7 and 8, it can be seen that the rotor unbalance vibration is stronger than normal motor, but the difference is not big, so it is difficult to discriminate the fault through vibration signal directly. Therefore, the vibration signal needs to be processed, and the authors of this paper choose the wavelet variation method to process and analyze the signal.

3.3. Wavelet Variation Parameters Determination. Before using wavelet variation for signal processing, three parameters m, r, N need to be determined first, usually $m = 2$, $r = 0.1 \sim 0.25SD_x$, (where SD_x is the standard deviation of the original data $u(i)$), and N is 100 to 5000.

3.4. Wavelet Variation Feature Vector Extraction. Based on the fault reproduction test, $4 \times 2 \times 20$ vibration samples of the motor are collected, and let each vibration signal be $x(i)$, so there are $x(1), x(2) \ldots, x(160)$ in total 160 points. The sample values measured at 2 measurement points in a state i are shown in T_i, $x_{k,j}$ where j denotes the measurement point and k denotes the kth sample of measurement point j.

$$T_i = \begin{bmatrix} x_{11} & x_{12} \\ \vdots & \vdots \\ x_{20,1} & x_{20,2} \end{bmatrix}. \tag{1}$$

There is a 20×2 sample matrix corresponding to the ith state. The corresponding matrix can be obtained by determining each wavelet change value based on (1). By dividing each row, 20 eigenvectors can be obtained. The feature vectors represent the wavelet variation values of the vibration signals corresponding to the 2 measurement points in a certain state. Therefore, there are 20 wavelet variation eigenvectors in a certain state of the motor, i.e. $S_i = [E_{i,1}, E_{i,2}, \ldots, E_{i,20}]^T$, and there are 4 states in the motor simulation test. That is, there are 80 wavelet variation eigenvectors.

3.5. Wavelet Packet Noise Reduction Steps for Signals. In this paper, wavelet packet analysis is used in fault diagnosis to reduce noise. The specific procedures are as follows:

(1) The wavelet packet method is used to decompose the signal. Wavelets are selected and the desired level of decomposition is determined, which is then decomposed into wavelet packets.

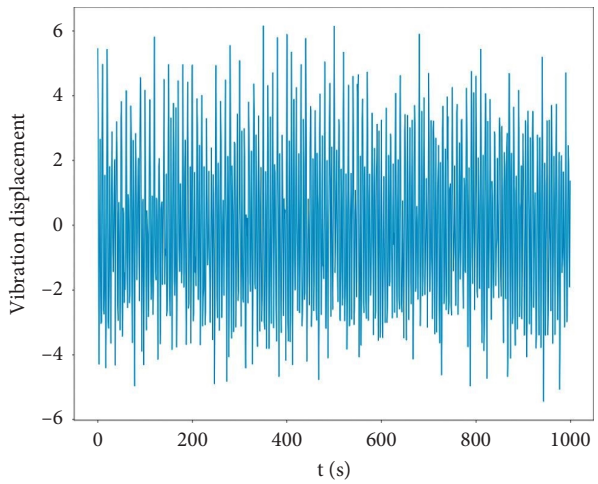


Figure 3: Normal working condition.

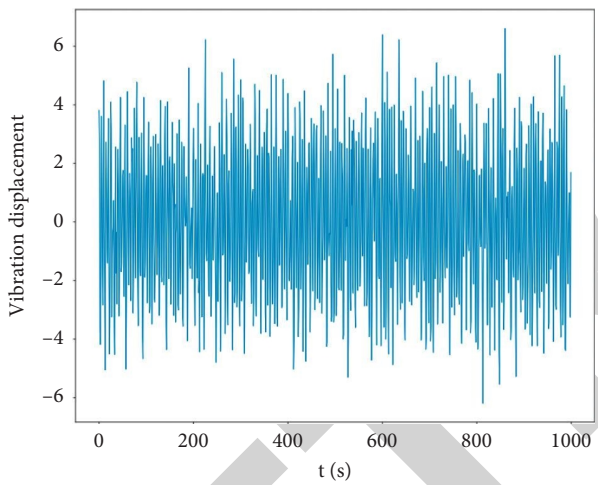


Figure 4: Rotor eccentricity.

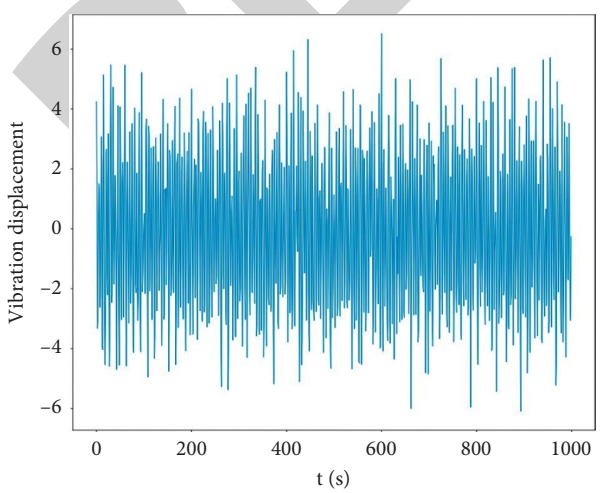


Figure 5: Stator short circuit.

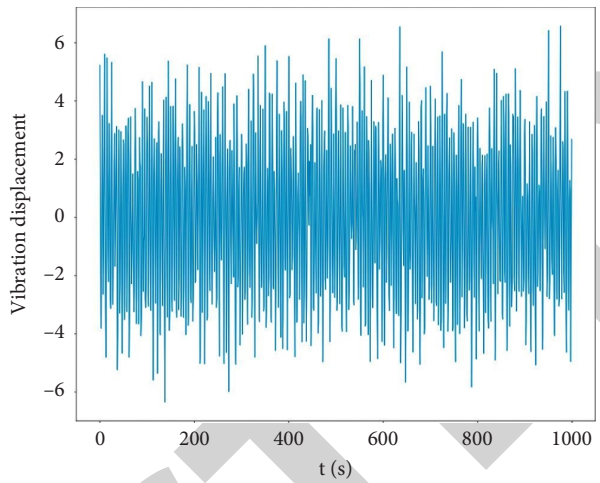


Figure 6: Bearing inner ring failure.

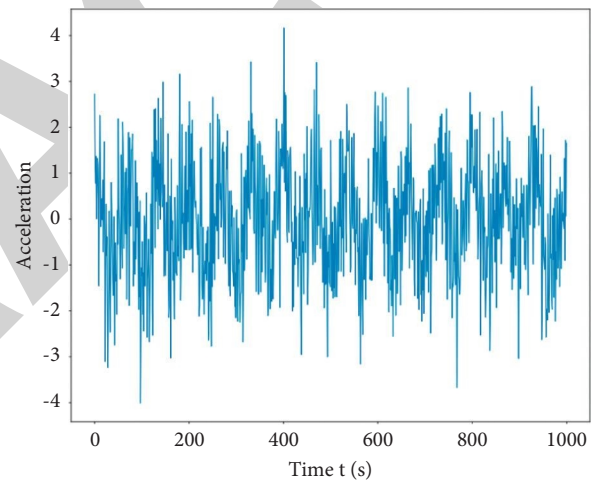


Figure 7: Normal motor time domain waveform.

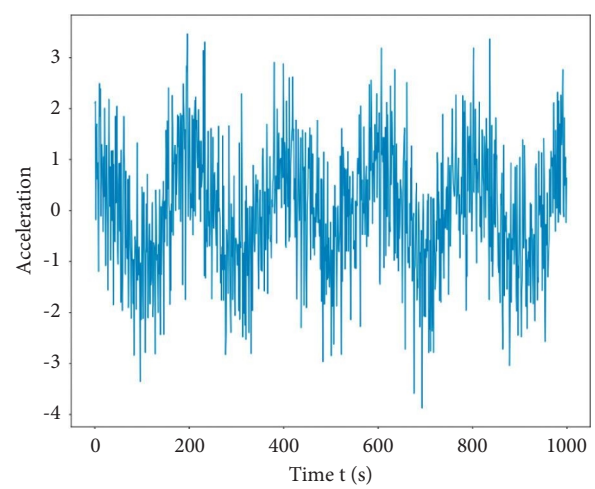


Figure 8: Time domain waveform of rotor unbalanced motor.

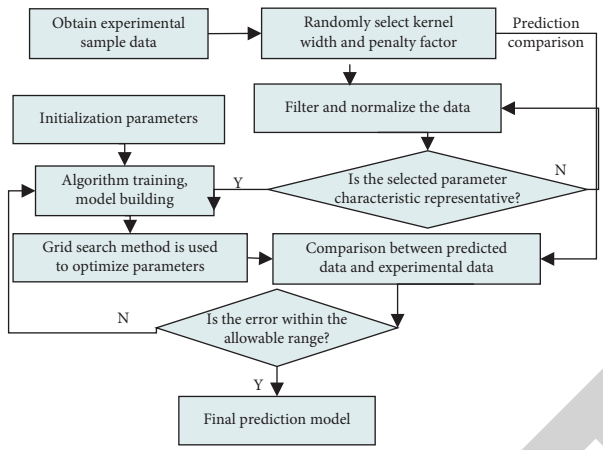


Figure 9: Flow of SVM to build a fault classifier.

(2) For a given entropy criterion, determine the optimal tree (this step is not essential and can be used selectively for different purposes).

(3) Threshold quantization of the decomposition coefficients of the wavelet packet and select a suitable threshold value for it.

(4) Perform wavelet packet reconstruction using the low-level decomposition coefficients of the wavelet packet and the quantized coefficients.

4. Case Study

Considering that the divided feature vector belongs to a small sample set and the fault type is clear, the SVM algorithm is selected to diagnose the fault.

4.1. Support Vector Machine (SVM) Algorithm. The motor fault classification problem is a linear regression problem. Calculated by the following equation:

$$f(x) = \omega^T x + b, \tag{2}$$

where x is the input vector, ω is the weight coefficient, and b is the bias.

The fault classification diagnostic function of the support vector machine, which can be expressed in the following equation:

$$\psi(\omega, \xi, \xi_i^*) = \frac{1}{2}\|\omega\|^2 + C\sum_{i=1}^{1}(\xi_i + \xi_i^*), \tag{3}$$

where ω is the weight coefficient, C is the penalty factor, and ξ, ξ_i^* are the upper and lower limits of the slack variables, respectively.

The constraint conditions are given by

$$\begin{aligned} &\text{s.t.}\,[(x_i \cdot \omega) + b] - y_i \le \xi_i + \varepsilon, \\ &i = 1, 2, \ldots, l, \end{aligned} \tag{4}$$

where: ξ_i is the relaxation variable, ω is the allowable error, y_i is the output corresponding to the ith sample [21–25].

Introducing Lagrange function

$$\sum_{i=1}^{i}(\alpha_i - \alpha_i^*) = 0. \tag{5}$$

Using the pairwise principle, Lagrange multiplier method, the pairwise form of the optimization problem can be obtained by substituting (6) into (3) as follows:

$$\begin{aligned} \max_{\alpha,\alpha^*}^{w} = &-\frac{1}{2}\sum_{i=1}\sum_{j=1}(\alpha_i - \alpha_i^*)(\alpha_j - \alpha_j^*)K(x_i, x_j) \\ &+ \sum[\alpha_i(y_i - \varepsilon) - \alpha_i^*(y_i + \varepsilon)]. \end{aligned} \tag{6}$$

That is, the output of the support vector machine is as follows:

$$f(x) = \sum_{i=1}(a_i, a_i^*)K(x_i, x) + b, \tag{7}$$

where a_i, a_i^* is the Lagrange multiplier; b is the bias.

4.2. Fault Diagnosis. The process of building a multifault classifier using support vector mechanism is shown in Figure 9. The model is as follows:

$$f_F(x) = \text{sign}\left[\sum_n y_i \alpha^* K(x_i \cdot x_j) + b^*\right], \tag{8}$$

where $K(x_i, x_j)$ is the kernel function, and the radial basis kernel function is chosen with the expression.

$$K(x, y) = \exp\left(-\frac{\|x - y\|^2}{2\sigma^2}\right), \tag{9}$$

where σ is the parameter that controls the width of the kernel function.

The SVM is used to diagnose 4 types of faults in motors; 20 groups of each of the 4 states, i.e., a total of 80

Table 1: Wavelet variation values for each state category of the motor.

Category	Group 1		Group 2		Group 3		Group 4	
	Measurement point 1	Measurement point 2	Measurement point 1	Measurement point 2	Measurement point 1	Measurement point 2	Measurement point 1	Measurement point 2
1	0.551	0.521	0.698	0.961	0.691	0.887	0.556	0.507
2	0.785	0.597	0.731	0.525	0.778	0.546	0.811	0.621
3	1.078	1.017	0.856	0.965	0.871	0.712	1.126	0.987
4	0.997	0.879	0.856	0.941	0.871	0.756	0.883	0.958

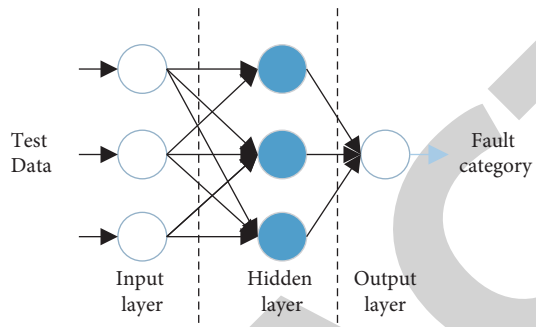


Figure 10: Improved BP neural network classification model.

Table 2: Mixing matrix of classifier output.

Prediction	Normal	Rotor eccentricity	Stator short circuit	Bearing inner ring
Normal	28	0	0	0
Rotor eccentricity	0	19	0	0
Stator short circuit	1	0	25	0
Bearing inner ring	0	0	0	25

groups of feature vectors. Then, the test samples are used for validation. For each state, 10 sets of test data are selected and tested four times. The wavelet variation values for each state category of the motor are shown in Table 1.

MATLAB 7.0 was chosen to set up the corresponding SVM program for diagnosis, and the penalty parameters of the model were C_i, kernel width σ_i ($i = 1, 2, \cdots, 9$). The highest training accuracy was achieved using training samples in MATLAB at $C_i = 100$, $\sigma_i = 1$. The data in Table 1 were validated and the classification results show that the SVM classification model was used to diagnose the test samples with an accuracy of 97.5%, and only one rotor bend was misclassified as a loose base fault [26].

4.3. Comparison of Different Fault Diagnosis Methods. The commonly used improved BP neural network diagnosis method is selected to train the sample set of wavelet variation feature quantity of the motor. The validation samples are used to verify and compare the diagnosis results with SVM. The commonly used 3-layer BP neural network model is selected to construct the motor fault classification prediction model as shown in Figure 10.

The validation samples are input into the BP classification model for diagnosis, and the classification results show that a normal motor condition is mistaken for a rotor unbalance fault. A rotor bending fault is mistaken for a loose base fault. A loose base fault was mistaken for a rotor bending fault, and a total of three state categories were misclassified, with an accuracy rate of 92.5%.

The results of diagnostic classification for different types of vibrations for normal operating conditions, rotor eccentricity, stator short circuit, and bearing inner ring faults are shown in Table 2. The mixture matrix output from the table shows that the faults are classified more accurately, and the diagnostic accuracy in this case is as high as 98.97%.

5. Conclusion

The motor vibration signal feature extraction is difficult and complex, and fault identification is difficult. The motor experiment platform is used to extract the motor normal and fault conditions data, and a motor fault diagnosis model based on wavelet transform and machine learning is proposed. The method is applied to the diagnosis and maintenance of motor faults. Based on failure analysis, reliability maintenance is performed to avoid overmaintenance and

undermaintenance of equipment, early detection of equipment failures, and better control of spare parts.

Wavelet analysis can decompose the signal into wavelet expanded fundamental functions, so that the signal can be locally refined in the high- and low-frequency parts, while maintaining the time-frequency characteristics of the signal, thus making the signal have better time-frequency characteristics, which can effectively identify the signal and thus achieve fault diagnosis of the signal. In this paper, wavelet analysis techniques are used for fault simulation of electric motors. Wavelet analysis is a new diagnostic technique.

There are two methods of combining wavelet analysis and neural networks.

(1) Loosely combined: wavelet analysis is a preprocessing method for traditional neural networks, which provides feature vectors for neural networks and trains and diagnoses them. The two are closely but independently related.

(2) Closely integrated: feedforward networks based on wavelet analysis. The basic idea is to use wavelet elements to substitute for neurons, i.e., the activation function is the wavelet function base, and the scale and translation parameters of the wavelet function replace the corresponding input layer to the hidden layer weight and hidden layer threshold [20].

Data Availability

The dataset used in this study can be obtained from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest regarding this work.

References

[1] M. Zhao, M. Kang, B. Tang, and M. Pecht, "Deep residual networks with dynamically weighted wavelet coefficients for fault diagnosis of planetary gearboxes," *IEEE Transactions on Industrial Electronics*, vol. 65, no. 5, pp. 4290–4300, 2018.

[2] B. Akhil Vinayak, K. Anjali Anand, and G. Jagadanand, "Wavelet-based real-time stator fault detection of inverter-fed induction motor," *IET Electric Power Applications*, vol. 14, no. 1, pp. 82–90, 2020.

[3] T. M. Alamelu Manghai and R. Jegadeeshwaran, "Vibration based brake health monitoring using wavelet features: a machine learning approach," *Journal of Vibration and Control*, vol. 25, no. 18, pp. 2534–2550, 2019.

[4] M. Ali, M. Shabbir, X. Liang, Y. Zhang, and T. Hu, "Machine learning-based fault diagnosis for single-and multi-faults in induction motors using measured stator currents and vibration signals," *IEEE Transactions on Industry Applications*, vol. 55, no. 3, pp. 2378–2391, 2019.

[5] J. Seshadrinath, B. Singh, and B. K. Panigrahi, "Incipient turn fault detection and condition monitoring of induction machine using analytical wavelet transform," *IEEE Transactions on Industry Applications*, vol. 50, no. 3, pp. 2235–2242, 2014.

[6] J. Cusido, L. Romeral, J. A. Ortega, A. Garcia, and J. R. Riba, "Wavelet and PDD as fault detection techniques," *Electric Power Systems Research*, vol. 80, no. 8, pp. 915–924, 2010.

[7] Y. Han, B. Tang, and L. Deng, "Multi-level wavelet packet fusion in dynamic ensemble convolutional neural network for fault diagnosis," *Measurement*, vol. 127, pp. 246–255, 2018.

[8] R. Chen, X. Huang, L. Yang, X. Xu, X. Zhang, and Y. Zhang, "Intelligent fault diagnosis method of planetary gearboxes based on convolution neural network and discrete wavelet transform," *Computers in Industry*, vol. 106, pp. 48–59, 2019.

[9] Y. Li, "Exploring real-time fault detection of high-speed train traction motor based on machine learning and wavelet analysis," *Neural Computing & Applications*, vol. 34, no. 12, pp. 9301–9314, 2021.

[10] R. Yan, R. X. Gao, and X. Chen, "Wavelets for fault diagnosis of rotary machines: a review with applications," *Signal Processing*, vol. 96, pp. 1–15, 2014.

[11] A. Ziaja, I. Antoniadou, T. Barszcz, W. J. Staszewski, and K. Worden, "Fault detection in rolling element bearings using wavelet-based variance analysis and novelty detection," *Journal of Vibration and Control*, vol. 22, no. 2, pp. 396–411, 2016.

[12] W. Deng, S. Zhang, H. Zhao, and X. Yang, "A novel fault diagnosis method based on integrating empirical wavelet transform and fuzzy entropy for motor bearing," *IEEE Access*, vol. 6, pp. 35042–35056, 2018.

[13] Z. Fu, X. Liu, and J. Liu, "Research on the fault diagnosis of dual-redundancy BLDC motor," *Energy Reports*, vol. 7, pp. 17–22, 2021.

[14] K. Chandra, A. S. Marcano, S. Mumtaz, R. V. Prasad, and H. L. Christiansen, "Unveiling capacity gains in ultradense networks: using mm-wave NOMA," *IEEE Vehicular Technology Magazine*, vol. 13, no. 2, pp. 75–83, 2018.

[15] J. Du, C. Jiang, Z. Han, H. Zhang, S. Mumtaz, and Y. Ren, "Contract mechanism and performance analysis for data transaction in mobile social networks," *IEEE Transactions on Network Science and Engineering*, vol. 6, no. 2, pp. 103–115, 1 April-June 2019.

[16] S. Alshareef, S. Talwar, and W. G. Morsi, "A new approach based on wavelet design and machine learning for islanding detection of distributed generation," *IEEE Transactions on Smart Grid*, vol. 5, no. 4, pp. 1575–1583, 2014.

[17] Y. Zou, Y. Zhang, and H. Mao, "Fault diagnosis on the bearing of traction motor in high-speed trains based on deep learning," *Alexandria Engineering Journal*, vol. 60, no. 1, pp. 1209–1219, 2021.

[18] Z. Zhang, Di Wu, and C. Zhang, "Study of cellular traffic prediction based on multi-channel sparse LSTM," *Computer Science*, vol. 48, no. 6, pp. 296–300, 2021.

[19] A. Revathy, C. S. Boopathi, O. I. Khalaf, and C. A. T. Romero, "Investigation of AlGaN channel HEMTs on β-Ga2O3 substrate for high-power electronics," *Electronics*, vol. 11, no. 2, p. 225, 2022.

[20] A. A. El-Saleh, A. Alhammadi, I. Shayea et al., "Measuring and assessing performance of mobile broadband networks and future 5G trends," *Sustainability*, vol. 14, no. 2, p. 829, 2022.

[21] N. Subramani, P. Mohan, Y. Alotaibi, S. Alghamdi, and O. I. Khalaf, "An efficient metaheuristic-based clustering with routing protocol for underwater wireless sensor networks," *Sensors*, vol. 22, no. 2, p. 415, 2022.

[22] U. Abubakar, S. Mekhilef, K. S. Gaeid, H. Mokhlis, and Y. I. Al Mashhadany, "Induction motor fault detection based on multi-sensory control and wavelet analysis," *IET Electric Power Applications*, vol. 14, no. 11, pp. 2051–2061, 2020.

Hindawi
Computational Intelligence and Neuroscience
Volume 2022, Article ID 9592200, 9 pages
https://doi.org/10.1155/2022/9592200

Research Article

A Computational Neural Network Model for College English Grammar Correction

Xingjie Wu

School of General Caliber-oriented Education, Wuchang University of Technology, Wuhan 430000, China

Correspondence should be addressed to Xingjie Wu; wuxingjie@wut.edu.cn

Received 4 July 2022; Revised 15 August 2022; Accepted 17 August 2022; Published 5 September 2022

Academic Editor: Ning Cao

For the error correction of English grammar, if there are errors in the semantic units (words and sentences), it will inevitably affect the subsequent text analysis and semantic understanding, and ultimately reduce the overall performance of the practical application system. Therefore, intelligent error detection and correction of the word and grammatical errors in English texts is one of the key and difficult points of natural language processing. This exploration innovatively combines a computational neural model with college grammar error correction to improve the accuracy of college grammar error correction. It studies the computational neural model in English grammar error correction based on a neural network named Knowledge and Neural machine translation powered College English Grammar Typo Correction (KNGTC). First, the Recurrent Neural Network is introduced, and the overall structure of the English grammatical error correction neural model is constructed. Moreover, the supervised training of Attention is discussed, and the experimental environment and experimental data are given. The results show that KNGTC has high accuracy in college English grammar correction, and the accuracy of this model in CET-4 and CET-6 writing can reach 82.69%. The English grammar error correction model based on the computational neural network has perfect function and strong error correction ability. The optimization and perfection of the model can improve students' English grammar level, which has certain practical value. After years of continuous optimization and improvement, English grammar error correction technology has entered a performance bottleneck. This mode's construction can break the current technology's limitations and bring a better user experience. Therefore, it is very valuable to study the error correction model of English grammar in practical application.

1. Introduction

Language processing technology is the product of the development and evolution of computer technology. It can enable computers to correctly understand and use natural language. It is a theoretical basis and method to realize the efficient communication between humans and computers. Specifically, to communicate in natural language is to enable computers to understand the ideas and meanings people convey in natural language, and to convey certain intentions and strategies in the written form of natural language. The first is to understand natural language, and the second is to generate natural language [1]. English is the most widely used international language, and writing is a crucial indicator of English level. Automatic check and correction of spelling errors of words and grammatical errors of sentences through computer natural language processing technology will be very conducive to improving students' English writing level [2]. Moreover, it also greatly saves the time and energy for teachers to review students' compositions, so that teachers can focus on the composition's overall structure and the narration content [3]. With the progress of automatic scoring technology for English essays, spelling and grammatical errors will become indispensable evaluation indicators, improving the objectivity and accuracy of automatic marking and scoring [4].

Experts and scholars have extensively researched grammatical errors in English sentences. Boyd's research showed that the ICICLE system detected grammatical abnormalities in sentences with the assistance of the constructed grammatical error recognition rules, and gave relevant prompt information [4]. Coley et al. checked

grammatical errors through the N-ary grammar model, established a list of the candidate recommended words, and prioritized the candidate words according to the random context-free grammar [5]. Fetaya et al. focused on the preposition errors. With the British National Corpus (BNC) corpus as the training set, they extracted the context of prepositions, then established feature vectors and applied them to the generated maximum entropy model for preposition error detection. Among the contextual features of prepositions, the contextual word features in the extraction window contribute more than collocation features and named entities [6]. Domestic Gan et al. proposed a method based on examples and the introduction of negative rules to check the syntax. It is found that among the sentences with grammatical errors, 55% belong to local grammatical errors and 18% belong to global grammatical errors [7]. To sum up, in the extraction of context features, different features have different contributions to grammar checking. If the allocation of features is unreasonable, the discrimination of features will decline and ultimately affect the comprehensive performance of the system, so it is crucial to build a reasonable feature validity allocation function.

First, the research background of college English grammatical error correction is analyzed. It is found that college English grammar error correction has the characteristics of limited information, various error types and relatively single alignment. According to its characteristics, a spelling error correction model based on a neural network is generated based on the Sq2seg network. The transition probability between keys is obtained through the analysis and processing of user data, and the standard Attention matrix is generated according to the transition probability of keys. The supervised training of the neural network Attention mechanism is realized by adding regular terms to the neural network loss function. Finally, the above model is verified by experiments.

The English grammar correction model established adopts the idea of using the computer to replace the manual operation, so that teachers can get rid of the heavy task of English homework correction. In addition, from the students' point of view, these correction suggestions can also help students improve their English writing ability without teachers' guidance.

2. Materials and Methods

2.1. Recurrent Neural Network (RNN). The neural network is like a simulation of the brain in processing information. Just like when people listen to a song or read a text, they must read information from front to back [8]. In traditional neural networks, each input is independent. However, in speech and natural language, the subsequent input is often affected by the previous input, which is called sequence information [9]. For example, when the next word in a sentence is predicted, the model must calculate according to the previously input word. RNN is a neural network for processing timing information, which is specially used to process timing information [10]. One of the main features of RNN is parameter sharing, which means there are many basic units in the RNN. These basic units have the same parameters, complete the same work, and transmit the information received by the model on different bases by transmitting the state of the hidden layer [11]. Therefore, at any time point, the model can obtain the input information of the current time, and take the previous input information into account. The characteristics of RNN make it especially suitable for processing natural language data [12].

The RNN basic unit consists of three main modules [13]. Figure 1 shows the basic structure:

(1) Input layer $X = \{x_0, x_1, \ldots, x_{t-1}, x_t\}$

(2) Output layer $Y = \{y_0, y_1, \ldots, y_{t-1}, y_t\}$

(3) Hidden layer $H = \{H_0, H_1, \ldots, H_{t-1}, H_t\}$

The RNN folded structure diagram in Figure 1 is expanded into the structure diagram in Figure 2:

Figure 2 shows that due to the structural characteristics of RNN, the parameters are mainly divided into three parts [14–16]:

(1) Connection weight W_{xh} from the input layer to the hidden layer

(2) Connection weight W_{hh} between hidden layers

(3) Connection weight W_{hy} from the hidden layer to the output layer

The main purpose of the RNN model structure is to process and predict sequence data. In the fully connected neural network or convolutional neural network model introduced before, the network structure is from the input layer to the hidden layer and then to the output layer. The layers are fully connected or partially connected, but the nodes between each layer are not connected. RNN network has been proved to perform well in multiple natural language processing tasks. Due to the parameter sharing mechanism, the RNN network can greatly reduce the number of parameters in the model [17, 18]. It suggests that the parameters W_{xh}, W_{hh} and W_{hy} are the same for each time step.

RNN also has some limitations. In the case of a deep network, due to the chain derivation rule, the updated value of parameters is often the result of the objective function multiplied by the gradient of several activation functions [19]. Figure 3 shows the commonly used activation functions in RNN networks:

Figure 3 suggests that the gradient of the activation function commonly used in neural networks is usually a value less than 1. The updated value of the parameters in front of the model will be multiplied by several activation function gradients, which will cause the updated value of the parameters to become smaller and smaller until it becomes 0, so that the parameters in front of the model cannot be effectively updated. This problem is called the gradient vanishing problem, also known as the long-distance dependence problem [20]. The gradient vanishing problem leads to even if the time series processed by the RNN network can be of any length in theory, with the deepening of the model, the parameter update rate of the former hidden

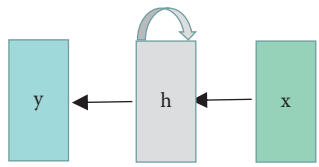


Figure 1: Folded view of RNN model structure.

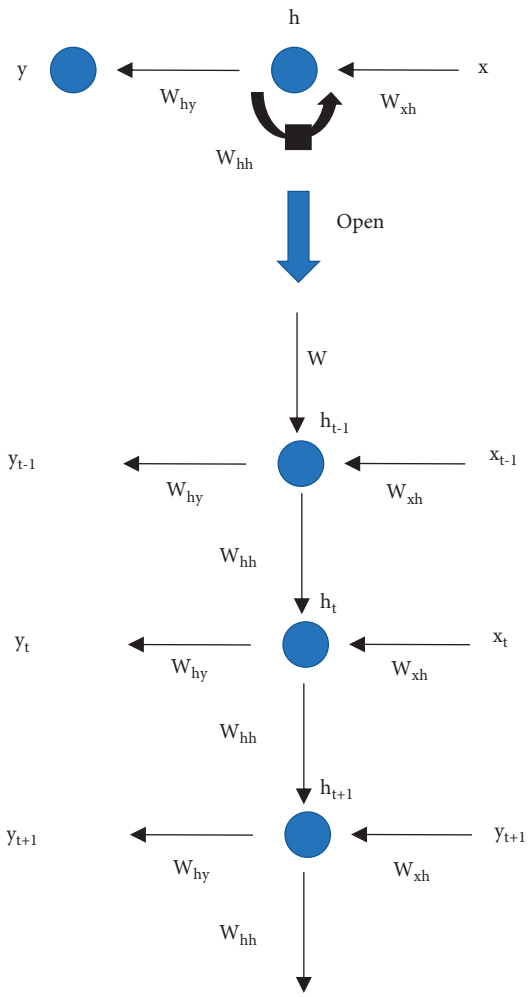


Figure 2: Expanded view of RNN model structure.

layer of the model is far lower than that of the latter hidden layer of the model, and the former input of the model is difficult to affect the latter input [21].

2.2. The Overall Framework of the Computational Neural Model for College English Grammar Correction. Currently, the key problems of college English grammar error correction are mainly as follows. First, the abbreviations and irregular word orders of entities (names of people and places) in English texts will affect the accuracy of computer clauses. Second, in the context of feature extraction, different features have different contributions to grammar checking. If the allocation of features is unreasonable, the discrimination of features will decline and ultimately affect the system's overall performance. Third, the

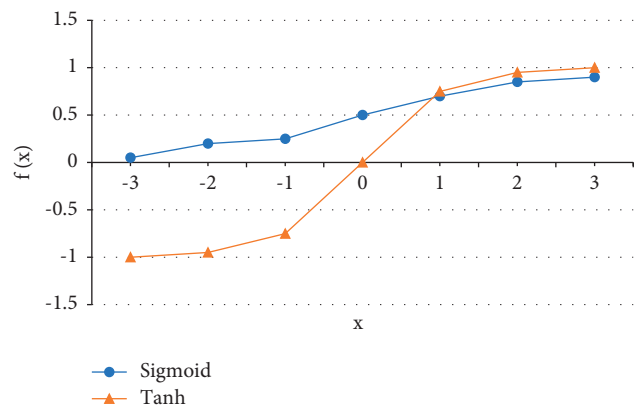


Figure 3: Common activation functions in RNN.

current research on grammar checking mainly focuses on the analysis of some types of grammatical errors, while the adaptability to other types of errors is not good enough. A Knowledge and Neural machine translation powered College English grammar Typo Correction (KNGTC) model is proposed for the problems in the existing methods. Neural Machine Translation (NMT) refers to a machine translation method that directly uses neural networks to carry out translation modeling in an end-to-end manner [22]. The error correction model based on NMT adopts a simple and intuitive method to complete the error correction, which has the following advantages. First, the process of input segmentation and processing one by one is omitted to avoid the uncertainty caused by segmentation errors. Second, it can effectively grasp the global information of user input, better fit the distribution of user input habits through the neural network, and have stronger adaptability and error correction ability.

College English grammar correction is regarded as a special translation task. It is determined that the input granularity of the model is character level and the output is statement level. Figure 4 shows the main model structure of KNGTC:

In Figure 4, $X = (x_1, x_2, \ldots, x_n)$ is adopted to express user input, where x_i represents a letter in input. $Y = (y_1, y_2, \ldots, y_{Ts})$ represents the result, and y_i represents a word. Error correction is conducted through the improved neural machine translation model. Besides, KNGTC can effectively improve error correction accuracy by combining the user's vector expression and the transition probability of adjacent keys.

KNGTC model is based on the RNN+ Attention structure in neural network machine translation. The Attention mechanism is to give a set of vector set values and a vector query. It is a mechanism that calculates the weighted sum of values based on the query. The focus of Attention is the calculation method of the "weight" of each value in the set values. Sometimes, this Attention mechanism is called query output, which focuses on (or takes into account) different parts of the original text. The RNN architecture is called the Seq2seq model. The Encoder part inputs a sequence $X = (x_1, \ldots, x_{TL})$ as the source language, and the coding network generates a series of hidden layer states $H=$

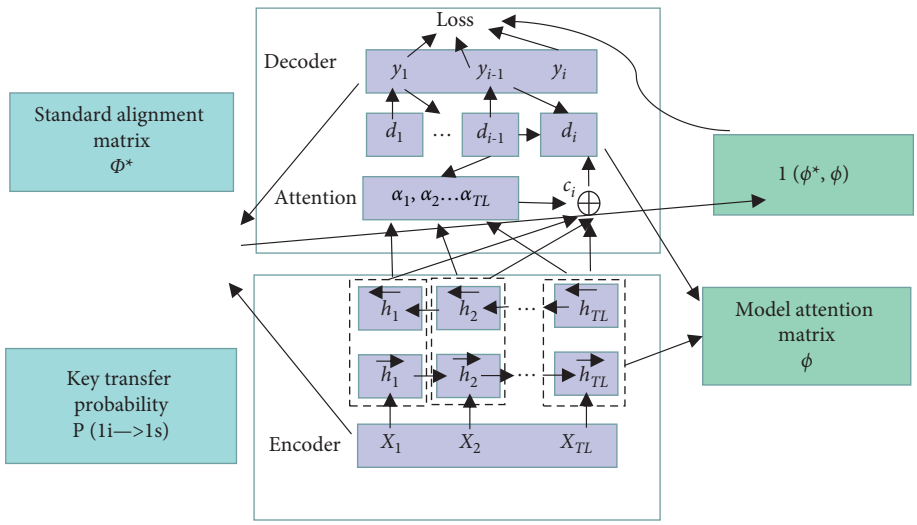


Figure 4: Overall network architecture of KNGTC.

($h_1, \ldots, h_{TL}$) as the vector expression of the input through a bidirectional RNN. The hidden layer h_t is the result of splicing the hidden layer results of the positive RNN and the hidden layer results of the negative RNN at time t. The expression of H is as follows [23]:

$$h_1 = \left[\overrightarrow{h_t}^T ; \overleftarrow{h_t}^T\right]^T, \tag{1}$$

$\overrightarrow{h_t}$ and $\overleftarrow{h_t}$ are calculated by the following equation [24]:

$$\begin{aligned} \overrightarrow{h_t} &= GRU_{\overrightarrow{enc}}\left(\overrightarrow{h_{t-1}}, x_t\right), \\ \overleftarrow{h_t} &= GRU_{\overleftarrow{enc}}\left(\overrightarrow{h_{t+1}}, x_t, x_t\right), \end{aligned} \tag{2}$$

where $GRU_{\overrightarrow{enc}}$ and $GRU_{\overleftarrow{enc}}$ represent the forward and reverse gated recurrent unit (GRU), and different parameters are used respectively.

In the Decoder phase, the model still uses GRU units. The hidden layer state in the decoding network is called $d = \{d_1, d_2, \ldots, d_{Ty}\}$, and the probability distribution information of the current output sequence $y_1, y_2, \ldots, y_{Ty}$ is as follows [25]:

$$p(y_i | y_1, \ldots, y_{i-1}, X) = g(y_{i-1}, d_i, c_i), \tag{3}$$

g is a function of the nonlinear multi-layer structure. The user calculates the probability distribution information of output y_i. d_i is the hidden layer state of RNN at the time i, and its equation is as follows:

$$d_i = GRU_{dec}(d_{i-1}, y_{i-1}, c_i). \tag{4}$$

The context information c_i used to predict y_i is equal to $\sum_{k=1}^{T_x} \exp\ (q_{ij})$. The weight a_{ij} of the hidden layer state h_i of the coding network is calculated by the following equation [26]:

$$a_{ij} = \frac{\exp\ (q_{ij})}{\sum_{k=1}^{T_x} \exp\ (q_{ij})}, \tag{5}$$

where $q_{ij} = a\ (d_{i-1}, h_j)$, a is a similarity calculation equation with back-propagation property, which updates the parameters together with the global network. The size of training data C is $|C|$ and each training data is composed of the form of (x, y). The loss function of the model is as follows:

$$Loss' = -\sum_{\langle x,y \rangle \in C} \sum_{l=1}^{T_y} logp(y_l | y_{<l}, x). \tag{6}$$

2.3. Supervised Training of Attention Mechanism. The supervised training of the Attention mechanism is a machine learning task of inferring functions from labeled training datasets. The training data consists of a set of training examples. In supervised learning, each example is a pair consisting of an input object (usually a vector) and the desired output value (also known as a supervised signal). A supervised training algorithm analyzes training data and generates an inference function, which can be used to map new examples. The purpose of supervised training of the Attention mechanism is to carry out supervised training of the Attention mechanism in combination with key transfer probability and input-output alignment information. This method can effectively correct click errors and the simplified spelling in the Pinyin input method.

Attention weights $\alpha_1, \alpha_2, \ldots$, and α_{TL} play an important role in predicting the next output in decoding networks. However, in the traditional Attention mechanism, only the information of the source language itself is considered, and other information is not effectively used. In the English grammar error correction task, a multi-level Attention mechanism is introduced to improve the error correction rate of the model at the word level and character level. In this project, alignment information and adjacent key transfer probability greatly impact spelling error correction tasks. Therefore, this section describes how to introduce alignment information and adjacent key transfer probability into the Attention mechanism as external information. A new

alignment model is proposed to construct a binary matrix representing the input-output alignment relationship, and it is initialized with "0." Then, the key transfer probability is adopted to fill in the aligned position, and the supervised training of the Attention mechanism is realized through this matrix [27].

For each training data, the standard alignment matrix automatically generated by the model is defined as middle ϕ^*, as shown in:

$$\begin{array}{c} \\ l_1 \\ l_2 \\ l_3 \\ l_4 \\ l_5 \\ eos \end{array} \begin{array}{c} \begin{array}{cccc} s_1 & s_2 & s_3 & eos \end{array} \\ \begin{bmatrix} 0.93 & 0 & 0 & 0 \\ 0.57 & 0 & 0 & 0 \\ 0 & 0.97 & 0 & 0 \\ 0 & 0.49 & 0 & 0 \\ 0 & 0 & 0.86 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \end{array}, \quad \begin{array}{c} \\ l_1 \\ l_2 \\ l_3 \\ l_4 \\ l_5 \\ eos \end{array} \begin{array}{c} \begin{array}{cccc} s_1 & s_2 & s_3 & eos \end{array} \\ \begin{bmatrix} 0.62 & 0 & 0 & 0 \\ 0.38 & 0 & 0 & 0 \\ 0 & 0.67 & 0 & 0 \\ 0 & 0.33 & 0 & 0 \\ 0 & 0 & 0.1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \end{array}. \tag{7}$$

The input of φ^* is $(l_1, \ldots, l_{\mathrm{TL}})$, the corresponding number of lines is $T_{\mathrm{L}} + l$ (representing T_{L} inputs and one <eos>), and the output is $(S_1, \ldots, S_{\mathrm{TS}})$. The corresponding number of columns is $T_{\mathrm{s}} + l$ (T_{s} outputs plus one <eos>). For $l < i < T_{\mathrm{L}}$, $l < j < T_{\mathrm{s}}$, $\varnothing^*_{ij}$ is as follows:

$$\varnothing^*_{ij} = \max_{1 \le k \le |s_j|} p_t\left(S_{jk} \longrightarrow l_i\right) A_{i,j}, \tag{8}$$

where A_{ij} indicates whether the character l_i is part of the spelling sentence S_j. If yes, it is l. Otherwise, it is 0. S_{jk} represents the k-th character in the word j. For $i = T_{\mathrm{L}} + l$ or $j = T_{\mathrm{S}} + l$:

$$\varnothing_{i,j} = \begin{cases} 1, & i = T_L + 1, j = T_S + 1, \\ 0, & \text{otherwise.} \end{cases} \tag{9}$$

First, the input sequence L and the corresponding word sequence S are obtained. Recursively, it is essential to traverse all possible separated results of L. Then, each result is scored. The scoring basis is to separate each corresponding part of the result and s, take the editing distance and sum, and take the result of the minimum score as the most reasonable segmentation result of the input string L. For example, if a division result of L is $\mathrm{Seg} = \{seg_1, seg_2, \ldots, seg_{T_s}\}$, and the CALCSCORE calculation equation in line 29 is [28]:

$$\tau = -\sum_{n=1}^{T_S} EditDistance\left(seg_n, s_n\right), \tag{10}$$

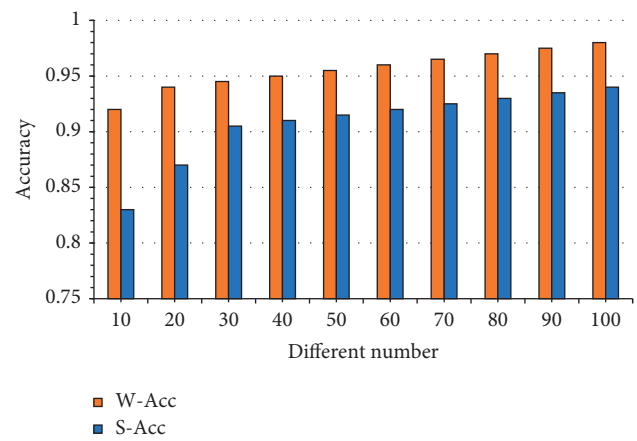


Figure 5: Accuracy of the model under the condition of taking the optimal results of different numbers.

where Edit Distance represents the editing distance, and refers to the minimum number of editing operations required to convert one string to the other between two strings. The greater their distance is, the more different they are. Editing operations include replacing one character with another, inserting characters, and deleting a character.

The traditional Attention matrix calculates the default alignment matrix ϕ^* based on the hidden layer of the coding network. The matrix distance between ϕ^* and ϕ' is calculated by the following equation [29]:

$$1\left(\varnothing^*, \varnothing\right) = \left\|\varnothing^* - \varnothing\right\|_2^2. \tag{11}$$

A new loss function is obtained by combining it with the original loss function:

$$Loss = -\sum_{\langle x,y \rangle > \in \mathbb{C}} \left\{ \sum_{l=1}^{T_y} \log p\left(y_l | y_{<l}, x\right) + 1\left(\varnothing^*, \varnothing\right) \right\}. \tag{12}$$

It reveals that the new loss function consists of two parts. The first part measures the accuracy of the error correction results, and the second part measures the accuracy of the alignment model.

2.4. Experimental Environment. The uniform distribution initializes the parameters in the experiment. Both the encoding and decoding networks use a 128 dimension hidden layer, and the dimension of the input layer is 256. The batch size is 64 samples, and the parameter optimization algorithm is Adam algorithm. The dropout method is used to prevent overfitting, and the dropout ratio is 0.2. KNGTC model has been trained on a Tesla K40 GPU for 48 hours, and has reached the convergence state after 350000 recursions. The decoding phase uses the Beam Search algorithm. Figure 5 shows the model's accuracy under different decoding widths in the Beam Search algorithm.

Figure 5 shows that the model accuracy increases with the increase of decoding width. However, due to the response time requirement of the input method, too many

candidate results will bring a lot of computation to the subsequent processing of the model. Therefore, in the decoding stage, the top ten results in the model output set is adopted to evaluate the model effect.

2.5. Experimental Dataset. The content used for testing in this section includes three parts: non-word error processing, true word error processing, and sentence grammar processing. Here, English texts with different degrees of difficulty are selected to test the impact of the selection of training corpus on the construction of dictionaries. The compositions of non-English majors (ST3 and ST4) in the Chinese learner corpus are taken as test examples of non-word error processing and sentence grammar processing. Four topics are selected, a total of 120 compositions (Table 1). These 120 compositions have non-lexical errors and grammatical errors in varying degrees, and have been manually marked with errors. Table 1 shows the specific symbol marking information related to the test. In addition, some common misspelled words are extracted from CET-4 and CET-6 exercises as supplementary test examples of non-word error handling. For the true word error test, the remaining 20% of the sentences related to the confusion set in the training corpus are taken as test examples.

3. Results and Discussion

3.1. Grammar Error Handling. In order to check and correct the grammatical errors of sentences, one of the most important problems is to deal with the text sentence segmentation. This system uses the rule-based sentence segmentation method, which adds a hypothetical boundary to the input text, and then uses the rule method to correct errors and realize the sentence segmentation function. The test examples come from 120 compositions in the experimental data. Table 2 shows the specific punctuation distribution.

Table 2 shows that if the question mark, exclamation marks and periods are directly taken as sentence boundaries, the sentence segmentation accuracy is only 66.43%. The result of sentence segmentation is the basis of sentence grammatical error analysis, so it is necessary to correct the hypothetical sentence boundary. The following is the sentence segmentation results of the test case by the system using the method based on error correction rules (Table 3).

The above error correction results show that the sentence segmentation accuracy is 98.96%, and the error correction effect is obvious.

By studying and analyzing the grammatical errors in CET-4 and CET-6 composition, this system uses the combination of a neural network model and artificial grammar rules to solve the common grammatical errors in writing. 110 test cases with these grammatical errors are extracted from the experimental data text, and some contain more than one kind of grammatical error. In the test, the system's error detection and error correction accuracy for sentence grammatical errors are mainly investigated. Figure 6 shows the test results.

Table 1: Main data sources of grammatically incorrect texts.

Composition topic	Source	Quantity
Practice makes perfect	ST3	30
Global shortage of fresh water	ST3	30
My view on job-hopping	ST4	30
My view on fake commodities	ST4	30

The data in Figure 6 show that the average correction accuracy of the system for CET-4 and CET-6 compositions is 82.69%. Through the observation of the data, the accuracy of the system in detecting and correcting prepositional errors and the inconsistency of singular and plural nouns is not high enough. The reasons are explained below.

The number of prepositions involved in English is relatively large, and the context in which they are used is not fixed, which makes the weight distribution of context features extracted in the training process more scattered, and the discrimination of features is low. Besides, during the training process, some prepositions contain similar contextual features, such as prepositions "in" and "on." Based on these two main reasons, the system is not easy to find the optimal preposition in prediction, which reduces the performance of error detection and correction.

The inconsistency between singular and plural nouns involves two cases: mistakenly writing uncountable nouns into plural form and mistakenly writing plural form of countable nouns into a singular form. Because some nouns can be treated as both countable and uncountable in actual writing, there will be some misjudgments and omissions. Besides, there are many kinds and numbers of leading words to identify whether the noun should be the plural form, and there are still some omissions in the definition of rules. It is believed that the system performance can be improved with the improvement of the rules.

3.2. KNGTC and Comparison Test Accuracy. The KNGTC model is compared with two comparative models to evaluate its performance. First, the probabilistic graphical model (PGM). The method of PGM is used to realize the training of English grammar error correction. Their model finds a joint global optimal correction type in the whole input statement sequence. Second, Google translate (Google *T*). Google *T* is a translation software developed by Google, which supports grammar checking. To some extent, it can represent the performance of mature language translation technologies in the current market. Figure 7 shows the accuracy comparison of the three modes:

Figure 7 shows that KNGTC outperforms the PGM model in almost all data types, especially on LAP and Google *T*. The main reason is that the translation information contains less information and is sparse. The PGM model can hardly deal with the user input in simplified form, while KNGTC can effectively generate reasonable semantic information and spelling form according to the global information. In addition, KNGTC can use key transfer probability and alignment information, so it can select more reasonable candidate results.

Table 2: The punctuation of the text to be separated.

	Number of question marks	Number of exclamation marks	Number of periods	The actual number of sentences	Sentence boundary ratio (%)
Test case	531	11	1475	1340	66.43

Table 3: Sentence segmentation results.

	Number of correctly found sentences	Number of all sentences found	Sentence segmentation accuracy (%)
Test case	1345	1359	98.96

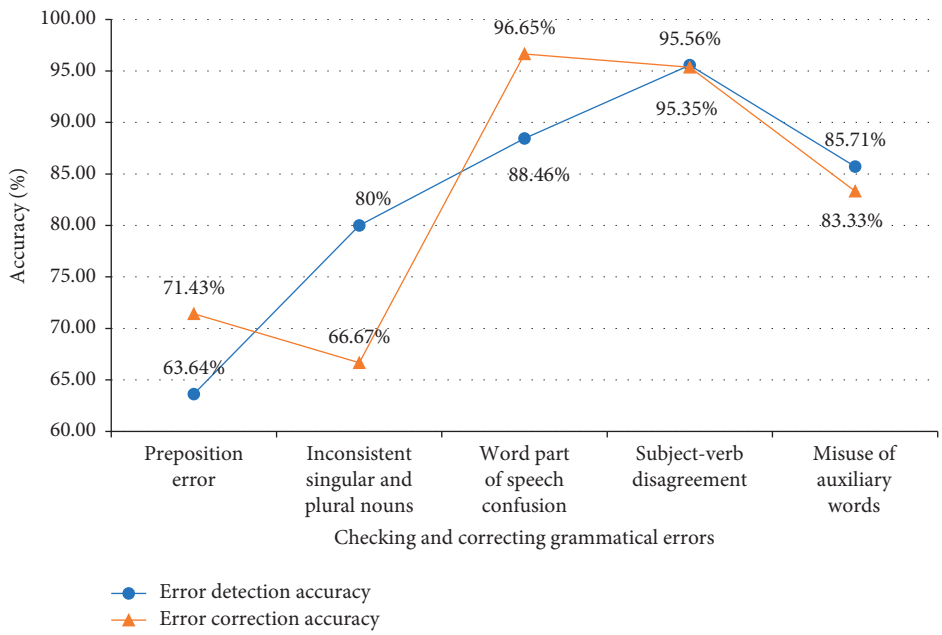


Figure 6: Grammar test results of the selected instance to be tested.

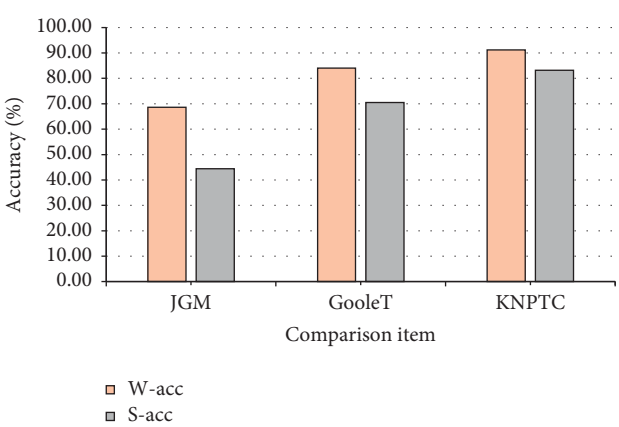


Figure 7: Accuracy of KNGTC and comparison models.

KNGTC performs as well as or even better than Google *T* in various data types, and achieves an accuracy improvement of nearly 17% in the comprehensive results. After analysis, it is found that the main reason for the poor performance of Google *T* is that the existence of simplified spelling or wrong input affects its accurate separation of user input. However, KNGTC overcomes this problem.

4. Conclusion

This work mainly studies the computational neural model of college English grammatical error correction. The main conclusions are as follows. (1) This work mainly summarizes RNN, establishes the overall architecture of the computational neural model of college English grammatical error correction, and studies the supervised training of Attention mechanism, experimental environment and dataset. (2) Through comparative experiments, it is found that the accuracy of grammar correction of KNGTC is higher than that of other models. The average accuracy of this model is 82.69%.

Due to the limited time and level, this work still has the following shortcomings. (1) The model can use a larger scale of corpus for training. The neural network model has a strong ability to fit data distribution. Theoretically, using more abundant user data can effectively enhance the error correction ability of the grammar error correction model. (2) The user's click position information is introduced into the model, and the input vector splicing method is used in the process. More ways can be considered to combine the prior information with the error correction model. (3) The

reordering model of prediction results based on the language model is not deep enough, and the structure is not complex enough, which leads to less obvious improvement of the prediction effect.

Data Availability

The dataset used in this paper are available from the corresponding author upon request.

Conflicts of Interest

The authors declared that they have no conflicts of interest regarding this work.

References

[1] A. Alishahi, G. Chrupała, and T. Linzen, "Analyzing and interpreting neural networks for NLP: a report on the first Blackbox NLP workshop," *Natural Language Engineering*, vol. 25, no. 4, pp. 543–557, 2019.

[2] M. Baroni, "Linguistic generalization and compositionality in modern artificial neural networks," *Philosophical Transactions of the Royal Society B: Biological Sciences*, vol. 375, Article ID 20190307, 2020.

[3] G. L. Bosco, G. Pilato, and D. Schicchi, "A neural network model for the evaluation of text complexity in Italian language: a representation point of view," *Procedia Computer Science*, vol. 145, pp. 464–470, 2018.

[4] A. Boyd, "Using Wikipedia edits in low resource grammatical error correction," *Proceedings of the 2018 EMNLP Workshop W-NUT: The 4th Workshop on Noisy User-Generated Text*, pp. 79–84, Brussels, Belgium, 2018.

[5] C. W. Coley, W. Jin, L. Rogers et al., "A graph-convolutional neural network model for the prediction of chemical reactivity," *Chemical Science*, vol. 10, no. 2, pp. 370–377, 2019.

[6] E. Fetaya, Y. Lifshitz, E. Aaron, and S. Gordin, "Restoration of fragmentary Babylonian texts using recurrent neural networks," *Proceedings of the National Academy of Sciences*, vol. 117, no. 37, pp. 22743–22751, 2020.

[7] C. Gan, L. Wang, and Z. Zhang, "Multi-entity sentiment analysis using self-attention based hierarchical dilated convolutional neural network," *Future Generation Computer Systems*, vol. 112, pp. 116–125, 2020.

[8] S. V. Gothe, S. Dogra, M. Chandra, and C Sanchi, "An efficient system for grammatical error correction on mobile devices," in *Proceedings of the 2021 IEEE 15th International Conference on Semantic Computing (ICSC)*, pp. 147–154, IEEE, Laguna Hills, CA, USA, January 2021.

[9] M. Heidari and S. Rafatirad, "Semantic convolutional neural network model for safe business investment by using bert," in *Proceedings of the 2020 Seventh International Conference on Social Networks Analysis, Management and Security (SNAMS)*, pp. 1–6, IEEE, Paris, France, December 2020.

[10] S. Islam, M. F. Sarkar, T. Hussain, M Mehedi Hasan, and M. F Dewan, "Bangla sentence correction using deep neural network based sequence to sequence learning," in *Proceedings of the 2018 21st International Conference of Computer and Information Technology (ICCIT)*, pp. 1–6, IEEE, Dhaka, Bangladesh, December 2018.

[11] S. N. Khan and I. Usman, "Amodel for English to Urdu and Hindi machine translation system using translation rules and artificial neural network," *The International Arab Journal of Information Technology*, vol. 16, no. 1, pp. 125–131, 2019.

[12] N. Koul and S. S. Manvi, "A proposed model for neural machine translation of Sanskrit into English," *International Journal of Information Technology*, vol. 13, no. 1, pp. 375–381, 2021.

[13] Y. Lakretz, D. Hupkes, A. Vergallito, M. Marelli, M. Baroni, and S. Dehaene, "Mechanisms for handling nested dependencies in neural-network language models and humans," *Cognition*, vol. 213, Article ID 104699, 2021.

[14] J. Lee, J. Lee, M. Lee, and G. J. Jang, "Named entity correction in neural machine translation using the attention alignment map," *Applied Sciences*, vol. 11, no. 15, p. 7026, 2021.

[15] P. Le-Hong and A. C Le, "A Comparative Study of Neural Network Models for Sentence Classification," in *Proceedings of the 2018 5th NAFOSTED Conference on Information and Computer Science (NICS)*, pp. 360–365, IEEE, Ho Chi Minh City, Vietnam, November 2018.

[16] S. Li, J. Zhao, G. Shi et al., "Chinese grammatical error correction based on convolutional sequence to sequence model," *IEEE Access*, vol. 7, pp. 72905–72913, 2019.

[17] N. Madi and H. S. Al-Khalifa, "A proposed Arabic grammatical error detection tool based on deep learning," *Procedia Computer Science*, vol. 142, pp. 352–355, 2018.

[18] N. Madi and H. Al-Khalifa, "Error detection for Arabic text using neural sequence labeling," *Applied Sciences*, vol. 10, no. 15, p. 5279, 2020.

[19] R. T. McCoy, R. Frank, and T. Linzen, "Does syntax need to grow on trees? sources of hierarchical inductive bias in sequence-to-sequence networks," *Transactions of the Association for Computational Linguistics*, vol. 8, pp. 125–140, 2020.

[20] A. F. d. S. Neto, B. L. D. Bezerra, and A. H. Toselli, "Towards the natural language processing as spelling correction for offline handwritten text recognition systems," *Applied Sciences*, vol. 10, no. 21, p. 7711, 2020.

[21] M. Qin, "A study on automatic correction of English grammar errors based on deep learning," *Journal of Intelligent Systems*, vol. 31, no. 1, pp. 672–680, 2022.

[22] H. Maleki, A. Sorooshian, G. Goudarzi, Z. Baboli, Y. Tahmasebi Birgani, and M. Rahmati, "Air pollution prediction by using an artificial neural network model," *Clean Technologies and Environmental Policy*, vol. 21, no. 6, pp. 1341–1352, 2019.

[23] M. A. Mohammed, K. H. Abdulkareem, S. A. Mostafa et al., "Voice pathology detection and classification using convolutional neural network model," *Applied Sciences*, vol. 10, no. 11, p. 3723, 2020.

[24] A. Roshanzamir, H. Aghajan, and M. Soleymani Baghshah, "Transformer-based deep neural network language models for Alzheimer's disease risk assessment from targeted speech," *BMC Medical Informatics and Decision Making*, vol. 21, no. 1, pp. 92–14, 2021.

[25] A. Solyman, W. Zhenyu, T. Qian, A. A. M. Elhag, M. Toseef, and Z. Aleibeid, "Synthetic data with neural machine translation for automatic correction in Arabic grammar," *Egyptian Informatics Journal*, vol. 22, no. 3, pp. 303–315, 2021.

[26] C. Tsoukala, M. Broersma, A. Van den Bosch, and S. L. Frank, "Simulating code-switching using a neural network model of bilingual sentence production," *Computational Brain & Behavior*, vol. 4, no. 1, pp. 87–100, 2021.

[27] X. Wang, C. Chen, and Z. Xing, "Domain-specific machine translation with recurrent neural network for software localization," *Empirical Software Engineering*, vol. 24, no. 6, pp. 3514–3545, 2019.

[28] A. Warstadt, A. Singh, and S. R. Bowman, "Neural network acceptability judgments," *Transactions of the Association for Computational Linguistics*, vol. 7, pp. 625–641, 2019.

[29] G. Zhang, "A study of grammar analysis in English teaching with deep learning algorithm," *International Journal of Emerging Technologies in Learning (iJET)*, vol. 15, no. 18, pp. 20–30, 2020.

Hindawi
Computational Intelligence and Neuroscience
Volume 2022, Article ID 6431776, 10 pages
https://doi.org/10.1155/2022/6431776

Research Article

Distributed Intelligent Learning and Decision Model Based on Logic Predictive Control

Yucheng Zhou,[1] **Wen Lu**,[2] **and Yingqiu Zhang**[3]

[1]*Chongqing Jiaotong University, Chongqing 400074, China*
[2]*School of Sciences, Zhejiang SCI-TECH University, Hangzhou, Zhejiang 310000, China*
[3]*School of Competitive Sports, Beijing Sport University, Beijing 100084, China*

Correspondence should be addressed to Yucheng Zhou; yczhou@cqjtu.edu.cn

Received 10 June 2022; Revised 18 July 2022; Accepted 25 July 2022; Published 30 August 2022

Academic Editor: Ning Cao

By the method of documentation and logical analysis, based on the data, based on logic and based on the knowledge of three kinds of artificial intelligence in the sports education, the intelligent learning system feedback delay are studied, combined with mobile communication which led to the artificial intelligence online sports games teaching, pattern recognition, and virtual technology combined with innovative teaching interaction and experience. Promoting the development of green PE teaching machine learning can identify the types of PE activities and realize efficient PE learning diagnosis. Intelligent decision support system can identify sports talents and improve the effect of personalized PE teaching evaluation. From the perspective of psychological development and education, the key problems to be solved in the integration of artificial intelligence and physical education are examined. Then, the consistent model predictive control for feedback delay of nonlinear sports learning multiagent system with network induced delay and random communication protocol is studied. Under the communication waiting mechanism designed, each agent has a certain tolerance of delay, and this tolerance can be determined by ensuring the stability of the system. At the same time, a random communication protocol is designed to ensure the ordered communication of the multiagent system. Finally, the effectiveness of the proposed algorithm is verified by numerical simulation. To solve the channel competition access problem of the sports intelligent learning system with special structure feedback delay model predictive control, a dual channel awareness scheduling strategy under the model predictive control framework was proposed, and the distributed threshold strategy of sensors and the priority threshold strategy of controllers were designed. It is proved that the sensor will eventually work at Nash equilibrium point under the policy updating mechanism, and the priority threshold strategy of the controller is better than the traditional independent and identically distributed access strategy. By avoiding the data transmission when the channel status is poor, the channel access of the system is efficient and saves energy.

1. Introduction

AI (artificial intelligence) technology is applied in training and competition stages, mainly through hardware terminals such as camera equipment and sensors for deep learning. In the process of sports training and competition, we maintain gentle state of mind which is a basic element, obtaining good results using machines to identify facial expressions in the video and image, and according to the facial features to predict detection machines, machine needs to face detection image processing tools and basic emotional states such as public data as the foundation; our country already has a mature machine for complex facial recognition task [1]. Foreign scholars use this method, using the tracking data, to construct the data-driven model. This model is mainly generated from the first person image reliable basketball action sequences, by way of unsupervised learning, and is widely used in the field of sports. Chinese scholars depend on the practical field of China sports development training, increase the intensity of research model, and understand the training and competition situation through artificial intelligence technology [2].

Networked model predictive control (NMPC) is a good method to deal with network induced delay and packet loss. The basic idea of model predictive control algorithm is to predict the future behavior of the controlled object based on

the real-time state information of the controlled object, and the prediction is rolling forward; that is to say, the controlled object will re-predict once every period of time [3]. When the induced delay and packet loss occur in the controlled network, the system can iteratively optimize the received information according to the previous time, so as to find the optimal state sequence. For multiagent systems, when there is network induced delay or packet loss between agents, this method can also be used to compensate effectively. However, compensating the missing information directly with the feasible solution of model predictive control will inevitably bring errors, so using a waiting mechanism to deal with the missing information can ensure the authenticity of data.

At present, online sports game teaching is still inadequate, its teaching effect depends on the specific type of game, the game may induce distraction, and sustainability of the teaching effect needs to be verified. At present, "5G + physical education" is ready to take off. Artificial intelligence will rely on the characteristics of 5G technology, such as large connection, ultralow delay, and ultrahigh speed, to innovate the content and communication mode of physical education and further promote the innovation of sports game teaching scenes and teaching experience.v

2. Related Work

The combination of big data and mobile communication technology makes online physical education flourish. Mobile Internet game courses represented by serious games and sports games innovate physical education teaching methods and can promote students' physical and mental development in multiple dimensions. Currently, sports game courses integrate big data physiological measurements to optimize exercise intensity and exercise plan, which can be extended to family sports. Serious games refer to video games for the purpose of education, health, or scientific research, aiming to stimulate and develop the motivation of physical education and exercise habits. A large number of studies have proved that serious games have positive effects on academic knowledge, cognitive ability, professional and technical knowledge, learning motivation, and academic achievement [4, 5]. Foreign researchers developed the game "Virtual Physical Education Teacher," which can design sports activities for different student groups based on big data feedback [6]. Sports games refer to games that combine body movement with electronic games based on motion sensor technology, represented by dance games. Because dance games can consume a lot of calories, they have been introduced into American youth sports courses and received positive feedback from parents and students. The skills acquired by young people in sports games can be transferred to promote physical, social, and cognitive development.

The combination of pattern recognition and model predictive control, feedback delay of sports intelligent learning system, augmented reality technology, and mixed reality technology provides a new window for the development of physical education technology. Pattern recognition technology captures and identifies movements through wearable devices to monitor physical education [7]. Sports courses with special requirements such as snow sports and golf break the limitation of teaching environment and reduce the risk of injury by combining the feedback delay of intelligent sports learning system with intelligent wearable devices and model predictive control. Augmented reality technology relies on wearable devices and obtains real-time multi-perspective image information through computer vision, so as to obtain 3D movement information of students and give evaluation. Mixed reality technology relies on wearable devices to achieve immersive physical education through user and environment interaction [8]. In addition, wearable devices can also be combined with virtual personal assistants to realize human-computer interaction in physical education teaching through emotional computing and cloud computing technology and develop personalized physical education teaching plans. The combination of pattern recognition and model predictive control with feedback delay of sports intelligent learning system can also meet the requirements of sports teaching for special groups. Recent studies have shown that the combination of brain science and model predictive control sports intelligent learning system feedback delay can stimulate the brain nerves of patients with lower body movement impairment, enabling them to imagine walking, obtain virtual movement experience, and promote rehabilitation [9]. At present, the combination of pattern recognition and model predictive control with feedback delay in sports intelligent learning system is still insufficient. For example, the teaching effect is limited by the type of sports skills. However, for open baseball, the migration effect of outfielders catching the ball is poor. At present, the feedback delay of model predictive control sports intelligent learning system may be more suitable for closed sports skill learning, and its ecological validity needs to be improved.

Machine learning is mainly used for intelligent classification and prediction. At present, machine learning is mostly used in competitive sports, and its influence on physical education is mainly reflected in two aspects. First, deep learning algorithms can be used for sports activity type recognition. For example, artificial neural networks can assess individual exercise metabolic equivalents and determine individual activity types (low intensity activity, exercise, vigorous exercise, and housework/other activities). Computational modeling can monitor and feed back muscle state in real time to predict fatigue and avoid sports injuries. Second, deep learning algorithm can be used for PE learning diagnosis and performance prediction. By mining the historical data of training and competition to predict the results of the competition [10], it provides data basis for hierarchical physical education. At present, PE learning diagnosis requires high precision and accuracy of deep learning algorithm, and its accuracy needs to be optimized.

Model predictive control algorithm is an advanced control technology developed recently [11]. The initial form of model predictive control is dynamic matrix control [12], model predictive heuristic control [13], etc. In [14], the generalized model predictive control was proposed by researchers. However, due to the underdeveloped science and technology at that time, model predictive control

(MPC) is difficult to be widely used in industry due to its large amount of computation. Until the information age, the relevant theories of model predictive control have been improved, so the examples closely combining model predictive control algorithm with practical application also show diversity [15]. Model predictive control with other optimization control algorithm: model predictive control for the cost function is to optimize the real time updated in each moment will be predicted, which is often said that it will roll forward in time domain of a time step, so the model predictive control is also known as rolling time domain control. Compared with traditional control methods, the special characteristics of model predictive control can be summarized as four points [16]. Finally, model predictive control has the ability to explicitly deal with all kinds of constraints of the system. In model predictive control, constraint conditions are written into the optimization problem in the form of mathematical expressions and solved directly in a mathematical way, so as to obtain the control quantity and state quantity satisfying the constraints. Therefore, model predictive control has attracted much attention in recent years [17, 18].

So far, comprehensive research on model predictive control at home and abroad was carried out. Model predictive control is roughly divided into three types: centralized model predictive control, decentralized model predictive control, and distributed model predictive control. In [19], a centralized model predictive controller is designed for multiagent systems. However, the decentralized model predictive control mentioned in [20] is based on a decentralized model, which only contains the input and output information of each agent without information interaction between agents [21]. Distributed model predictive control allows communication and information sharing among agents to achieve global optimality of agent systems. Therefore, DMPC reduces the complexity of optimization problems and improves work efficiency. In addition, model predictive control can not only deal with multiple constraints, but also calculate the control input sequence through online solution [22, 23]. The prediction and evaluation results are mainly nonlinear models based on machine learning algorithms. This is to improve the fitting ability of the regression model and accurately predict the explained variables. A prediction model of sports performance based on machine learning algorithm is proposed [24]. Its working principle is to improve the status quo of performance prediction, improve the evaluation level of physical fitness, and make the prediction results accurate by using the principle of structural risk minimization.

3. Feedback Delay Diagnosis of Sports Intelligent Learning System Based on Model Predictive Control

3.1. Sports Intelligent Learning Feedback Delay Diagnosis. Machine learning is mainly used for intelligent classification and prediction. At present, machine learning is mostly used in competitive sports, and its influence on physical education is mainly reflected in two aspects. First, deep learning algorithms can be used for sports activity type recognition. For example, artificial neural networks can assess individual exercise metabolic equivalents and determine individual activity types (low intensity activity, exercise, vigorous exercise, and housework/other activities). Computational modeling can monitor and feed back muscle state in real time to predict fatigue and avoid sports injuries. Second, deep learning algorithm can be used for PE (physical education) learning diagnosis and performance prediction. By mining the historical data of training and competition, we can predict the result of the competition and provide data basis for stratified physical education. At present, PE learning diagnosis requires high precision and accuracy of deep learning algorithm, and its accuracy needs to be optimized. The time-delay frame diagram of sports intelligent learning feedback is shown in Figure 1.

As shown in Figure 1, model predictive control is an important development direction in sports training activities, which mainly refers to the simultaneous acquisition, processing, editing, storage, and display of two different types of information media technologies. Multimedia computer-assisted instruction needs to be based on the artificial intelligence technology and complex program; in the process of sports training, multimedia computer technology is used to realize the three-dimensional model, and the plane model interacts and promotes the teaching effect. The main purpose is to provide different stimulus in the process of sports training for athletes, control direct teaching environment to keep their nervous system excitability, make them interested in sports, help athletes broaden their horizons, train athletes' intelligence, and promote the development of athletes' personality, with greater superiority. Multimedia computer aided instruction belongs to the new trend of sports technology development, which effectively creates a relaxed environment for athletes and makes the feedback delay of sports intelligent learning system play a practical role.

Artificial intelligence technology has been applied in athlete training. Statistics of sports training are based on numbers, and automatic digital report is the development direction of sports training, which can be realized only by artificial intelligence technology. Artificial intelligence technology can completely change live broadcasting. According to sports events, artificial intelligence technology can choose the right successive angle, understand the real-time situation of athletes, and timely find the emergency situation in the training process.

It is an effective way to solve dynamics problems by using neural network intelligent technology. Dynamics is an important content in sports training. Dynamics mainly uses dynamic data during the execution and separation of 90-degree cutting movements to provide theoretical support for athletes' sports training plans. The principle of neural network intelligent technology is to use feedforward neural network to predict the torque and ground response of joints, knee joints and manic joints, and record the trajectory and joint angle as the input data of neural network intelligent

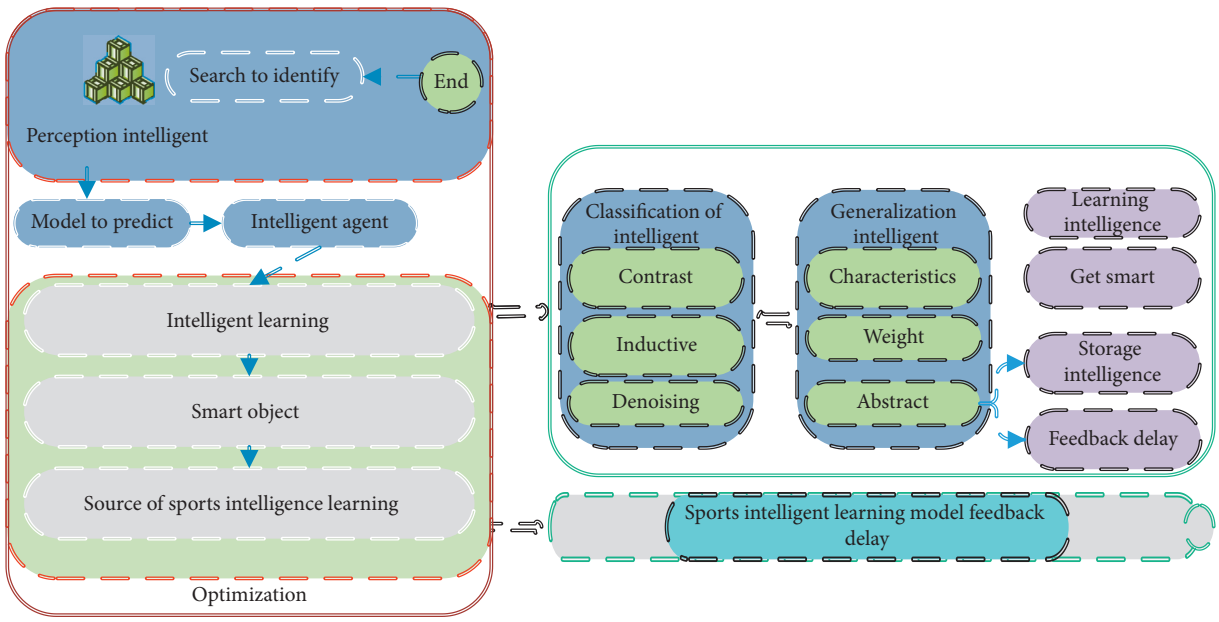


Figure 1: Time-delay frame diagram of sports intelligent learning feedback.

technology. In the actual process of sports training, the training data and prediction data of athletes have a strong correlation in the process of sports. The joint angle of an athlete is used as an input parameter. The higher the prediction accuracy of joint torque, the higher the prediction accuracy of ground reaction force. Through the use of inertial sensors, the rapid changes of movement direction and mechanical parameters can be predicted to help the athlete improve the sports training program.

3.2. Research on Model Predictive Control Algorithm. The structure of the sports intelligent learning system of model predictive control is shown in Figure 2, in which n independent control systems are controlled by remote controllers installed on the cloud computing platform, and the sensing data and control data are transmitted through wireless channels. The channel from the sensor side to the controller side is called the measurement channel, and the channel from the controller side to the actuator side is called the control channel.

There are two main reasons for packet loss in the sports intelligent learning system based on model predictive control. One is packet collision; that is, the wireless channel allows only one user access at most at any time, and the simultaneous access of multiple users will lead to transmission failure of all users. The other is transmission error, where packets may be discarded due to errors in transmission due to interference and signal fading.

The cost function is as follows:

$$R(x(n), y(n)) = |x(n|k) - y(n|k)|^2. \quad (1)$$

In the standard of rolling optimization mechanism, inequality is usually meet; however, if the given reference trajectory is a cycle of time-varying reference trajectory, it may be a sudden change, so it is difficult to find a suitable weighting matrix to ensure inequality always meet, and inequality cannot be sure; it is difficult to prove the system stability. Therefore, a time-varying weight matrix is considered in this section to ensure the inequality. Design the following corresponding terminal cost function, terminal domain, and local controller:

$$V_j(n|k)t = |x(n) - v(n|k)|^2. \quad (2)$$

The model predictive control algorithm is shown in Algorithm1:.

In sports training activities, the musculoskeletal system injury of athletes is more serious, and artificial intelligence technology can be used to monitor the quality of human movements. In the process of movement analysis of tennis special sports technology, AI technology is the future development direction of tennis training. Tennis needs to be separated according to the technology and movement of the sport purpose, and the important content of each subitem is determined. The key direction is to establish excellent technical model, based on the technical movement parameter data of excellent athletes. The weighted value of node connection in the second part represents the connection signal. The output mainly depends on the network connection mode, and the weighted value changes with the output function. In motion learning, the artificial neural network is used to record data for many times to obtain the motion sequence, and the neuron output is used. It is helpful to analyze the movement pattern of athletes.

The combination of expert system and decision support system can form intelligent decision support system, which can assist in solving the decision-making and evaluation problems of physical education. Intelligent decision support system can be used for personalized

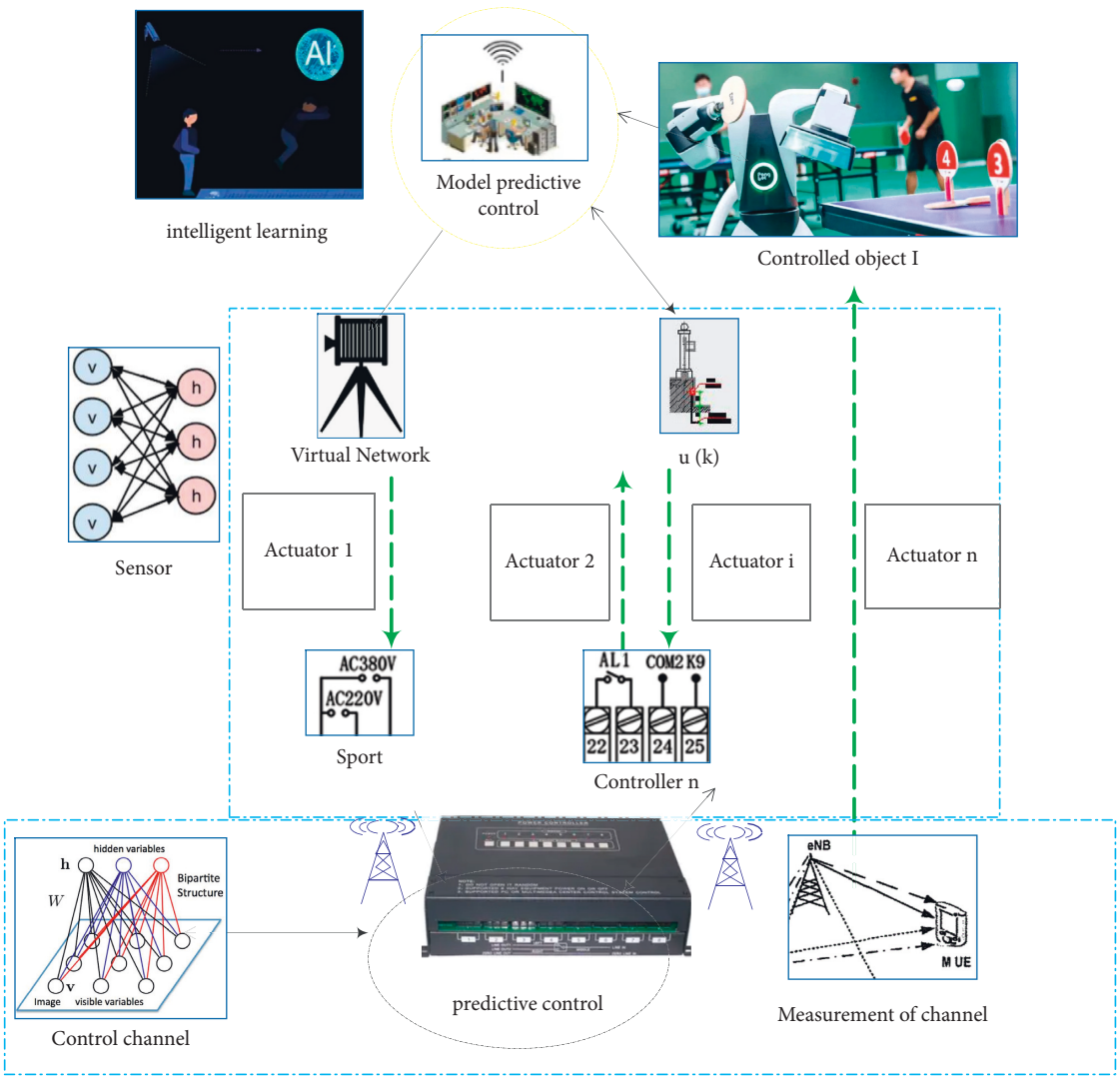


Figure 2: Structure of sports intelligent learning system based on model predictive control.

Offline: determine the parameters Q, P, R, S of the nonlinear model predictive control scheme.
Online: time $k = 0, 1, 2$.
(1) The initial state $x(k)$ measured at the initial moment $k = 0$ and the reference trajectory of a given period, wherein the period T of the reference trajectory of a given period is known.
(2) The time-varying matrix is obtained by solving the linear matrix inequality.
(3) By solving the optimization problem, an artificial reference trajectory and an optimization control sequence can be obtained.
(4) For all I, the controlled object adopts control input.
(5) The $k = k + 1$.
(6) Return to Step 1.

Algorithm 1: Overview of nonlinear model predictive control algorithm.

sports training. The expert system of hurdle sports can generate the training load of hurdle sports and predict the result. There are also studies on the use of big data mining and analysis, analysis of learning preferences, and learning strategies, to achieve personalized physical education. Intelligent decision support system can be used for sports material selection.

4. Sports Intelligent Learning System Feedback Delay

From the basic framework of sports intelligent learning system, we can know that predicting the behavior of other agents plays a very important role in the learning results of agents. Classical Q learning assumes that the environment is

static, so model-free reinforcement learning can be adopted. However, for agent I, the generalized environment is dynamic, so model-free learning cannot be adopted. Model-based reinforcement learning learns the optimal strategy by modeling the environment. Because it can reflect the characteristics of the environment in time and the state prediction mechanism can reflect the dynamics of other agents, the combination of the two can build a model that reflects the dynamic environment, namely, the generalized environment. Based on the above ideas, if the environmental dynamics can be accurately and timely estimated and predicted, then, to a certain extent, the model of generalized environment is accurate and stable. Therefore, the generalized environment can be regarded as a nearly static environment, and then the optimal strategy can be obtained by dynamic programming method.

Dynamic corresponding agent: because the environment is caused by other intelligent beings, as a result, accurate and timely strategy to estimate the other agent and the dynamic characteristics of the environment to be effective are modeled, and then, using the reinforcement learning for optimal decision to adapt to the dynamic characteristic, the comparison is carried out with the classical control theory of tracking control. Before designing the state prediction mechanism, the basic definition of state prediction is given.

$$P(x, a) = (1 - \lambda)P\left(x, a^{j}\right). \tag{3}$$

In the feedback delay framework of sports intelligent learning system based on model predictive control, the Q function will become a dynamic process due to the existence of other agents. This means that due to the change of other agents' strategies, the optimal action in the current state may become the worst action in the next experience, and the corresponding Q table will also change. However, this situation does not exist in single-agent reinforcement learning and multiagent learning with joint states and joint actions. Therefore, the action selection mechanism mentioned above cannot be used to deal with the dynamic environment in which the feedback delay framework of the model predictive control sports intelligent learning system exists. Otherwise, action selection oscillations will occur in the agent's behavior, and even the agent cannot complete the task in an effective time. Therefore, a new action selection mechanism needs to be designed. The improved action selection mechanism fully considers the dynamic nature of the environment and can effectively guide agents to balance the "exploration-utilization" problem. See Algorithm 2.

AI can change the structure of the brain and improve cognitive performance. For example, action video games using computer vision can enhance attention, improve multi-objective tracking, and speed up reaction times. AI can also boost motor cognitive processing and improve performance. For example, electroencephalography can be used to explore brain activity during exercise. Neurofeedback training can improve attention and working memory, so as to develop soccer shooting skills and archery skills.

There is a spiral relationship between the application of artificial intelligence and cognitive development, and the two promote each other. Smart PE teaching should conform to students' cognitive level, consider the potential impact of pulse technology on cognitive development, timely adjust teaching requirements, and achieve accurate teaching. Follow-up research should focus on the role of artificial intelligence in the development of sports core literacy and focus on exploring the relationship between artificial intelligence, cognition, and physical education, so as to provide a cognitive perspective for in-depth analysis of the mechanism of artificial intelligence in physical education.

Artificial intelligence promotes the development of ubiquitous learning and adaptive learning, and physical education is further personalized and fragmented, which brings the difficulty of monitoring the teaching effect. The contradiction between personalized learning and collectivized teaching system is becoming more and more prominent. Objectively, it is necessary to constantly upgrade technical monitoring means and provide accurate and systematic big data monitoring support services; subjectively, physical education teachers should combine the students' personality characteristics, reconstruct the teaching content and the teaching process, and seek a multi-teaching evaluation method combining individuation and collectivized learning.

In sports, key note recognition is an important content of sports analysis, which provides effective sports analysis and auxiliary analysis for athletes. Initially, it mainly tracks each individual key gesture, using the characteristics of every athlete research group. In the later development process, its main purpose is to group the image processing and sampling and access global information and processing; data were collected to exceptions and are classified. The traditional technology has been unable to meet this requirement.

5. Example Verification

Because the feedback delay algorithm of model predictive control sports intelligent learning system adopts model-based reinforcement learning, Friend-Q belongs to model-free reinforcement learning. Generally speaking, model-based reinforcement learning has higher learning efficiency than model-free reinforcement learning because it records the learning experience of agents and can make full use of such historical information. In order to ensure the consistency of the comparison between the two algorithms, Friend-Q is implemented by model-based reinforcement learning, so that both algorithms belong to model-based reinforcement learning algorithm.

Similarly, in order to ensure the consistency of the premise of comparison as much as possible, the parameter settings with the same meaning in the above two algorithms are also the same, including the discount coefficient of 0.8, the threshold of 0.7, and the prediction function RME of 0.8. The experimental results are repeated 50 times, and the average value is taken.

The other two areas of comparison are the average number of steps required to reach the target state and the

(1) Initialization.
(2) Repeat the following steps:
(a) Observe the current environmental state s.
(b) Select actions according to the improved action selection mechanism.
(c) Observe and record the new state s', individual action A, other agent actions, and instant reward; then, update the state prediction function and state transition probability function, and calculate the strategy change function of agent J.
(d) Enter (e) if the target is reached; otherwise, return to (a).
(e) Compute the Q function using the priority scan algorithm.

Algorithm 2: Model prediction action selection mechanism algorithm.

average reward earned. As shown in Figures 3 and 4, the feedback delay algorithm of the sports intelligent learning system based on model predictive control can learn the optimal tracking strategy within shorter learning steps and obtain higher rewards at the same time. Therefore, the learning speed of the feedback delay of the sports intelligent learning system based on model predictive control is obviously faster than that of Friend-Q. In addition, the feedback delay of the sports intelligent learning system of model predictive control can learn the optimal strategy and reach the target state with a minimum of 7 steps (excluding the last step when reaching the target).

Figures 5 and 6 show the state response curve and real control input curve under the distributed access mechanism of sensors, the random priority threshold strategy of the controller, and PBMPC algorithm of the model in the measurement control and sports intelligent learning system. All systems are found to be stable and satisfy control constraints.

Compared with the networked fixed time domain, the efficiency of the control scheme can be verified. The simulation results can be seen in Figure 7 which shows that the time-varying prediction time domain obtained by the feedback delay algorithm of the model predictive control motion intelligent learning system is good. Figure 8 shows the calculation time of each step of the two control schemes and evaluates the computational complexity of the algorithm accordingly.

The feedback delay of sports intelligent learning system based on model predictive control belongs to the computer system simulation technology for creating and experiencing virtual world. It mainly integrates computer graphics, multimedia technology, artificial intelligence technology, and man-machine interface technology to provide support for creating and experiencing virtual world. The feedback delay of the model predictive control sports intelligent learning system enables athletes to have a deep understanding of sports with the help of virtual technology. In the virtual environment, athletes can get a change from perceptual rationality, so that they can have a deep understanding of sports spirit, actively explore and collect information, and develop their imagination ability. Sports intelligent learning system feedback delay of model predictive control in the process of sports training has a wide range of applications, for example, establishing physical exercise prescription sites and analyzing the practical development of sports. A three-dimensional

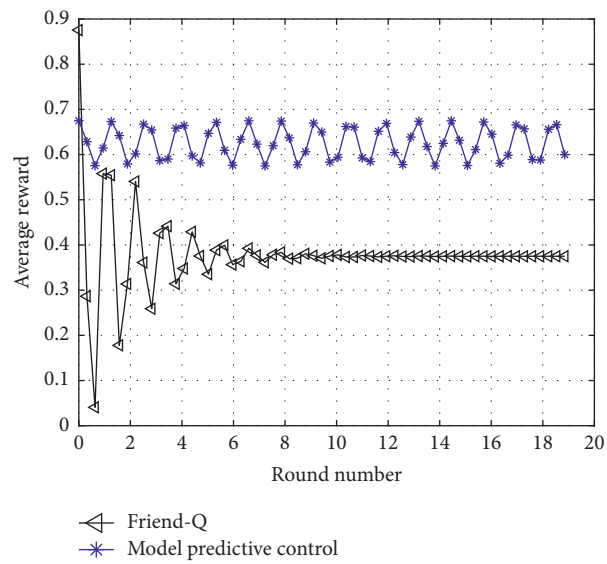


Figure 3: Average predictive control reward obtained by completing the task.

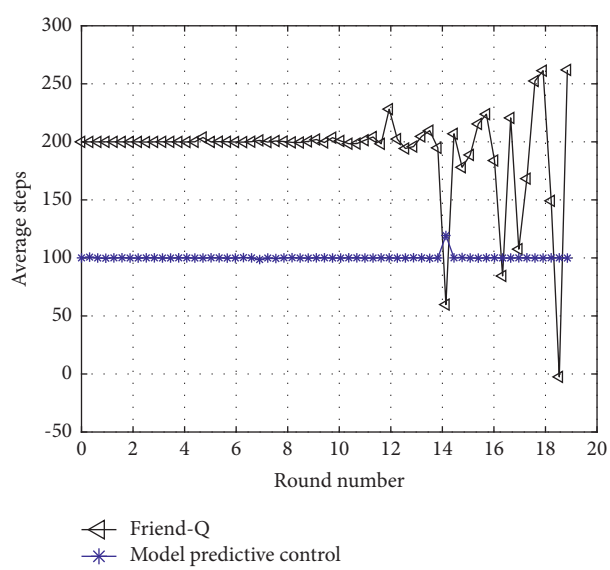


Figure 4: Average predictive control steps obtained by completing the task.

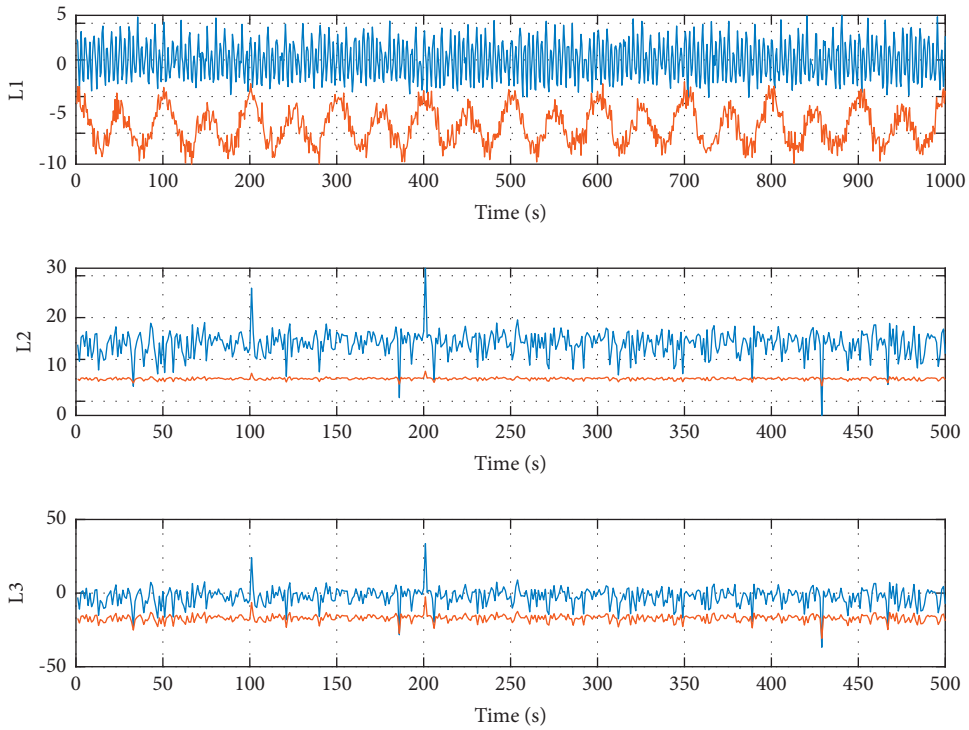


Figure 5: State response curve (the blue line is generated by the algorithm in this section; the red line does not have any control compensation scheme).

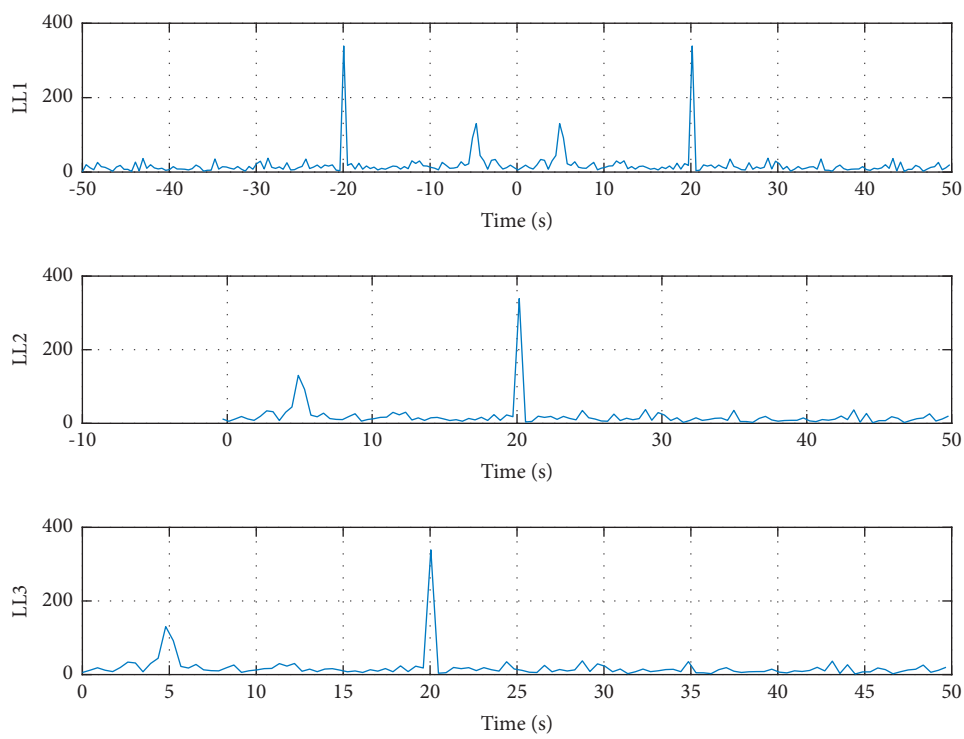


Figure 6: The real control object input of the system under the model pre-side control algorithm.

intelligent learning system with feedback delay is established by using model predictive control of motion, and a human motion model is established to make robots and athletes find problems in the process of motion. Thirdly, the feedback delay of the model predictive control sports intelligent learning system is simulated to the real world, and it is used to surpass the real world, so that the athletes can achieve the goal of high efficiency in the virtual environment, enhance the athletes'

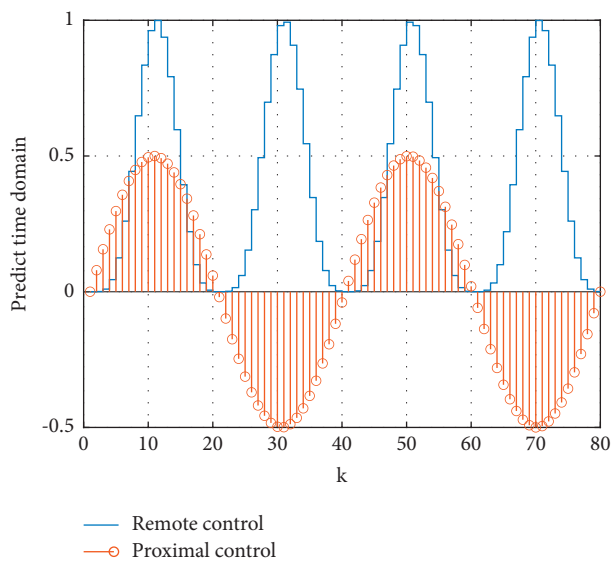


Figure 7: Time-domain variation curve of feedback delay of sports intelligent learning system based on model predictive control.

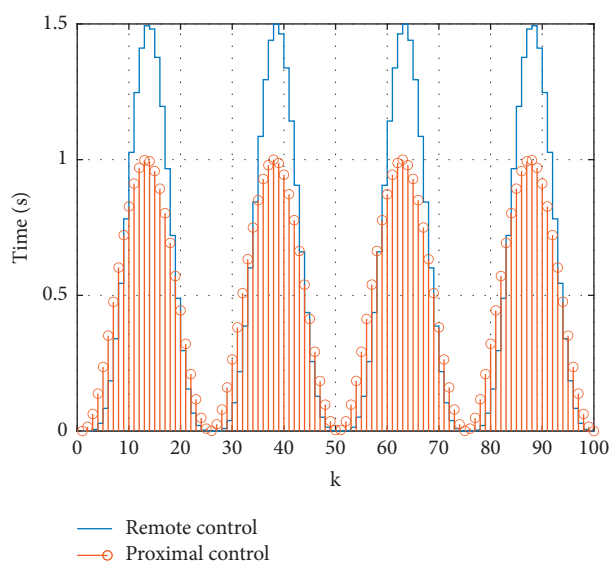


Figure 8: Solution time of feedback delay of sports intelligent learning system based on model predictive control.

experience of the virtual environment, and strengthen their understanding and mastery of sports.

6. Conclusion

In the process of integrating artificial intelligence technology into sports, the sports effect has made continuous progress, and artificial intelligence technology has been developing continuously. At present, important breakthroughs have been made in basketball, tennis, and skiing, and it is an important development direction to integrate artificial intelligence technology into sports. Data packet loss between agents is described by distribution, and compensation is made by using predictive sequence data loss of model predictive control. In order to avoid chaotic communication between agents, the largest prediction time domain is chosen as the next prediction time domain of the whole multi-intelligence system at each update moment. Finally, a numerical example is given to verify the effectiveness of the proposed algorithm. The integration of physical education and artificial intelligence can promote the construction and interaction of massive physical education resources, innovate teaching methods and interaction, provide physical education learning diagnosis and talent identification, promote the reform of physical education, and cultivate innovative sports talents. The future development of intelligent physical education requires that we should attach importance to the spiraling interaction between artificial intelligence and students' cognitive development, comprehensively consider the advantages of personalized and collectivized learning, dynamically examine teachers' information literacy, promote teachers' empowerment, and support multidimensional evaluation of students' physical ability. Condition forecasting method used in this study is based on the accurate observation of other agent history behavior after the forecast; in some practical applications, to the presence of large amounts of agents in the environment or obstacles leads to an agent's behavior observation which is not accurate, so we cannot update forecast function because of the deviation of decision-making. Therefore, the next step is to design a more robust state prediction method in order to improve the robustness of the whole system decision.

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 or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

The work of this paper was supported by the following: National Social Science Foundation of China, Grant no. 20BTY119: "Research on the System Construction and Operational Mechanism of Self-Motivation in Teenagers' Sports Exercise" and Research on Higher Education and Teaching Reformation in Chongqing, no. 213186, project name: "Research on the Construction and Application of Physical Education in Promoting Physical and Mental Health under Moral Education Perspective."

References

[1] X. Shen, G. Shi, Y. Zhang, and S. Weng, "Wireless volatile organic compound detection for restricted internet of things environments based on cataluminescence sensors," *Chemosensors*, vol. 10, no. 5, p. 179, 2022.

[2] X. Gou, W. Zhang, F. Zhang, J. Zhang, J. Zhang, and J. Zhang, "Research and analysis of step intelligent model predictive control of generator excitation system based on the field of building installation," *IOP Conference Series: Earth and Environmental Science*, vol. 632, no. 4, pp. 042001–042023, 2021.

[3] J. Cheng and X. Wang, "Artificial intelligence based on effectiveness of inverted classroom teaching of college sports [J]," *Journal of Intelligent and Fuzzy Systems*, vol. 40, no. 2, pp. 21–31, 2020.

[4] D. Zhang, Z. H. Ye, P. C. Chen, and Q. G. Wang, "Intelligent event-based output feedback control with Q-learning for unmanned marine vehicle systems," *Control Engineering Practice*, vol. 105, no. 5–6, pp. 104616–104634, 2020.

[5] X. Shen, G. Shi, H. Ren, and W Zhang, "Biomimetic vision for zoom object detection based on improved vertical grid number YOLO algorithm," *Frontiers in Bioengineering and Biotechnology*, vol. 10, no. 5, Article ID 905583, 2022.

[6] W. Ling, H. Ni, and R. Yang, "Intelligent virtual reference feedback tuning and its application to heat treatment electric furnace control[J]," *Engineering Applications of Artificial Intelligence*, vol. 46, no. NOV, pp. 121–129, 2015.

[7] W. Li, Z. Cao, and C. Zhu, "Intelligent feedback cognition of greengage grade based on deep ensemble learning[J]," *Nongye Gongcheng Xuebao/Transactions of the Chinese Society of Agricultural Engineering*, vol. 33, no. 23, pp. 276–283, 2017.

[8] L. Dejournett and J. Dejournett, "Silico testing of an artificial-intelligence-based artificial pancreas designed for use in the intensive care unit setting[J]," *Journal of Diabetes Science & Technology*, vol. 2016, pp. 193–229, 2016.

[9] D. He and B. Peng, "Gaussian learning-based fuzzy predictive cruise control for improving safety and economy of connected vehicles[J]," *IET Intelligent Transport Systems*, vol. 14, no. 5, pp. 763–778, 2020.

[10] B. M. Reese and E. G. Collins, "A graph search and neural network approach to adaptive nonlinear model predictive control," *Engineering Applications of Artificial Intelligence*, vol. 55, pp. 250–268, 2016.

[11] J. Liu, H. Qin, and G. Wang, "Control algorithm of permanent magnet direct drive belt conveyor system for mining based on reduced order model[J]," *International Journal of Pattern Recognition and Artificial Intelligence*, vol. 35, no. 14, pp. 672–683, 2021.

[12] Z. Ping, Z. Yin, and X. Li, "Deep Koopman model predictive control for enhancing transient stability in power grids[J]," *International Journal of Robust and Nonlinear Control*, vol. 31, pp. 892–903, 2021.

[13] B. Weismüller, M. Ghio, K. Logmin et al., "Effects of feedback delay on learning from positive and negative feedback in patients with Parkinson's disease off medication," *Neuropsychologia*, vol. 117, pp. 46–54, 2018.

[14] V. Pueyo, T. Pérez-Roche, E. Prieto et al., "Development of a system based on artificial intelligence to identify visual problems in children: study protocol of the TrackAI project," *BMJ Open*, vol. 10, no. 2, pp. e033139–e033153, 2020.

[15] H. Li, J. Ma, R. Huang et al., "Prevalence of vitamin D deficiency in the pregnant women: an observational study in Shanghai, China," *Journal of Intelligent and Fuzzy Systems*, vol. 78, no. 2, pp. 31–41, 2020.

[16] A. Guo and C. Yuan, "Network intelligent control and traffic optimization based on SDN and artificial intelligence," *Electronics*, vol. 10, no. 6, pp. 700–723, 2021.

[17] L. Zhang and X. Zhuan, "Model predictive control method of a parallel electromagnetic isolation system based on the improved genetic algorithm:[J]," *Journal of Vibration and Control*, vol. 26, no. 21-22, pp. 2001–2012, 2020.

[18] C. J. Ostafew, A. P. Schoellig, T. D. Barfoot, and J Collier, "Learning-based nonlinear model predictive control to improve vision-based mobile robot path tracking," *Journal of Field Robotics*, vol. 33, no. 1, pp. 133–152, 2016.

[19] F. Arnold and R. King, "State–space modeling for control based on physics-informed neural networks," *Engineering Applications of Artificial Intelligence*, vol. 101, no. May 2021, pp. 104195–104208, 2021.

[20] G. Guo and Y. Wang, "An integrated MPC and deep reinforcement learning approach to trams-priority active signal control," *Control Engineering Practice*, vol. 110, no. 11, pp. 104758–104767, 2021.

[21] M. Klauco, M. Kaluz, and M. Kvasnica, "Machine learning-based warm starting of active set methods in embedded model predictive control[J]," *Engineering Applications of Artificial Intelligence*, vol. 77, no. JAN, pp. 132–148, 2019.

[22] A. H. Mazinan and A. R. Khalaji, "A comparative study on applications of artificial intelligence-based multiple models predictive control schemes to a class of industrial complicated systems," *Energy Systems*, vol. 7, no. 2, pp. 237–269, 2016.

[23] X. Z. Xia and L. Cheng, "Adaptive Takagi-Sugeno fuzzy model and model predictive control of pneumatic artificial muscles [J]," *Science China Technological Sciences*, vol. 2021, no. 10, pp. 323–334, 2021.

[24] A. Owczarkowski and D. Horla, "A comparison of control strategies for 4DoF model of unmanned bicycle robot stabilised by inertial wheel[J]," *Advances in Intelligent Systems and Computing*, vol. 351, pp. 211–221, 2015.

Hindawi
Computational Intelligence and Neuroscience
Volume 2022, Article ID 5611456, 11 pages
https://doi.org/10.1155/2022/5611456

Research Article

A Multimodal Convolutional Neural Network Model for the Analysis of Music Genre on Children's Emotions Influence Intelligence

Wei Chen and Guobin Wu

Changchun Humanities and Sciences College, Changchun 130117, Jilin, China

Correspondence should be addressed to Guobin Wu; wuguobin@ccrw.edu.cn

Received 9 July 2022; Revised 6 August 2022; Accepted 9 August 2022; Published 29 August 2022

Academic Editor: Ning Cao

This paper designs a multimodal convolutional neural network model for the intelligent analysis of the influence of music genres on children's emotions by constructing a multimodal convolutional neural network model and profoundly analyzing the impact of music genres on children's feelings. Considering the diversity of music genre features in the audio power spectrogram, the Mel filtering method is used in the feature extraction stage to ensure the effective retention of the genre feature attributes of the audio signal by dimensional reduction of the Mel filtered signal, deepening the differences of the extracted features between different genres, and to reduce the input size and expand the model training scale in the model input stage, the audio power spectrogram obtained by feature extraction is cut the MSCN-LSTM consists of two modules: multiscale convolutional kernel convolutional neural network and long and short term memory network. The MSCNN network is used to extract the EEG signal features, the LSTM network is used to remove the temporal characteristics of the eye-movement signal, and the feature fusion is done by feature-level fusion. The multimodal signal has a higher emotion classification accuracy than the unimodal signal, and the average accuracy of emotion quadruple classification based on a 6-channel EEG signal, and children's multimodal signal reaches 97.94%. After pretraining with the MSD (Million Song Dataset) dataset in this paper, the model effect was further improved significantly. The accuracy of the Dense Inception network improved to 91.0% and 89.91% on the GTZAN dataset and ISMIR2004 dataset, respectively, proving that the Dense Inception network's effectiveness and advancedness of the Dense Inception network were demonstrated.

1. Introduction

Emotions are fundamental to our daily lives and play an essential role in rational decision-making, perception, healthcare, and human intelligence [1]. However, affective states are primarily overlooked, so the use of computers to analyze and process affective signals from sensors to identify a person's affective state has become a hot topic of contemporary research. There is no strict range calibration to differentiate musical works of different genres in terms of rhythmic and genre-specific expression features. However, other musical results of the same genre express their properties similarly. By capturing the similarity patterns unique to the same genre, it is possible to determine the genre attributes in musical expression. Such similarity patterns are universally applicable, allowing classification operations with more significant data sizes to be accomplished [2]. Traditional music genre classification methods are mainly achieved through manual or social annotation. Manual classification requires participants to have a certain degree of music expertise. Although it can ensure the classification effect, manual sort consumes much labor and time and is also costly. Although this classification mode saves costs, it is not easy to guarantee its classification effect [3].

One of the critical aspects of multimodal sentiment recognition is the fusion strategy. In machine learning, multimodal fusion usually employs feature, decision, and

hybrid fusion [4]. Although these methods can perform well in multimodal sentiment recognition, they are shallow fusion models that perform poorly in uniting and modeling multiple input features. Some new approaches have been recently investigated in the field of emotion recognition, which fuses the information of all available modalities directly with deep learning networks for feature extraction and optimization, which can achieve the fusion of multiple modalities into a single modality to enhance the stability of the recognition process [5]. A profound sentiment recognition model based on multimodal decomposition bilinear pooling is proposed. In this model, we first select the channels of EEG signals to reduce the interference caused by redundant channels. Then, convolutional neural networks extract each modality's convolutional features [6]. Finally, a multimodal decomposition bilinear pooling method is used to fuse and optimize the characteristics of each modality. The proposed fusion strategy allows all elements of each modality to interact to express each modality's complex internal relationships effectively. The model has an average accuracy of 93.22% on the DEAP dataset and 90.50% on the MAHNOB-HCI dataset, demonstrating that the model can improve multimodal sentiment recognition performance and significantly outperform existing techniques.

Neural networks have led the development of artificial intelligence in recent years, among which Convolutional Neural Networks (CNN) is a leader in image recognition, which can find out the features in complex images with high accuracy, improvement space, and sound processing efficiency in the face of high-dimensional data, and its architecture and each component element are still developing rapidly. On the other hand, there are many good deep algorithms available. On the other hand, many excellent deep learning frameworks are no longer challenging to build an easy-to-use deep learning system [7]. The development of computer hardware also gradually decreases the computational cost, making deep learning algorithm models such as convolutional neural networks closer to daily life [8]. This paper uses music genre classification and recognition as the research direction; we use short-time Fourier transform, Meier transforms, and constant Q transform to process one-dimensional audio files to generate spectrum and related data and use a convolutional neural network to automatically learn to extract acoustic features such as rhythm, pitch, and chord from images to build a music genre classification model. Applying deep learning techniques combined with EEG signals for multimodal emotion recognition methods has become a research trend [9]. This paper addresses the issues related to feature extraction and feature fusion in multimodal emotion recognition: on the one hand, we propose a layered fusion feature extraction to improve the recognition rate of emotion states; on the other hand, we use deep learning models to automatically extract and fuse multimodal signal features based on bilinear pooling of multimodal decomposition to achieve intelligent analysis of the influence of music genres on children's emotions.

2. Related Works

With its unique artistic elements of melody, rhythm, and harmony, music is significantly more complex than ordinary audio in terms of sound composition. The different ways of combining other artistic elements in a musical work while designing various musical works also produces different genres to which the musical works belong [10]. When enjoying a musical work, in addition to the singer's voice and the performer's song, there will be the presence of other combinations of elements such as natural sounds and frequencies outside the range of the human ear's vocal range. In the early days, when musical genre classification arose, humans could determine the genre affiliation of different music by analyzing it. However, the way humans understand and perceive music is challenging to construct through scientific methods, so it is not easy to implement automated music genre classification to allow computers to process it according to human determinations [11]. Since the introduction of automatic genre classification, extracting music's audio signal and processing it by machine learning models can effectively realize the automatic determination of music genre attribution [12]. The research on music genre classification has continued to focus on the extraction of features expressing musical attributes and attribute decisions based on the extracted features, among which the extraction of features describing musical details is the most significant [13]. In music genre classification, the study of feature engineering has been proposed since the 1990s. Gu et al. analyzed and verified the effectiveness of genre classification models based on sparse representation by introducing audio features such as spectral variance and combining them with principal component analysis methods for data dimensionality reduction [14]. Li et al. analyzed the GTZAN dataset and found some problems through targeted experimental design, which provided the relevant basis for the subsequent academic work of researchers [15].

Compared with unimodal sentiment analysis, multimodal sentiment analysis can overcome the noise effect carried by unimodal modality on the one hand and retain the practical information of each modality based on the complementary characteristics of different modalities on the other hand. Pan et al. combined LSTM and attention mechanism to extract the emotional relationship between contextual discourse in the video and introduced multimodal context-dependent information further to improve multimodal sentiment analysis systems [16]. Although all the research works in recent years have advanced the development of multimodal sentiment analysis to some extent, there are still problems such as limited characterization ability of unimodal sentiment features, insufficient robustness of fusion models, and no characterization ability of fused multimodal sentiment features. Based on the previous work, there is a need and feasibility to continue in-depth research [17]. Perceptual inconsistencies make it insufficient to predict highly personal variables' central (average)

effective category [18]. Gonzalez et al. propose that two multimodal affective computing tasks can be performed to deal with the subjectivity challenge: predicting each viewer's individualized affective perception and assigning multiple influential labels to each stimulus [19].

The child's inner emotional power system is the pillar of a balanced education. Moral, intellectual, emotional, aesthetic, creative, and physical elements constitute the spiritual world of each child [20]. When influencing children, it is necessary to educate them not only about knowledge but also about the development of their personality. Therefore, the product of emotions is an essential guarantee for the child's comprehensive development [21]. To improve the one-sidedness and deformation of educational development, we must pay attention to the role of emotions in the educational process. He believes that the key to the stimulation of children's "emotional motivation" lies in the teacher, who should have the art of mastering children's thinking, be highly perceptive and flexible, change teaching methods at any time, be good at encouraging children in teaching activities, care about children's life and health, their interests and happiness, and their complete spiritual life. Teachers and students should make spiritual contact, stimulate children's sense of self-esteem, self-confidence, and trust in teachers and make children enjoy learning success.

3. Multimodal Convolutional Neural Network Model Construction

Convolutional neural networks (CNNs) are now a common type of feedforward neural network and have an important position in the field of computer vision; the convolutional effect of CNNs can handle the correlation information in adjacent local receptive fields, and in computer vision problems, CNNs do not need to use one-to-one connections between all pixel units (i.e., like most neural networks), but instead use grouped local links. In addition, another feature of CNNs is the shared weights, where the convolutional kernel uses the same importance at all positions in the same layer, so the size of the network parameters can be effectively reduced [22]. As the depth of the network increases, the increase in the size of the sensory domain allows the network to represent more abstract features of the input. The convolutional layers focus on the edge portion of the object and then process the entire thing at a higher level in the hierarchy. Gradient descent combined with error backpropagation is essential for solving multilayer networks. A convolutional neural network is a multilayer structure that contains a feature extractor consisting of several convolutional layers, subsampling layers, and a connection layer. A neuron is connected to only some of its neighboring neurons in a convolutional layer. A convolutional layer in a CNN usually contains several feature maps, each consisting of several rectangularly arranged neurons that share the weights of the same feature map. The shared consequences are the convolution kernels. The convolution kernel is usually initialized in the form of a random fractional matrix, and the seed will learn reasonable weights during the training process of the network.

$$w_{(k-1)} = \sum \frac{\sqrt{\partial e\,(w-1)+\alpha}}{\sqrt{w_{k-1}-\partial w}}. \tag{1}$$

The gradient is calculated on the training set. The gradient descent equation can be easily solved using the chain rule of derivatives. An image input of $25 * 25$ pixels in size is an example of handwritten digit recognition. The information in the convolutional network is arranged as an array of $25 * 25$ design warp elements arranged in a rectangle, and their values correspond to the $25 * 25$ pixels used as the input. By convention, the input pixels are connected to the implicit layer neurons. However, CNN differs from fully connected layers in that CNN establishes connections.

$$\sigma = \sum \frac{w_1 - m_a - a_{x,y}}{\sqrt{b-l-m}} + \sqrt{k-m}, \tag{2}$$

where σ is the activation function of the neuron, possibly a sigmoid function, b is the shared bias, $w_{1,m}$ is the $5 * 5$ shared weight matrix. $a_{x,y}$ denotes the input excitation at positions xy. All neurons in the first hidden layer detect a particular feature at a different location in the input image. To justify this setup, assume that a certain weight and bias, i.e., a hidden neuron, can detect vertical edges. This ability may also be functional elsewhere in the image. Thus, the same feature detector can be applied anywhere in the picture. To perform image recognition, we need multiple feature mappings, and one convolutional layer corresponds to many different feature mappings, as shown in Figure 1.

This structural design detects many different features and can be performed from any position in the graph. One outstanding advantage of shared importance and tendencies in convolutional neural networks over fully connected neural networks is that it dramatically reduces the size of the network parameters. If every input pixel were fully connected to a neuron in the hidden layer, it would require roughly 40 times more parameters in this example. The training process of deep convolutional neural networks can also converge quickly. The mathematical equation describes the process of moving neurons from input to output.

$$y = \sum_{i=1} (n - x_i) \times w_{i+b} + \sqrt{f-1}. \tag{3}$$

A neuron's input can come from the input signal and the output of other neurons. This fully connected neural network has many neurons, divided into three layers, from top to bottom. The network structure has only one hidden layer; this neural network is called single hidden layer feedforward neural network. In deep learning, multiple hidden layers can be set up, and each hidden layer has a different number of neurons according to the actual situation to improve the learning ability.

For the hidden and output layers, the connection weight matrix of each layer with the previous layer and the output value of the neurons in the last layer are multiplied by the bias term of that layer to obtain the linear output [23]. Then the activation function of that layer is nonlinearly transformed to get the work of the neurons in that layer. The computational equation describes the process of the neuron

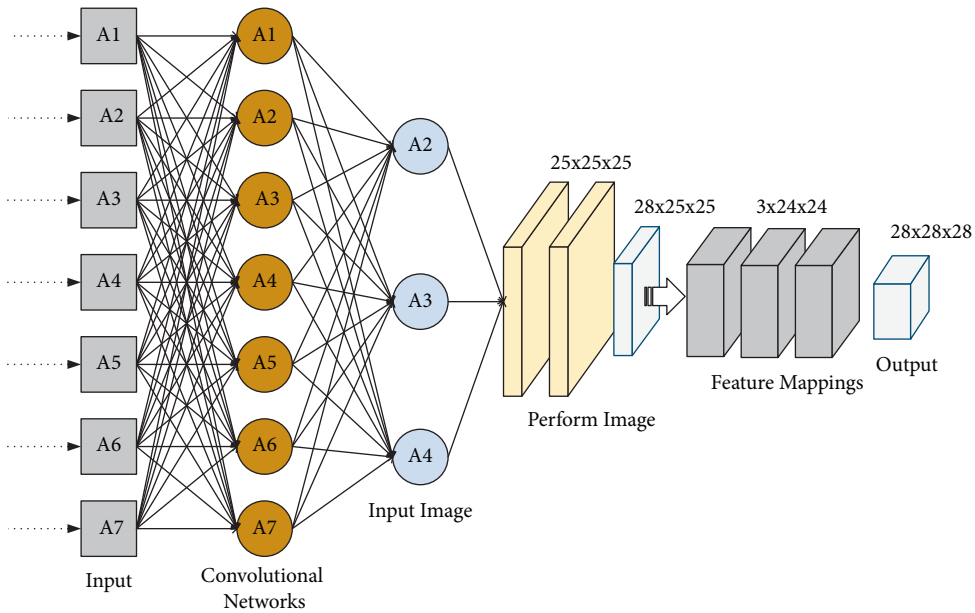


Figure 1: Convolutional layers corresponding to different feature mappings.

in each layer, from receiving the input to computing the output.

$$z^l = \int_{l=1} \frac{b^L - 1}{w^L + a^{L+1}}, \tag{4}$$

$$a^l = \sum \frac{f^L + z^L}{\sqrt{f^L - z^L}}. \tag{5}$$

Equation (4) is the linear output vector of the neuron at the first layer, calculated from the output vector of the neuron at the coating $L - 1a^{L-1}$, the connection weight matrix at the layer LW^L , and the bias term at the L layer b^L. In equation (5), a^L is the nonlinear output vector of the neuron at layer L obtained by passing the linear output of the neuron at layer Lz^L through the activation function $Lf^L(\cdot)$. Starting from the input layer, along with the input to output direction, as in the above process, the input vector, connected weight matrix, and bias term of each layer are subjected to a series of linear and activation operations and computed layer by layer backward until the target prediction result is obtained in the output layer, such a process is the forward propagation process. An LSTM network contains three gates: forget, input, and output. The forget gate is used to remove redundant information, while the input gate is responsible for updating the old cell state, and the output gate affects which part of the output of the updated cell state. Advantageous in sequence modeling problems with long-time memory. Simple to implement. Solve the problem of gradient disappearance and gradient explosion in the long sequence training process.

$$Z = \sum \tanh \frac{1 - b}{\sqrt{\theta x^T - zwh^{T-1} - B}},$$

$$z - l = \sum \tanh \frac{zwh^{T-1}}{\theta x^T - zwh^{T-1} + B}, \tag{6}$$

$$z^f = \sum \frac{Bf - 1}{\sqrt{\theta_f x^T - W_f h^{T-1} + Bf}}.$$

4. Design of an Intelligent Analysis Model of the Influence of Music Genres on Children's Emotions

The most significant difficulty in feature selection for music classification can be solved by deep learning models, which allow computers to automatically learn the pattern features needed to correctly classify music, making it possible to design effective models without relying on a deep knowledge of audio signal processing and music theory. Many studies follow image processing experience in designing convolutional neural networks, lacking an effective method for feature attributes in music audio spectrograms [24]. In this section, this paper proposes a residual module and a dual-attention fusion module for music feature extraction that combines residual the model in extracting high-dimensional abstract features of music by stacking residual modules and enhancing the directionality of feature extraction by dual-attention fusion modules, thus optimizing the training efficiency of the model and thus improving the effectiveness of music genre classification.

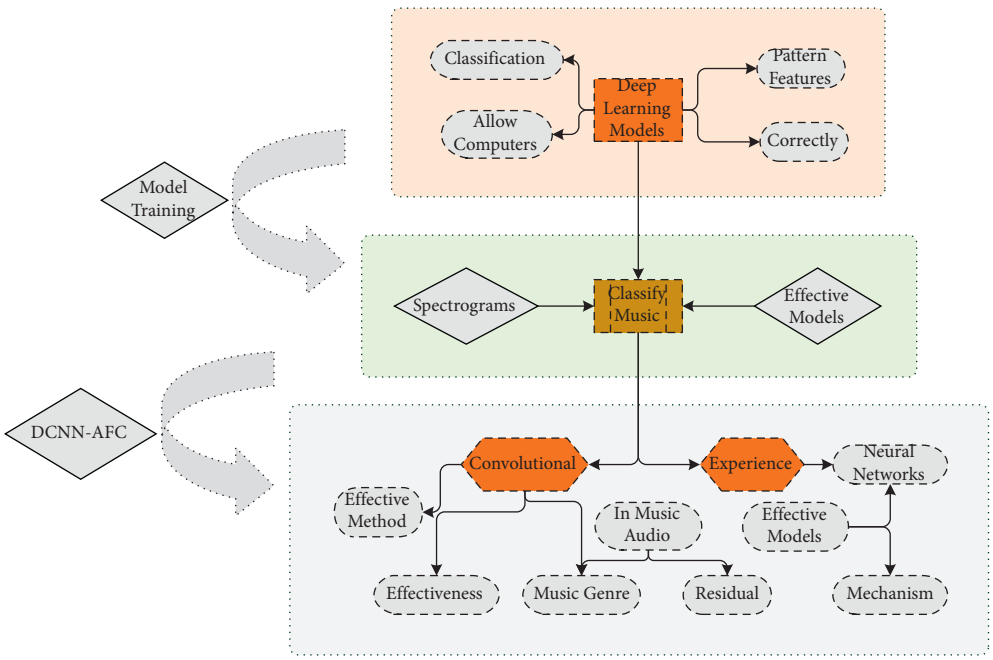


Figure 2: DCNN-AFC model flow.

The DCNN-AFC model performs Mayer filtering on the original audio signal, converts the filtered audio signal into an audio spectrogram, and then splits it into DCNN. Model training, the model, completes the training and validation of the corresponding music data on the training and validation sets as a batch for multiple iterations and outputs the model when it reaches the specified collection. The DCNN-AFC model flow is shown in Figure 2.

Given the ability of DCNN to extract features in the audio spectrogram and the performance of audio spectrogram classification, the residual structure, channel attention mechanism, and spatial attention mechanism are introduced to enhance the model's performance in extracting audio spectrogram features. The partial design of the network is improved to propose a deep convolutional neural network model consisting of an effective combination of the residual module, and dual-attention fusion module.

Six features are selected: chromatic frequency, root mean square of the spectrum, the spectral center of mass, acoustic spectrum roll-off coefficient, Mel frequency cepstral coefficient, and over-zero rate. Their mean values are calculated, and the mean values of 20 features of the Mel frequency cepstral coefficient are added to form a total combination of twenty-six features to describe the feature information of a piece of music. And the classification model is designed as a hybrid model containing an LSTM submodule and one-dimensional convolutional submodule according to the characteristics of feature combination, and the performance of traditional LSTM and LSTM containing attention mechanism in the model is compared. Through continuous exploration of acoustics, we have extracted acoustic features from various aspects, and digitized music can often remove hundreds of dimensional acoustic features. The audio spectrum of an audio signal is shown in Figure 3.

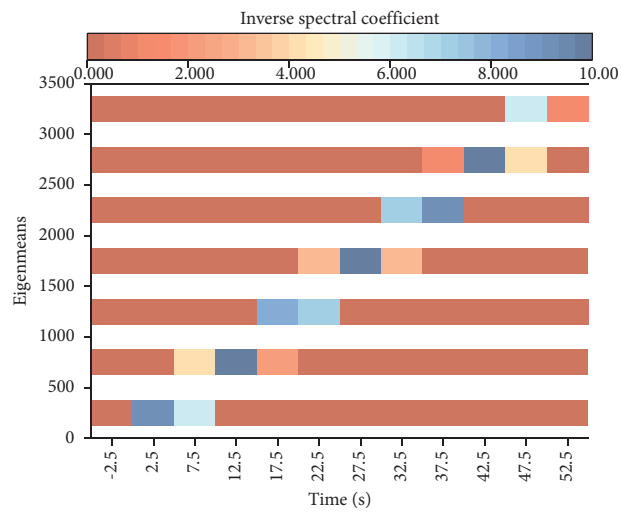


Figure 3: Audio spectrum of the audio signal.

Music teaching is not only for all children but should also be able to take special care of and accommodate children. This is not for music teachers to treat children differently, but to teach in detail so that children no longer feel alone but love and care from the inside, making them feel happy and successful in learning music. The content of music teaching should be appropriately integrated with children's lives, able to express children's hearts, pay attention to children's interests, attitudes and needs, strengthen the emotional experience, improve children's ability to externalize their emotions, and make children more adept at using music to regulate and express their feelings. Music teaching should be carried out in a relaxed, friendly, and pleasant environment, focusing on nurturing; the teaching language should be mainly encouraging, providing children with rich ways of

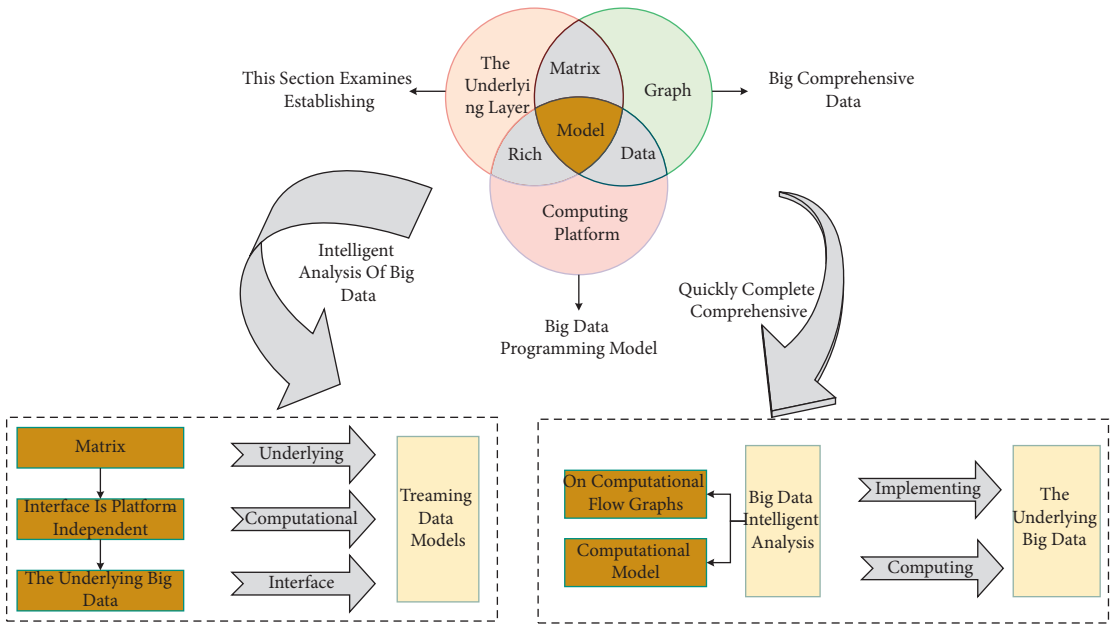


Figure 4: Data intelligence analytics programming model and programming framework.

participating in musical activities and opportunities for expression, cultivating their cheerful character, enhancing children's self-confidence, cultivating their strong will and courage to face, and overcome difficulties, and developing their optimistic attitude toward life and living. To realize big comprehensive data intelligent analysis, this section examines establishing a cross-platform unified big data programming model covering the table, matrix, tensor, graph, and streaming data models. It then further explores the construction of a programming method for big data intelligent analysis based on computational flow graphs. The big data brilliant analysis programming model and framework are shown in Figure 4.

At the same time, the upper-layer analytic algorithms and applications are implemented based on platform-independent high-level programming computational models and interfaces [25]. Through the cross-platform unified high-level abstract big data programming computing model and interface, the upper-layer data analysis algorithms and applications can be decoupled from the underlying distributed parallel computing platform, hiding the details of the underlying big data computing platform from the upper-layer data analysts, realizing the transparency of the underlying platform to the upper-layer algorithm design and application development, and thus significantly improving the ease of use of the big data processing system platform.

5. Analysis of Results

5.1. Analysis of Multimodal Convolutional Neural Network Models. To verify the effectiveness of the multimodal convolutional neural network proposed in this chapter, 62 channels of EEG data and eye-movement data of the subject children in the SEED-IV dataset were used as experimental data. In the unimodal signal-based emotion recognition experiments, the MSCNN network was used to classify the 62-channel EEG data for emotion. Single-modality image alignment refers to the floating two images acquired with the same imaging device. It is mainly applied to the alignment between different MRI-weighted images and the alignment of image sequences, etc. Multimodal image alignment refers to the floating of two images from other imaging devices. The LSTM network was used to realize the emotion classification based on the eye-movement signal; the MSCNN-LSTM model was used for the multimodal signal emotion classification experiments. Three experimental data for each child at different periods were conducted separately, with a total of 45 experiments. Each sample size of EEG data was 5×62, each sample size of eye-movement data was 1×31, and the total sample size of session 3 was 812. First, the samples were preprocessed. After normalization with zero-mean normalization, the training set and test set are divided according to the ratio of $8:2$, and the number of experimental iterations is set to 300. In terms of emotion classification, the eye-movement signal contains less obvious emotion recognition features than the EEG signal and multimodal signal, and the EEG signal and multimodal signal are more suitable as data for emotion classification studies; second, the experimental results of different signals all that the accuracy rates of the first experiment and the second experiment are generally higher than those of the third experiment, which may be due to the acquisition environment of the third experiment or other factors may have influenced the data acquisition in the third experiment. The comparison of the classification results of unimodal and multimodal signals is shown in Figure 5.

The advantages of multimodal signals relative more intuitively to unimodal signals, the average accuracy of subjects' emotion classification in three experiments under three different signals and the average classification accuracy

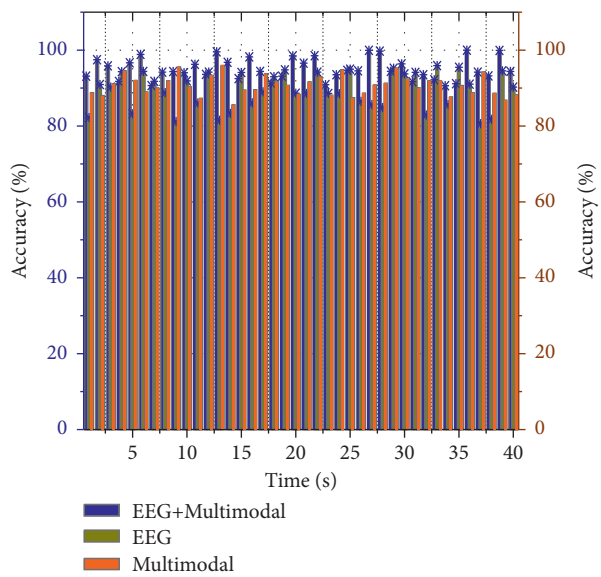


Figure 5: Comparison of classification results between unimodal and multimodal signals.

in 45 experiments under three separate signals, where the three different signals were 62-channel EEG signal, eye-movement signal and 62-channel EEG signal fused with eye-movement signal in the multimodal signal. From the experimental results, in the first eight subjects and the last two subjects, the classification accuracy of the emotion of multimodal signal is higher than that of the unimodal signal. In contrast, the classification results of subject 9, subject 10, subject 11, subject 12, and subject 13 showed that the classification accuracy of the emotion of multimodal signal is lower than that of unimodal signal, which means that multimodal convolutional neural network has subject variability from the side; the classification accuracy based on the multimodal signal. The average accuracy of 45 experiments based on multimodal signals, EEG signals, and eye-movement signals are 95.2139%, 96.2315%, and 93.0641%, respectively. However, multimodal signals have individual differences; for most subjects, the accuracy of multimodal-based emotion classification is higher than that of unimodal signals, which verifies the effectiveness of multimodal signals and verifies the efficacy of multimodal convolutional neural networks for emotion classification of multimodal signals.

When the parameters in the DEAP dataset were optimal, the average accuracy of the wakeup dimension was 83.28%, and the valence dimension was 84.71%. When the parameters in the MAHNOB-HCI dataset are optimal, the accuracy of the wake dimension is 88.28%, and the valence dimension is 89.00%. It can be concluded that the model performs better in the valence dimension than in the wake dimension using the same dataset with optimized parameters, which indicates that the model proposed in this chapter is more active in the valence dimension [26]. When the classifier parameters are optimal, the accuracy of the model presented in this chapter applied to the MAHNOB-HCI dataset is higher than that of the model applied to the DEAP dataset. The experimental results obtained using a layered fusion convolutional neural network to extract features are significantly better than those obtained using only the convolutional neural network alone. The reason is that although convolutional neural networks can automatically learn and extract multilayer feature representations from the original data, the neural network does not focus on one feature per neuron; instead, a group of neurons focuses on one part. The variation curves of classifier accuracy with parameter settings are shown in Figure 6.

5.2. Intelligent Analysis of the Impact of Music Genres on Children's Emotions Realized. The music genre includes many sensory systems, such as visual, auditory, and tactile. The first sense that needs to be used is visible, which transmits the notes on the music score to the brain's nervous system. This is followed by the sense of hearing, which is the inner sense of hearing, a mental activity based on the perception of music and the purpose of shaping musical images; that is to say, the inner meaning of hearing at this time is a kind of auditory imagination of the expected music and the effect of the music style. Next, the natural sound is put out through kinesthetic and tactile senses, and only then does the honest auditory feedback appear. Therefore, the coordination of the sensory and motor systems is the most critical psychological basis for improving the coordination of musical movements. First, sight-reading skills can be enhanced through sight-singing and ear-training exercises, which allow children to connect music and sound and vision and hearing. Second, children can perform the sighted scores so that they can establish connections between the scores and performance, visual and kinesthetic senses, and strengthen their inner auditory and tactile senses so that the sensory system and the motor system will gradually coordinate, and the coordination of movements in music performance will improve progressively. Of course, it is also possible to use the Dalcroze somatic rhythm teaching method to enhance the connection between children's auditory and kinesthetic senses so that children can feel the relationship between music and movement and improve the coordination between hearing and kinesthetic senses. To illustrate the effect of the number of iterations on model training, experiments were conducted on the validation set with a learning rate of 0.01 and 37,000 iterations. The genre classification accuracy plot for 37000 iterations is shown in Figure 7.

The audio fragment-based prediction method can make the model more flexible to cope with different time-length music samples, which is very important in practical use. According to the previous settings, 66560 sampling points at 22050 Hz audio are used as an audio clip. The audio clips are selected by randomly selecting the starting point and sequentially intercepting the audio clips. There is a 50% overlap between the audio clips, which avoids the loss of information partially located near the audio clip interception point across the two audio clips due to nonoverlapping cutting. The model's performance was tested in the prediction phase using different numbers of segments from 1 to 18, and the classification accuracy in the prediction phase

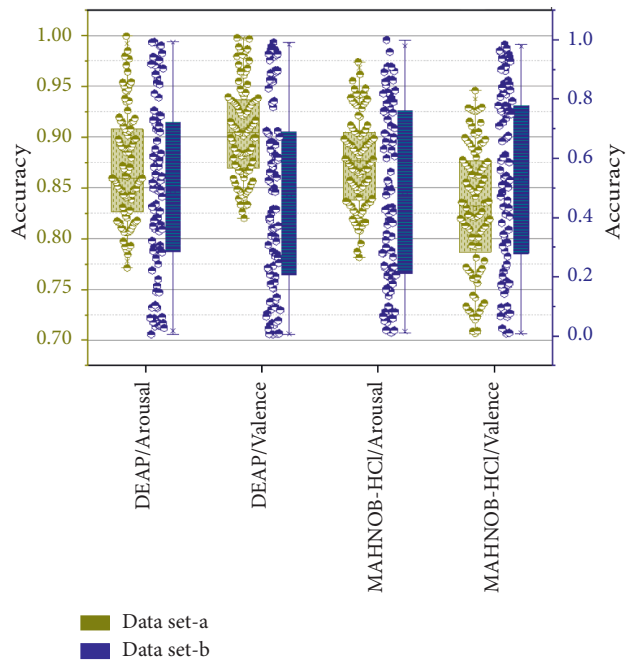


Figure 6: Variation curve of classifier accuracy with parameter settings.

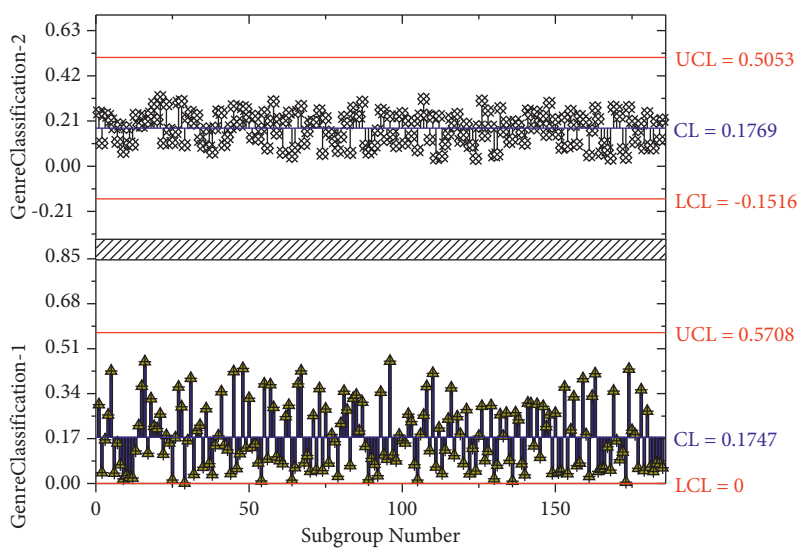


Figure 7: 37000 iterations of genre classification accuracy graph.

using other numbers of features is shown in Figure 8. The accuracy of the model increases rapidly as the number of audio segments increases when the model uses a small number of pieces for prediction; from using one audio component for prediction to using ten audio details for prediction, the accuracy improves by about 3.87% after smoothing (the accuracy improves by about 5.1% before filing). As the number of audio clips increases, the model's accuracy increases slower, from using ten audio clips for prediction to 18 audio clips, with an improvement of about 8% after smoothing (and about 0.4% before filing). Theoretically, increasing the number of audio clips used for the prediction can improve accuracy. Still, as the number of audio clips used increases, each audio clip will bring less and less additional valid information and will contribute less and less to the accuracy improvement. Eventually, the accuracy will converge to a stable interval, so it is not meant to increase the number of audio clips used blindly. To ensure a high accuracy rate while avoiding a meaningless increase in the number of audio clips and combining the characteristics of the GTZAN dataset, 18 is finally chosen as the number of audio clips used in the prediction stage of the model.

The intelligent analysis dataset was divided into equal proportions according to the category labels with the ratio of 80%, 10%, and 10% for the training set, validation set, and test set, respectively. In total, 200 rounds were conducted on the training set using a learning rate of 0.001, and the learning process was supervised using the validation set. The

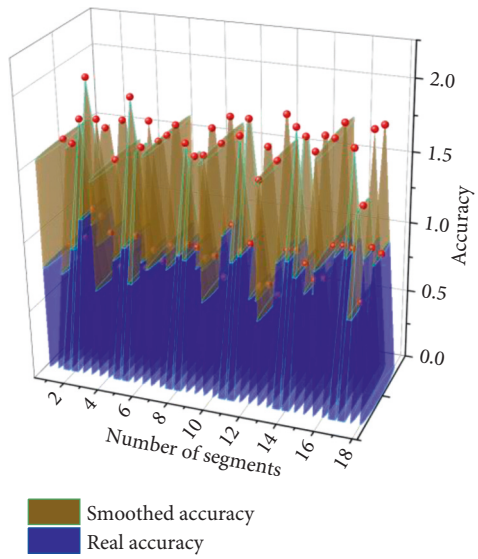


Figure 8: Classification accuracy when using a different number of segments in the prediction stage.

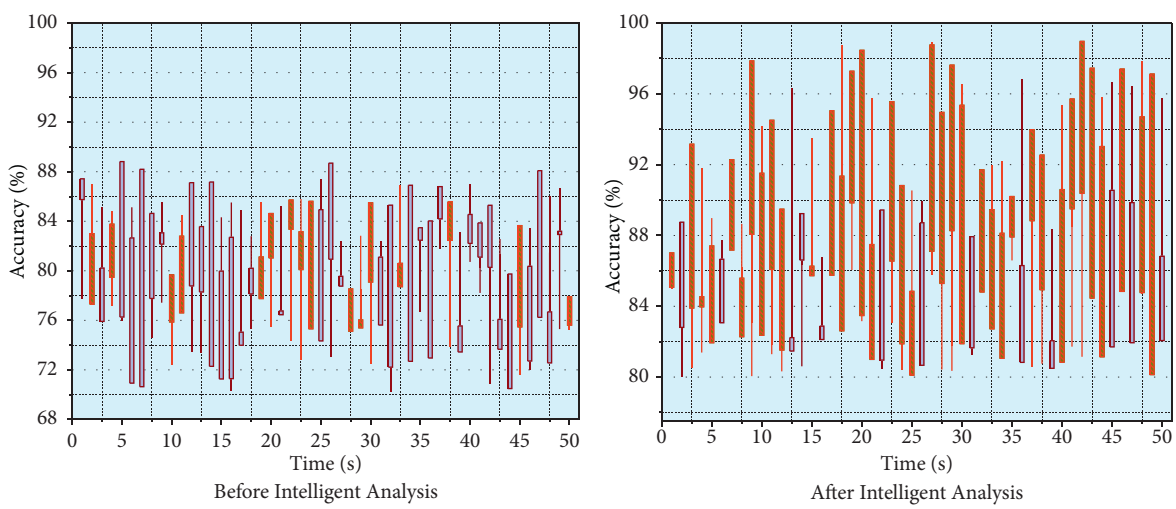


Figure 9: Accuracy comparison before and after using intelligent analysis dataset.

training was terminated early when the accuracy performance on the validation set did not improve for more than 20 rounds. A total of about 219 hours of training was conducted, and the activity was terminated early after 8 games of movement without any further improvement in accuracy on the validation set. The pretrained model was used as the base model for migration learning. The last fully connected layer was replaced according to the number of categories in the GTZAN, ISMIR2004 dataset, and retrained as a starting point. After the same training process with the exact details described in the previous section, the model classification accuracy was improved substantially. The model's 10-fold cross-validation classification accuracy was improved to 91.0% and 89.91% on the GTZAN and ISMIR2004 datasets, respectively, as shown in Figure 9. Music genres begin with auditory sensations and express auditory content by triggering emotions through auditory senses. Thus, music genres strengthen the connection between aural and kinesthetic feelings and music and emotions. It is believed that the problem of children's coordination of movements in musical genres is caused by the lack of close coordination between the neurological system of the child's brain and the parts of the body and the separation of the child's mind and body from the music felt in musical genres. In music genres, children can feel the fusion of emotions and music, experience muscle tension and relaxation, strengthen the connection between the brain's nervous system and body parts, and improve the coordination of body parts, which are precisely the causes of children's emotions. Therefore, it is believed that the use of the music genre has a vital role in promoting the coordination of movements in the emotional impact of children.

6. Conclusion

As a critical technology in the development of computer intelligence, emotion recognition has significant application research value for guiding machines to serve human beings better. However, because human emotions are complex and diverse, it is challenging to accurately grasp specific emotional attributes and categories. Numerous challenges in this field still need to be tackled and solved by researchers. In this paper, we design a deep convolutional neural network-based music genre classification (DCNN-AFC) model to determine the genre attributes of musical works by processing the audio spectrogram of audio signals through dimensionality-reduced Mel filtering and capturing the gene expression features of music files. In this paper, a sub-dataset generated from the MSD dataset is used to pretrain the Dense Inception network, and then the pretrained model is used to train on the GTZAN dataset and the ISMIR2004 dataset. This pretraining step improves the stability of the model and the convergence during training while significantly improving the classification accuracy of multiple datasets. Comparing before and after pretraining using the MSD dataset, the accuracy of the Dense Inception network improved from 88.7% and 87.68% to 91.0% and 89.91% on the GTZAN dataset and the ISMIR2004 dataset, respectively. The use of computer-based interactive teaching not only updates the traditional teaching methods and enriches the teaching contents but also makes the abstract music realistic and lively, which significantly broadens children's musical horizons, enlivens their musical thinking, stimulates their enthusiasm for learning music, and increases the efficiency of analyzing children's emotional impact. Although the multimodal emotion recognition system designed and developed in this paper functionally realizes the recognition and visualization of students' learning emotions, the form of visualization presentation is relatively single, and consideration can be given to increasing the diversity of visualization presentation content. Since the system is operated based on separate computer local storage, the management of student data and course data is cumbersome, after which the use of a database for user data and course data should be considered. Control, to simplify the operation of users, while the effect of providing feedback to teacher users should be analyzed and designed according to the actual user needs, and more feasible functions should be added based on giving visual input to improve the practicality of the system.

Data Availability

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

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work was supported by Changchun Humanities and Sciences College.

References

[1] A. Dhillon and G. K. Verma, "Convolutional neural network: a review of models, methodologies and applications to object detection," *Progress in Artificial Intelligence*, vol. 9, no. 2, pp. 85–112, 2020.

[2] J. V. Tembhurne and T. Diwan, "Sentiment analysis in textual, visual and multimodal inputs using recurrent neural networks," *Multimedia Tools and Applications*, vol. 80, no. 5, pp. 6871–6910, 2021.

[3] J. Ramírez and M. J. Flores, "Machine learning for music genre: multifaceted review and experimentation with audioset," *Journal of Intelligent Information Systems*, vol. 55, no. 3, pp. 469–499, 2020.

[4] K. Somandepalli, T. Guha, V. R. Martinez, N. Kumar, H. Adam, and S. Narayanan, "Computational media intelligence: human-centered machine analysis of media," *in Proceedings of the IEEE*, vol. 109, no. 5, pp. 891–910, 2021.

[5] F. Shen, G. Dai, G. Lin, J. Zhang, W. Kong, and H. Zeng, "EEG-based emotion recognition using 4D convolutional recurrent neural network," *Cognitive Neurodynamics*, vol. 14, no. 6, pp. 815–828, 2020.

[6] L. Chen, W. Su, M. Wu, W. Pedrycz, and K. Hirota, "A fuzzy deep neural network with sparse autoencoder for emotional intention understanding in human–robot interaction," *IEEE Transactions on Fuzzy Systems*, vol. 28, no. 7, pp. 1252–1264, 2020.

[7] V. Gupta, S. Juyal, and Y. C. Hu, "Understanding human emotions through speech spectrograms using deep neural network," *The Journal of Supercomputing*, vol. 78, no. 5, pp. 6944–6973, 2022.

[8] A. Yadav and D. K. Vishwakarma, "Sentiment analysis using deep learning architectures: a review," *Artificial Intelligence Review*, vol. 53, no. 6, pp. 4335–4385, 2020.

[9] K. Zhang, Y. Li, J. Wang, E. Cambria, and X. Li, "Real-time video emotion recognition based on reinforcement learning and domain knowledge," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 32, no. 3, pp. 1034–1047, 2022.

[10] G. Tu, Y. Fu, B. Li, J. Gao, Y. G. Jiang, and X. Xue, "A multi-task neural approach for emotion attribution, classification, and summarization," *IEEE Transactions on Multimedia*, vol. 22, no. 1, pp. 148–159, 2020.

[11] O. Agbo-Ajala and S. Viriri, "Deep learning approach for facial age classification: a survey of the state-of-the-art," *Artificial Intelligence Review*, vol. 54, no. 1, pp. 179–213, 2021.

[12] S. Amiriparian, N. Cummins, M. Gerczuk, S. Pugachevskiy, S. Ottl, and B. Schuller, "Are you playing a shooter again?!" deep representation learning for audio-based video game genre recognition," *IEEE Transactions on Games*, vol. 12, no. 2, pp. 145–154, 2020.

[13] R. Lotfian and C. Busso, "Curriculum learning for speech emotion recognition from crowdsourced labels," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 27, no. 4, pp. 815–826, 2019.

[14] X. Gu, Z. Cao, A. Jolfaei et al., "EEG-based brain-computer interfaces (BCIs): a survey of recent studies on signal sensing technologies and computational intelligence approaches and their applications," *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, vol. 18, no. 5, pp. 1645–1666, 2021.

[15] J. Li, S. Qiu, C. Du, Y. Wang, and H. He, "Domain adaptation for EEG emotion recognition based on latent representation

similarity," *IEEE Transactions on Cognitive and Developmental Systems*, vol. 12, no. 2, pp. 344–353, 2020.

[16] L. Pan, S. Wang, Z. Yin, and A. Song, "Recognition of human inner emotion based on two-stage FCA-ReliefF feature optimization," *Information Technology and Control*, vol. 51, no. 1, pp. 32–47, 2022.

[17] X. Jiang, S. C. Satapathy, L. Yang, S. H. Wang, and Y. D. Zhang, "A survey on artificial intelligence in Chinese sign language recognition," *Arabian Journal for Science and Engineering*, vol. 45, no. 12, pp. 9859–9894, 2020.

[18] L. Xu, X. Wen, J. Shi et al., "Effects of individual factors on perceived emotion and felt emotion of music: based on machine learning methods," *Psychology of Music*, vol. 49, no. 5, pp. 1069–1087, 2021.

[19] H. A. Gonzalez, R. George, S. Muzaffar et al., "Hardware acceleration of EEG-based emotion classification systems: a comprehensive survey," *IEEE Transactions on Biomedical Circuits and Systems*, vol. 15, no. 3, pp. 412–442, 2021.

[20] C. Bian, Y. Zhang, F. Yang, W. Bi, and W. Lu, "Spontaneous facial expression database for academic emotion inference in online learning," *IET Computer Vision*, vol. 13, no. 3, pp. 329–337, 2019.

[21] D. Chaudhary, N. P. Singh, and S. Singh, "Automatic music emotion classification using hashtag graph," *International Journal of Speech Technology*, vol. 22, no. 3, pp. 551–561, 2019.

[22] B. Sonawane and P. Sharma, "Review of automated emotion-based quantification of facial expression in Parkinson's patients," *The Visual Computer*, vol. 37, no. 5, pp. 1151–1167, 2021.

[23] I. Shahin, A. B. Nassif, N. Nemmour, A. Elnagar, A. Alhudhaif, and K. Polat, "Novel hybrid DNN approaches for speaker verification in emotional and stressful talking environments," *Neural Computing & Applications*, vol. 33, no. 23, Article ID 16033, 2021.

[24] S. S. Panicker and P. Gayathri, "A survey of machine learning techniques in physiology based mental stress detection systems," *Biocybernetics and Biomedical Engineering*, vol. 39, no. 2, pp. 444–469, 2019.

[25] S. M. Rajendram and T. T. Mirnalinee, "Contextual emotion detection on text using Gaussian process and tree based classifiers," *Intelligent Data Analysis*, vol. 26, no. 1, pp. 119–132, 2022.

[26] J. Liu, S. Snodgrass, A. Khalifa, S. Risi, G. N. Yannakakis, and J. Togelius, "Deep learning for procedural content generation," *Neural Computing & Applications*, vol. 33, no. 1, pp. 19–37, 2021.

Hindawi
Computational Intelligence and Neuroscience
Volume 2022, Article ID 2424279, 11 pages
https://doi.org/10.1155/2022/2424279

Research Article

Adaptive Biological Neural Network Control and Virtual Realization for Engineering Manipulator

Hao Guo,[1] **Hongyang Liu**,[2] **Dashuai Zhou**,[1] **and Yao He**[1]

[1]*Departmentof Mechanical Manufacturing, School of Mining and Mechanical Engineering, Liupanshui Normal University, Liupanshui City, Guizhou Province 553004, China*
[2]*Department of Mining Engineering, School of Mining and Mechanical Engineering, Liupanshui Normal University, Liupanshui City, Guizhou Province 553004, China*

Correspondence should be addressed to Hao Guo; hg229@njit.edu

Received 14 July 2022; Revised 15 August 2022; Accepted 18 August 2022; Published 29 August 2022

Academic Editor: Ning Cao

By analyzing the feasibility of the digital twin technology in the assembly of construction machinery, the assembly process of the construction manipulator in the engineering environment is discussed. According to the application criteria and modeling requirements of digital twin, the overall framework of digital twin engineering manipulator assembly modeling and simulation is constructed from three aspects: model layer, data layer, and application layer. According to the operation task characteristics of space engineering manipulator, the feasibility of the control method based on joint angular velocity is analyzed, and the task environment of space engineering manipulator based on Markov model is defined. Aiming at the application of the algorithm in the control task of the space engineering manipulator, a reward function with the addition of the angular velocity soft bound term is designed, which improves the strategy optimization process of the algorithm and obtains a better control effect of the engineering manipulator. The motion trajectory of the end of the engineering manipulator is directly given on the simulation platform, and the expected motion of each joint of the engineering manipulator is calculated through the kinematics of the engineering manipulator. It can be seen from the simulation results that the controllers designed in this study can achieve ideal control effects. With the help of Baxter robot platform, the control algorithm designed in this study is applied to the actual engineering manipulator control, and the effectiveness of the control algorithm is further proved by the actual control effect.

1. Introduction

As the most commonly used robot in the industry, the construction manipulator is mostly controlled by point-to-point control in the early stage [1]. This control scheme is suitable for scenarios with low precision requirements. With the increase of production technology, industrial manufacturing puts forward higher requirements for the control accuracy of the construction manipulator, and at the same time requires the end of the construction manipulator to track the given reference trajectory motion [2]. On the other hand, traditional engineering manipulators are generally made of rigid materials with large volume and mass, which cannot complete high-precision work tasks, and are often used for mechanical work such as handling or assembly. With the advent of the era of intelligent manufacturing, the work tasks of the construction manipulator have become refined, and industrial production has put forward higher requirements for the precision of the construction manipulator [3]. At the same time, the adaptability of the production line to different work tasks is also increasing, and the scenarios in which humans and robots cooperate to complete production tasks are gradually increasing [4].

In order to improve the control accuracy and ensure the safety of human-computer interaction, a new type of engineering manipulator system, the flexible engineering manipulator system, is proposed. Flexible engineering manipulators are divided into flexible link engineering manipulators and flexible joint engineering manipulators

[5]. The flexible link engineering manipulator is that the link of the engineering manipulator is composed of elastic materials, and the connecting rod itself has the characteristics of flexibility; the flexible joint engineering manipulator is that the connecting rod of the engineering manipulator is still composed of rigid materials but exists at the joints of the engineering manipulator. The spring device makes its joint flexible. This study will take the flexible joint engineering manipulator as the research object. Compared with the rigid construction manipulator system, the flexible joint construction manipulator has the advantages of light structure, high control precision, high load-to-weight ratio, and quick response. At the same time, due to the elastic brakes installed at the joints, it has good flexibility [6]. When encountering obstacles during the movement, its contact force will be much smaller than that of the rigid engineering robot arm, so it can effectively protect the operator [7].

To realize the precise control of the engineering manipulator, it is necessary to model the system accurately and then carry out state feedback control through modern control theory [8]. However, in practical applications, the kinematics and dynamic models of engineering manipulators inevitably have uncertainties, and it is difficult to achieve accurate modeling. For the problem of uncertain items in the system, there is usually a scheme to identify the unknown nonlinearity of the system by using fuzzy logic or neural network. However, it is a cumbersome process to train the neural network and adjust the parameters. At the same time, in actual production, it is very common for the engineering manipulator to perform the same or similar tasks repeatedly. How to save the knowledge of neural network identification system dynamics, realize the learning of system dynamics, and avoid repeated training of neural network is a research direction of great theoretical significance. Because of the limitations of sensor deployment and the influence of external interference, we often cannot obtain all the state variables of the system. Therefore, it is of great theoretical and practical value to study the controller design of the flexible joint engineering manipulator system whose model contains unknown dynamics and unobtainable state quantities.

This study expounds on the main problems existing in the design stage of the current engineering manipulator assembly process and analyzes the necessity of introducing digital twin technology into traditional assembly. By analyzing the application requirements of digital twin assembly, the main process of construction manipulator assembly under the background of digital twin is planned, and the construction of the construction manipulator assembly process based on digital twin is completed from the model layer, data layer, and application layer. This section introduces a deep deterministic policy gradient algorithm for continuous motion control in the control of space engineering manipulators for the multi-degree-of-freedom control system of space engineering manipulators. The method of adding the angular velocity soft bound to the reward function effectively solves the problem of neural network divergence. The simulation results show that the engineering manipulator can be stably controlled at the target point. The engineering manipulator controller designed in this study is tested on the engineering manipulator platform with two degrees of freedom as an example. Through the engineering manipulator platform, from the ideal system model and state to the unfavorable application conditions of the unknown system model and state, it is proved that the control algorithm proposed in this study can achieve the ideal control effect. Moreover, the adaptive control based on the RBF neural network can effectively fit the unknown model under the condition of unknown system model and achieve an ideal control effect. Finally, the Baxter robot arm system is used to verify the two adaptive neural network control algorithms designed in this study.

2. Related Work

For the control of engineering manipulators, a local linearization method is usually used to linearize the nonlinear part of the dynamics of the engineering manipulator near the target trajectory [9]. However, due to the strong time-varying, nonlinear, and strong coupling characteristics of engineering manipulators, local linearization cannot guarantee the global stability of the system. In order to solve this problem, some scholars have proposed a feedback linearization method [10]. This method mainly uses differential geometry and spatial coordinate transformation to make the input and state or input and output of the nonlinear system approximately satisfy the linear relationship and then use mature linearity. The system control method makes the system satisfy a certain robustness. Related scholars discussed the problem of robust tracking control of rigid engineering manipulators with uncertain dynamics using nominal feedback controllers and variable structural compensators [11]. The results show that the method can eliminate the influence of large system uncertainty and ensure the asymptotic convergence of the output tracking error. The researchers further, using a multiloop version of the small gain set, can obtain robust trajectory tracking under the assumption that the deviation of the model from the real system satisfies some norm inequalities [12]. Related scholars use gap metric analysis to derive the robustness and performance margins of feedback linearized controllers [13]. Unlike previous stability analyses, it incorporates the case of outputting nonstructural uncertainties and derives general stability conditions that can be applied to stable and unstable systems.

Related scholars use the specific structure of engineering manipulator dynamics to develop a simple global convergence adaptive controller, design PD feedback part and full dynamic feedforward compensation part, perform online estimation of unknown manipulator and payload parameters, this method is simple to calculate, the joint accelerations are not known, and there is no need to estimate the inverse of the inertia matrix [14]. The researchers designed a compensatory control rate and a nonlinear filter feedback term to obtain a globally progressively stable tracking effect [15]. Related scholars have proposed a robust control method for n-link engineering manipulators with uncertain upper bounds [16]. This method does not need to identify all

the physical parameters of the engineering manipulator but only needs to estimate several parameters of the upper bound function. For multilink engineering manipulators, this method is much less computationally intensive. However, the discontinuous control amount caused by this method increases, and it is difficult to suppress disturbance. Further, related scholars have introduced an estimation rate of exponential changes such as random parameters and tracking error, which further enhances the robustness of the system [17].

Relevant scholars pointed out that the inversion control algorithm is a control algorithm for high-order complex nonlinear systems [18]. It combines the design of the controller with the selection of the Lyapunov function and divides the high-order system into several low-order subsystems in series according to the order of the system. The inversion control algorithm reverses step by step for each low-order subsystem in series and finally obtains the output of the overall controller.

The work of the restricted engineering manipulator is constrained by the environment, and its position, speed, and other states need to be restricted, which puts forward high requirements for the control design of the engineering manipulator. In recent years, a large number of scholars have devoted themselves to the control research of nonlinear systems with constraints and have achieved rich research results. At present, there are two common methods to solve the constraint problem, one is the obstacle Lyapunov function, and the other is the method based on function transformation [19].

The method based on function transformation adopts a class of nonlinear functions to directly transform restricted objects into equivalent unrestricted objects and then performs control design for the unrestricted objects. This method does not need to indirectly realize the constraint effect through the limited error but directly converts the limited physical quantity, so the control scheme is less conservative. Related scholars use the method based on function transformation to transform the tracking error of the system, so that the tracking error can reach the preset transient performance [20]. Subsequently, the method was further extended to solve the constrained tracking control problem of a class of strict feedback systems by combining different control techniques such as adaptive control, dynamic surface control, and neural network control.

3. Methods

3.1. Analysis of Engineering Manipulator Assembly Process Design Problems. In the field of construction machinery product assembly, with the gradual application of intelligent manufacturing theory and computer-aided process design, the digitization and visualization of product assembly process have higher requirements. As an important working device of large-scale construction machinery, the construction manipulator contains a variety of parts and connectors, with complex structure and various forms, and has extremely high requirements for design accuracy. The primary task of the assembly process design work is to meet the assembly quality of the product and then save the assembly cost and shorten the assembly cycle as much as possible through system planning and deployment. As the most intuitive quality evaluation index of assembly process documents, assembly accuracy is not only related to the manufacturing accuracy of parts but also affects the overall economy of the assembly system. At this stage, most manufacturing companies still have many problems in the formulation of assembly process regulations. For the field of engineering manipulator assembly, it is mainly reflected as follows:

(1) In the assembly process of the engineering manipulator, the traditional two-dimensional assembly design method lacks the real-time acquisition and processing of dynamic data, the synchronization between the assembly process and the data update is poor, and there are also many shortcomings in theoretical methods. Accurate prediction and judgment of failure phenomena and failure causes are realized, and the system flexibility and real-time performance are poor.

(2) The construction manipulator is assembled as a moving component. Compared with the conventional assembly, the mechanism assembly has the following characteristics:

①Affected by the motion characteristics and force characteristics of the construction manipulator during operation, a certain deviation will accumulate between the components, so a certain assembly gap needs to be reserved at the assembly connection to avoid affecting the subsequent work stability. ②Due to the influence of the manufacturing accuracy of the parts during the matching process of the construction manipulator, there may be a certain deviation in the actual assembly state; that is, there is a matching gap. In addition, the parts that cooperate with each other will generate various deviations and change directions during the movement process, forming different deviation transmission paths and making the accuracy prediction of the mechanism more difficult.

(3) There is a certain backwardness in the assembly of traditional engineering robotic arms in terms of assembly tools and assembly methods, resulting in an increase in the demand for manpower and material resources in the assembly process and poor economy. At the same time, the automation of assembly equipment is low, resulting in low production efficiency and insufficient production capacity.

3.2. Assembly Process of Construction Manipulator in Engineering Environment. With the rapid development of digital manufacturing technology, all kinds of intelligent equipment and virtual manufacturing technology are gradually applied to the field of traditional construction machinery assembly. The traditional manufacturing model has undergone a new transformation, forming a human-machine

interactive collaborative analysis and decision-making integrated system. As an important new technical means in the field of intelligent manufacturing, digital twin is mainly used to realize the interconnection of physical space and digital space by relying on models, data, and sensors and digitally define and analyze the characteristics and activities of physical entities in virtual models. The digital twin model mainly consists of three parts, including digital space, physical space, and the correlation mechanism between the two. Combined with the actual assembly conditions of the production site and a large amount of data information in the physical manufacturing process, the dynamics is realized through the information interface between the two parts.

The role of the basic elements of physical space construction in digital twin assembly is to provide model and process support for the construction of digital space and provide mutual feedback with digital space to achieve program optimization. In the physical space, it is necessary to clarify the implementation steps of the assembly process and the tool requirements of the assembly site. Through the assembly sequence planning and assembly process planning of the construction robot arm, the actual operation is strictly carried out according to the assembly requirements, and other smart labels can be configured to facilitate the recording and follow-up of information in the digital space.

As a real mapping of physical space, digital space has roughly the same composition and structure as physical space. Virtual production lines, machining layouts, and assembly processes can be completed in third-party modeling and simulation software, including models of geometry, behavior, and rules, and related simulation, optimization, and analysis activities. The data information in the digital space is generally connected to the service platform through multiple external interfaces, forming a multifaceted and multidomain application of data integration and fusion. Figure 1 shows the overall idea of digital twin assembly process control.

After reasonable collection and processing, the assembly process planning of the digital space is guided through the information interface. The engineering manipulator assembly based on digital twin mainly realizes the continuous iteration and optimization of assembly model design and assembly process parameters in the process of modeling and simulation.

Through dynamic data acquisition and modeling and simulation of the assembly process, the quality and accuracy of the model are affected. The assembly process parameters are analyzed, and continuous optimization feedback is carried out according to the condition update until the design requirements are met.

Through the research on the digital twin technology composition and digital twin assembly process, the digital twin-based engineering manipulator assembly process planning method is described in detail. In the process of engineering manipulator assembly simulation, digital twin technology is used to visualize the process flow of the assembly site, evaluate the pros and cons of the scheme, analyze whether it is reasonable, and realize the precise control of the whole process from the design of the assembly model of the engineering manipulator to the assembly production.

3.3. General Framework of Digital Twin Engineering Robotic Arm Assembly. By summarizing the assembly process problems of the construction manipulator and the assembly process under the digital twin technology, combined with the preparation process of the assembly process of a certain type of construction manipulator, according to the assembly process design method, the assembly process design process of the construction manipulator in the engineering environment is based on its characteristics and characteristics. The composition is divided into three main parts:

(1) *Data Layer*. We analyze the design size information, assembly process information, and dimensional tolerance information of the engineering robot arm and provide basic data support for the construction of the engineering robot arm assembly model and the control of the assembly process. Through data acquisition methods such as human-computer interaction, hardware acquisition terminals, and sensors, the assembly resources included in the assembly process of the construction robot arm are equipped with intelligent cores such as electronic labels and barcodes, and the data in the assembly production process of the construction robot arm are collected to realize the assembly information.

(2) *Model Layer*. When facing the assembly process of the engineering manipulator, it is necessary to fully consider the multilevel information model in the engineering environment and the needs of the physical assembly site. Starting from the three-dimensional model of the engineering manipulator, according to its assembly characteristics and technical requirements, the assembly level and assembly sequence of the product are determined, so as to obtain an accurate assembly model. By acquiring detailed assembly knowledge such as products, tooling equipment, process knowledge, and logistics, 3D modeling of the assembly production line is performed based on the digital simulation platform to realize the mapping and connection from the physical space to the digital space from the engineering manipulator body to the entire production and assembly process.

(3) *Application Layer*. The design, assembly, and application of the engineering robot arm are a complete life cycle, and an intelligent virtual environment is built in combination with digital twin technology. Through simulation, the specific working conditions of the assembly process are analyzed and fed back to the staff in real time, so as to realize the control and optimization of production capacity and production bottlenecks. The overall framework of the system is shown in Figure 2.

Because in the actual assembly process, the final assembly quality of the product will be affected by some unpredictable actual factors. In order to ensure the dynamic unity of the digital twin and the physical entity, it is necessary to build a digital twin model with all the elements of

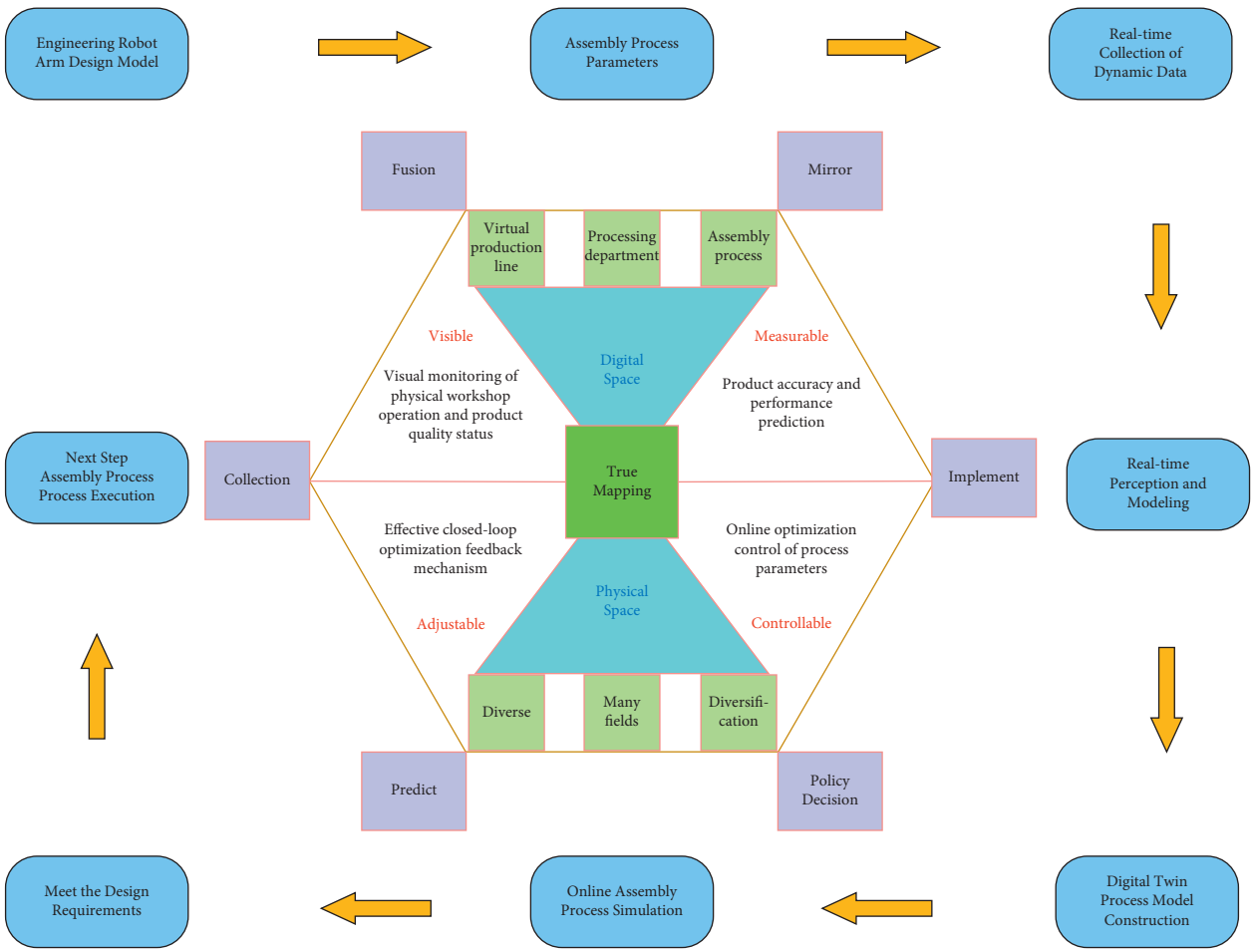


Figure 1: Digital twin assembly process control process diagram.

the physical entity space and realize the data and information exchange between the two. The constructed digital twin model meets the following technical requirements:

(1) *Single Mapping*. The virtual entity structure in the digital twin space needs to be in one-to-one correspondence with the physical product, and the geometric feature information (such as shape, size, and tolerance) and manufacturing process information contained in the physical entity must be in accurate representation in the digital twin.

(2) *Dynamics*. As an effective judgment model for physical entities, the product digital twin is required to reflect the current state of the system in real time during the entire process of production and assembly, so it is necessary to ensure the dynamic unity of the two.

(3) *Predictability*. The process execution process of the whole life cycle of product manufacturing can be established by building a virtual simulation environment, the possible design defects and performance defects can be predicted through the detected real-time data, and the parameters can be adjusted in time.

On the basis of the above framework, they obtain manufacturing resource data such as equipment, products, processes, and logistics and use 3D design tools to establish a digital virtual assembly production line.

3.4. Adaptive Biological Neural Network Learning Method for Discrete Action Output. The Q-learning algorithm is a widely used reinforcement learning algorithm that can be used to optimize strategies in solving Markov decision processes. Taking action a_t for the current state s_t will not only affect the immediate reward r_t but also affect the reward obtained in the future and obtain the Q-value corresponding to the state s_t and the action at

$$Q(s_t, a_t) = \frac{1 + \sum_{i=1}^{n} \gamma^i}{r_t} + \frac{1 + \sum_{i=1}^{n} \gamma^i}{r_{t+1}} + \frac{1 + \sum_{i=1}^{n} \gamma^i}{r_{t+2}} + \cdots + \frac{1 + \sum_{i=1}^{n} \gamma^i}{r_{t+i}}. \quad (1)$$

Among them, the discount function indicates that the current action will weaken the series of rewards obtained in the future as the time step increases.

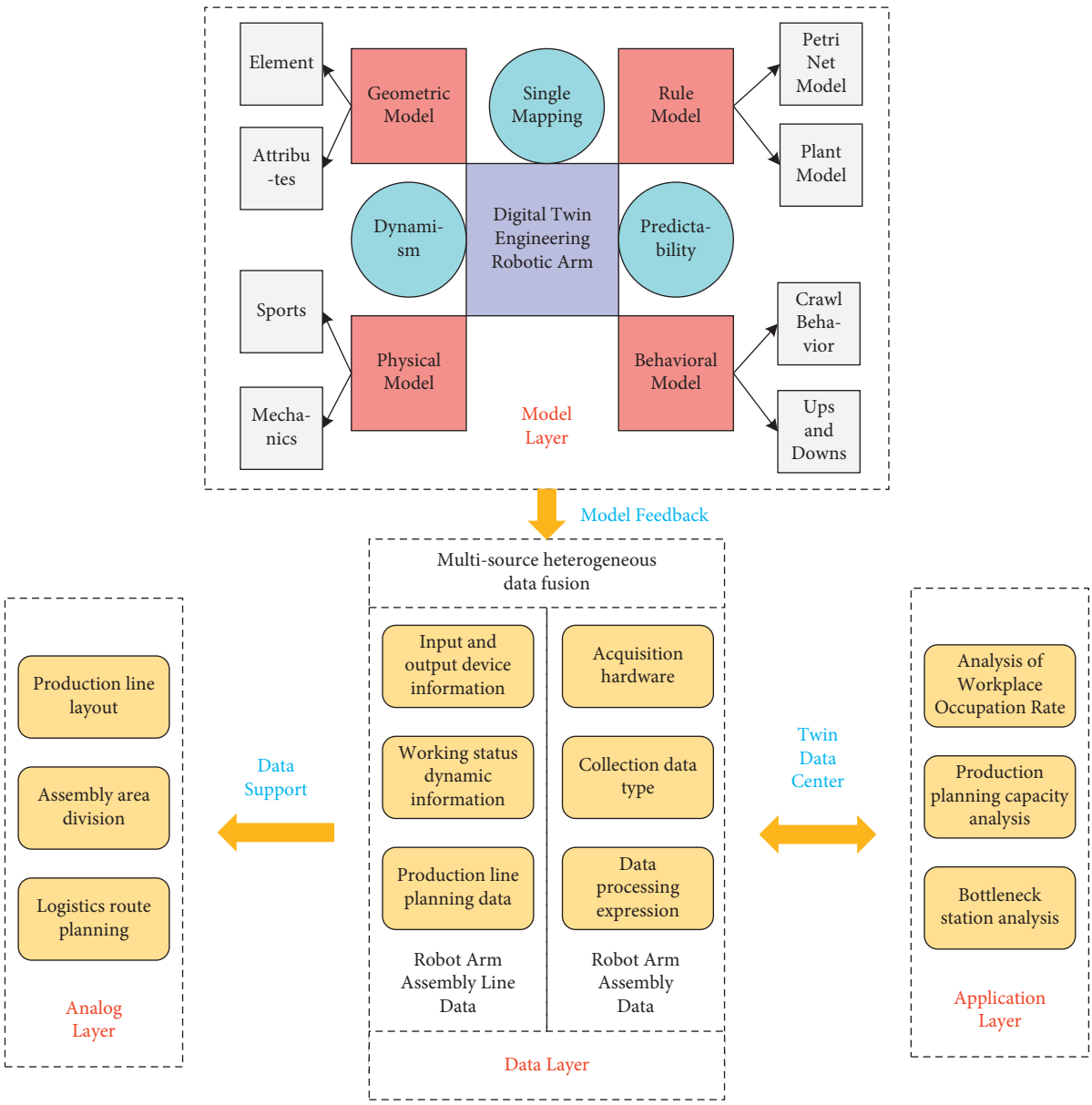


Figure 2: The overall framework of the digital twin engineering robotic arm assembly.

The state for time $t+1$ is obtained as follows:

$$Q(s_{t+1}, a_{t+1}) = \frac{1+\sum_{i=1}^{n}\gamma^{i}}{r_{t+1}} + \frac{1+\sum_{i=1}^{n}\gamma^{i}}{r_{t+2}} + \frac{1+\sum_{i=1}^{n}\gamma^{i}}{r_{t+3}} + \cdots + \frac{1+\sum_{i=1}^{n}\gamma^{i}}{r_{t+1+i}}. \tag{2}$$

The Bellman iteration equation for the state-action value function can be obtained as follows:

$$Q(s_t, a_t) = Q(s_{t+1}, a_{t+1}) - \frac{1+\sum_{i=1}^{n}\gamma^{i}}{r_t} + \frac{1+\sum_{i=1}^{n}\gamma^{i}}{r_{t+1+i}}. \tag{3}$$

In the Q-learning algorithm, the corresponding update equation is as follows:

$$\frac{Q(s_t, a_t)}{\alpha_t(s_t, a_t)} r_t \sqrt{\frac{\gamma \max_a Q(s_{t+1}, a_{t+1})}{Q(s_t, a_t)}} \longrightarrow Q_{t+1}(s_t, a_t). \tag{4}$$

The Q-table implements a mapping strategy from the state s to the optimal action. When the number of states of the environment is large, the storage space required by the Q table will become large. Neural networks can fit large nonlinear functions and have generalization capabilities. Instead of bulky Q tables, neural networks can be used to fit mapping functions from states to actions.

Taking the space engineering manipulator with 6 degrees of freedom as an example, if the motion of each joint is simply discrete into 3 actions of "forward rotation, reverse rotation, and stop," the combination of actions generated by

all 6 joints is 36 = 729. If a finer control is to be obtained, a correspondingly finer discrete action is required, and the number of actions generated will be greater. As a result, the adaptive biological neural network cannot be directly applied to the continuous motion control of the space engineering manipulator.

Deterministic strategy neural network can express deterministic control strategy, the input layer of neural network is state s, and the output is deterministic optimal action.

The deterministic policy gradient neural network does not need to calculate the action value corresponding to each action, so it does not need to discretize the action, so it can output the continuous control action of the multi-degree-of-freedom spatial engineering manipulator.

The deterministic policy gradient neural network adjusts the parameters of the neural network according to the policy optimization gradient. However, it is difficult to obtain the gradient direction of the policy optimization parameter update, which makes the training of the deterministic policy gradient neural network difficult to achieve.

3.5. Deep Deterministic Policy Gradient Algorithm Design. This study will use the deep deterministic policy gradient algorithm to learn and optimize the control strategy of the space robot arm. The DDPG algorithm is based on the actor-critic training system and can be adapted to the continuous control of multi-degree-of-freedom spatial engineering manipulators. The actor part uses a deterministic policy gradient neural network to output the continuous control action of the multi-degree-of-freedom engineering manipulator. The critic part uses an adaptive biological neural network to fit the action Q-value function.

The process of training the critical neural network is as follows: the critical action value function network is used to fit the action value function. The input of the neural network is the state s_t at time t and the action a_t, and the output is the fitting value of the neural network. The target value uses the Bellman iteration equation Q. Therefore, the sum of squares of the difference between the target value and the fitted value is the loss value. The process of fitting the action value function by the network is to update the parameter C of the network to reduce the loss value L.

By adding a soft bound on the angular velocity to the reward function, the reward function is of the following form:

$$r = \frac{d}{\sum_{i=1}^{n} \max\left(0, v_b/\left|v_i\right|\right)}. \tag{5}$$

4. Results and Analysis

4.1. Control Algorithm Simulation Platform Verification. The control system continuously obtains the parameters of the controlled object, the state of the system, the tracking error, and other information, compares it with the expected state and performance, and obtains the estimated value of the uncertain parameter according to the preset estimation rate. The estimated value is used as part of the controller, and the controller is then corrected so that the system achieves the desired tracking performance.

In this section, the controller will be simulated and verified on the engineering manipulator simulation platform. The derivative term of the expected trajectory involved in the simulation is described by the finite difference method. The specific form of the finite difference method to describe the derivative of a function is as follows:

$$\begin{aligned}\dot{f}\left(x_i\right) &= \lim_{\Delta x \longrightarrow 0} \frac{f\left(x_i\right) + f\left(\Delta x - x_i\right)}{\Delta x} \\ &= \lim_{\Delta x \longrightarrow 0} \frac{f\left(x_i + \Delta x\right) - f\left(x_i\right)}{\Delta x}, \\ \ddot{f}\left(x_i\right) &= \lim_{\Delta x \longrightarrow 0} \frac{f^2\left(x_i + \Delta x\right) - 2f\left(x_i + \Delta x\right)f\left(x_i\right) + f^2\left(x_i\right)}{\Delta x^2}.\end{aligned} \tag{6}$$

First, the control effect of the controller based on the system model is tested. The motion effect diagram of the engineering manipulator is shown in Figure 3. It can be seen from these figures that the model-based controller can effectively control the motion of the engineering manipulator and can make its end track the given trajectory well, especially from the detailed diagram of the trajectory tracking of the end of the motion of the engineering manipulator. The actual motion trajectory of the end of the engineering manipulator is almost completely fitted with the expected trajectory.

The controller can control the engineering manipulator to move on the YOZ plane and make its execution end track the given desired motion trajectory. The trajectory of the end of the construction manipulator is tracked in detail, the actual trajectory of the end of the construction manipulator is not completely fitted with the expected trajectory, and there is a slight error. The error may come from the estimation error of the actual system model by the neural network and the choice of control parameters. In general, the controller can also achieve good control results when the system model is unknown.

The controller can control the construction manipulator to move on the YOZ plane and make the end of the construction manipulator track the given desired trajectory. The actual trajectory of the end of the engineering manipulator is not completely fitted with the expected trajectory, and there is a slight error. The error may originate from the estimation error of the actual system model by the neural network, the estimation error of the system velocity term by the state observer, and the selection of control parameters. But in general, when the system model is unknown and the system state quantity is unmeasurable, the controller can achieve a good control effect.

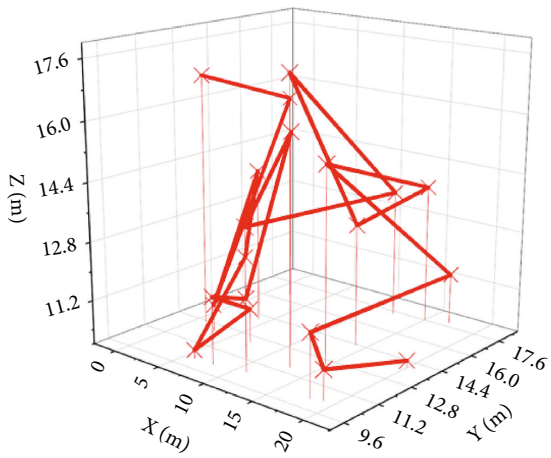


Figure 3: 3D diagram of the motion of the construction manipulator.

The control adjusts the network parameters in real time during the control process, and the parameters are kept within a certain range.

From the above test results of the three controllers, it can be seen that under ideal conditions, the model-based controller can achieve the best control effect, followed by the control effect of the adaptive neural network controller when the model is unknown. Finally, when the system model is unknown and the state quantity is unmeasurable, the control effect of the adaptive neural network controller with the state observer is adopted. In practical applications, if the system model and state are known, the model-based controller can be applied. However, in the actual system, the model is often not accurately determined, and the system state is not necessarily completely measurable, so the above two adaptive controllers have more practical application significance.

4.2. Verification of Control Algorithm Virtual Experiment Platform. In order to further illustrate and verify the effectiveness of the engineering manipulator control algorithm designed in this study, the designed controller is applied to the actual engineering manipulator system. Due to the complexity of the construction of the actual arm system and the progress of the project, the arm system of SRU3 has not been built yet and cannot be used for experimental verification of the control algorithm. Therefore, this study will borrow the existing Baxter robot as the experimental verification platform for the control algorithm.

The Baxter robot is designed with good safety and flexibility, which can facilitate application development and research, and can be flexibly used to deploy industrial production lines and scientific research in universities. The Baxter robot is a robot with two multi-degree-of-freedom structural arms. Each arm is a redundant structure with seven degrees of freedom, and its multiple degrees of freedom greatly improve the flexibility of the arm, which can better simulate the structure of the human arm, and can complete complex operation tasks.

For the convenience of description, the degrees of freedom of each joint of the arm are named. The joints of the arm are driven by a series of elastic drivers, which are different from the direct drive joints. This design can make the engineering manipulator have a certain flexibility and sensing ability. In the process of interacting with people, it can better ensure the safety of personnel. Among the seven joint degrees of freedom, the four joint degrees of freedom, S0, S1, S2, and E0, are mainly used to adjust the trajectory of the end of the arm movement. The three joint degrees of freedom of W0, W1, and W2 are mainly used to adjust the arm after reaching the desired trajectory. Compared with the designed SRU3 engineering manipulator system, the three degrees of freedom of S0, S1, and S2 can be compared to the three degrees of freedom of the shoulder joint of the SRU3 engineering manipulator. SRU3 is the two degrees of freedom of the wrist joint, and W3 is the redundant degree of freedom. S1 and E0 have the same axis of rotation, and the combined motion of the two joints is on the same plane, which makes it easier to see the motion effect.

The control system as a whole is divided into two parts, the development workstation and the robot main body. The development workstation is the user's computer, which is used to write and solve the robot control algorithm; there is an independent computer on the robot body, which is used to schedule robot tasks and generate control commands for the embedded controller on the robot. The mechanism executes the control instructions given by the user. The development workstation and the main robot computer are connected through Ethernet to control the robot.

Baxter is a robot platform developed based on the robot operating system. ROS is an open source system software with certain generality. It relies on the Linux operating system to manage computer resources and provides functions such as underlying driver management, task management and scheduling, and process communication scheduling for robots in the process of robot development. Based on the ROS system, it is convenient to use its interface for the secondary development of the robot. The computer system environment on the Baxter robot is "Gentoo + ROS + SDK," which runs the ROS master independently to manage the robot's control tasks, robot status, and underlying hardware drivers and provides communication and development interfaces. Therefore, any computer capable of running ROS can connect to Baxter through the ROS network interface to operate and develop it. Baxter's SDK provides users with an interface to access underlying hardware such as robot motors and sensors through ROS. The computer system environment of the development workstation is "Ubuntu14.04 + ROS + SDK," and the control algorithm of the robot can be developed and implemented on the user's computer.

The joint trajectory tracking angular velocity of the action process of the model-based controller is shown in

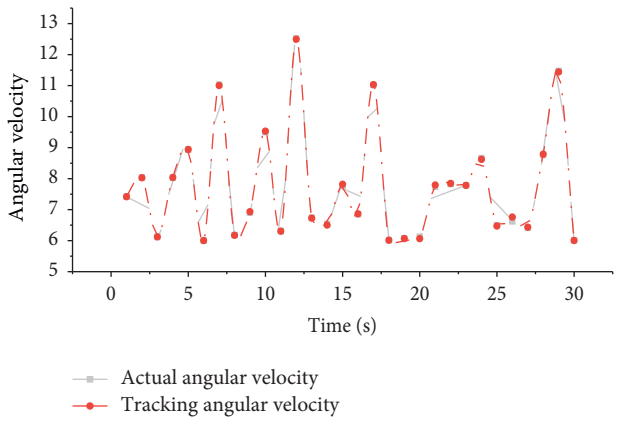


Figure 4: Angular velocity tracking renderings.

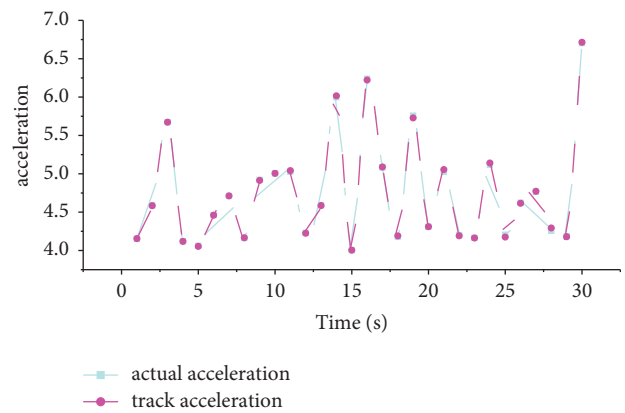


Figure 5: Adaptive neural network controls acceleration tracking effect.

Figure 4. From the tracking effect diagram of the system output tracking the preset trajectory, it can be seen that the model-based controller can almost perfectly track the desired trajectory, and the tracking error is in the range of 0.01. The control input of the system is bounded and stable in the neighborhood of a certain value.

4.3. Experimental Results and Analysis. In this experiment, the S1 and E0 joints of the Baxter robot arm were selected as the controlled objects to verify the effectiveness of the controller. Considering the actual application conditions, it is impossible to establish an accurate system model. In this experiment, only the adaptive neural network controller based on state observation will be experimentally verified. Since the adaptive neural network controller does not need an exact system model or precise system parameters in the application process, it is possible to test the effect of the controller by directly inputting the desired trajectory for the joint without paying attention to the system parameters.

Figure 5 shows the acceleration tracking effect of adaptive neural network control. It can be seen from the figure that the adaptive neural network controller designed in this study can control the engineering manipulator to track the trajectory after setting the appropriate control parameters. The actual tracking error is shown in Figure 6. It can be seen from the figure that the tracking error range is kept within 0.01.

Figure 7 shows the output torque of the controller. It can be seen that the adaptive neural network controller designed in this study can achieve a better control effect in the actual engineering manipulator system control.

When the system state model is unknown and the system state is unmeasurable, appropriate control parameters are selected. The adaptive neural network controller based on state observation designed in this study can also control the engineering manipulator for trajectory tracking. The actual tracking error situation is shown in Figure 8. The absolute value of the maximum tracking error does not exceed 0.02/rad, and the conversion to angle value is also around 1.

The adaptive neural network controller designed in this study can also realize the desired trajectory tracking control

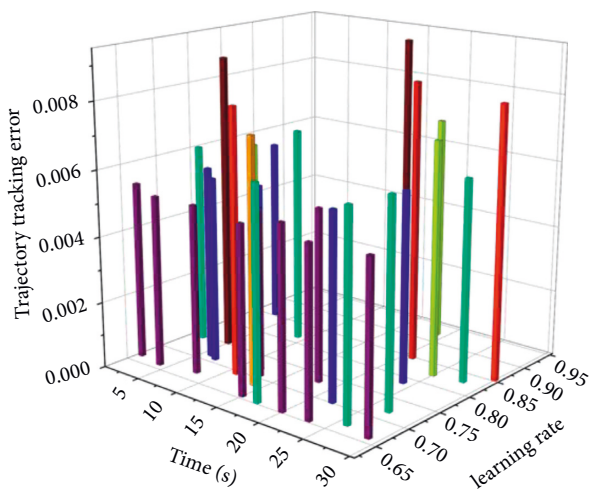


Figure 6: Adaptive neural network control trajectory tracking error.

of the engineering manipulator in the actual engineering manipulator system control.

Since too much information about the precise system model is not required, the above controller has good applicability and can be used to solve control problems of various similar engineering manipulator systems.

Comparing the tracking errors of the two controllers, it can be seen that the controller has a better control effect when the system state is measurable.

But at the same time, when the system model is unknown and the system state is unmeasurable, the adaptive neural network control strategy based on state observation can also provide an effective control scheme.

The above experimental results show that under the action of the two controllers, the trajectory tracking of the engineering manipulator has a certain error. The source of the error may have the following reasons: the controller has many parameters, which may not guarantee that the adjusted control parameters are optimal; at the same time, Baxter's arm joints are all flexible joints, and there will be certain changes in the movement process. The flexible vibration may also have a certain impact on the control accuracy.

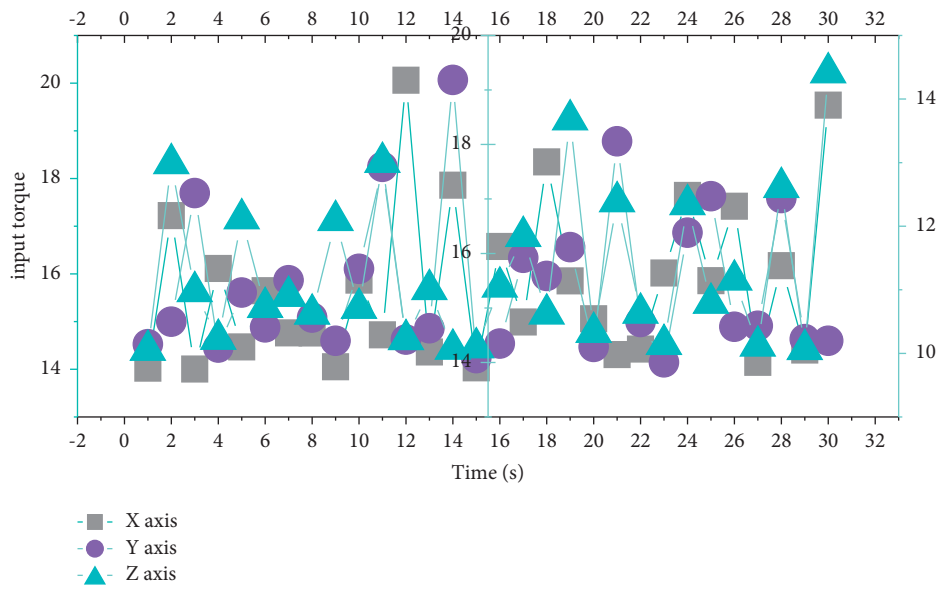


Figure 7: Adaptive neural network controls input torque.

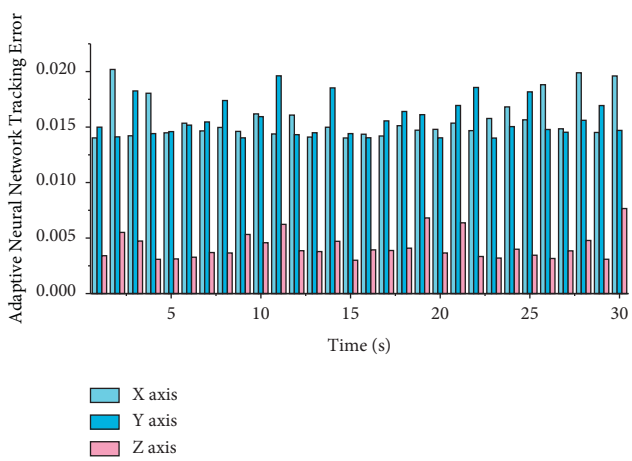


Figure 8: Adaptive neural network control trajectory tracking error based on state observation.

5. Conclusion

In this study, a modeling and simulation framework of engineering manipulator assembly based on digital twin is designed. On the basis of summarizing the research status of digital twin and construction machinery assembly, the shortcomings of traditional construction machinery assembly are analyzed. According to the core concept, application criteria, and model requirements of digital twin, a design process and method of engineering manipulator assembly process based on digital twin are proposed. By studying the modeling requirements of the model layer, data layer, and application layer, the general idea of the digital twin assembly process management and control is put forward, and the assembly modeling and simulation framework of the engineering robot arm in the engineering environment is established. The kinematic control framework of the space engineering manipulator is designed, and the control input of the space engineering manipulator is the desired angular velocity of the joint. Since the control at the kinematic level is not affected by the dynamic parameters of the engineering manipulator, the dynamic parameter difference between the simulation platform and the real physical system has no effect on the transfer of the control strategy. Therefore, the simulation platform can be used to control a large number of engineering manipulators and collect data samples, and these data samples are used to train the deep policy neural network; after the neural network is trained, it can be directly transferred to the control of the real engineering manipulator system. It can achieve the same control effect as the simulation platform, thus avoiding the equipment wear and tear problem caused by relying on the real physical system to generate training data. This study designs a model-based control algorithm and an adaptive control algorithm for the engineering manipulator, considering the control requirements when the system model is known and the model is unknown. The adaptive method in the adaptive control adopts the radial basis function neural network and uses the fitting performance of the neural network to fit the engineering manipulator system. At the same time, the two situations of completely known and unknown system state are considered separately to solve the control under various environmental conditions in practical applications. The Lyapunov stability principle and MATLAB simulation are used to analyze and verify the effectiveness of the controller and system stability. Finally, before the construction of the engineering manipulator platform has been completed, the two adaptive neural network controllers designed in this study are experimentally verified with the help of the Baxter robot arm system, which provides further proof of the effectiveness of the algorithm.

Data Availability

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

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work was supported by the Education Department of Guizhou Province Fund (Qianjiaohe KY Zi [2017] no. 265, Qianjiao XKTJ [2020] no. 23), the Science and Technology Department of Guizhou Province Fund (Qiankehe Platform Talent [2019] no. 5620, Qiankehe Platform Talent-YSZ [2021] no. 001), Design and Research on 7DOF Redundant Manipulator of Picking Robot (Qianjiaohe Zi [2020] no. 128), and Mechanical Equipment Application and R&D Technology Innovation Team (LPSSYKJTD201802).

References

[1] W. H. Tian, C. C. Jhan, M. Inokuma, S. Dohta, T. Akagi, and S. Shimooka, "Development of a tetrahedral-shaped soft robot arm as a wrist rehabilitation device using extension type flexible pneumatic actuators," *Journal of Robotics and Mechatronics*, vol. 32, no. 5, pp. 931–938, 2020.

[2] F. A. Raheem, A. T. Sadiq, and N. A. F. Abbas, "Robot arm free Cartesian space analysis for heuristic path planning enhancement," *Int. J. Mech. Mechatron. Eng*, vol. 19, pp. 29–42, 2019.

[3] P. Vestartas and Y. Weinand, "Laser scanning with industrial robot arm for raw-wood fabrication," *Proceedings of the International Symposium on Automation and Robotics in Construction. IAARC Publications*, vol. 37, pp. 773–780, 2020.

[4] M. Pollák and J. Dobránsky, "Structural design and material cutting using a laser end effector on a robot arm," *Tem Journal*, vol. 9, no. 4, pp. 1455–1459, 2020.

[5] M. Yin, X. Liu, Y. Liu, and X Chen, "Medical image fusion with parameter-adaptive pulse coupled neural network in nonsubsampled shearlet transform domain," *IEEE Transactions on Instrumentation and Measurement*, vol. 68, no. 1, pp. 49–64, 2019.

[6] C. J. Liang, K. M. Lundeen, W. McGee, C. C. Menassa, S. Lee, and V. R Kamat, "A vision-based marker-less pose estimation system for articulated construction robots," *Automation in Construction*, vol. 104, pp. 80–94, 2019.

[7] H. Yu, I. S. Choi, K. L. Han, J. Y. Choi, G. Chung, and J Suh, "Development of a upper-limb exoskeleton robot for refractory construction," *Control Engineering Practice*, vol. 72, pp. 104–113, 2018.

[8] L. Vasey, B. Felbrich, M. Prado, B. Tahanzadeh, and A Menges, "Physically distributed multi-robot coordination and collaboration in construction," *Construction Robotics*, vol. 4, no. 1-2, pp. 3–18, 2020.

[9] H. Huang, C. Yang, and C. L. P. Chen, "Optimal robot–environment interaction under broad fuzzy neural adaptive control," *IEEE Transactions on Cybernetics*, vol. 51, no. 7, pp. 3824–3835, 2021.

[10] J. F. Buhl, R. Grønhøj, J. K. Jørgensen et al., "A dual-arm collaborative robot system for the smart factories of the future," *Procedia Manufacturing*, vol. 38, pp. 333–340, 2019.

[11] J. F. Qiao, Y. Hou, L. Zhang, and H. G. Han, "Adaptive fuzzy neural network control of wastewater treatment process with multiobjective operation," *Neurocomputing*, vol. 275, pp. 383–393, 2018.

[12] I. Dakhli, E. Maherzi, and M. Besbes, "Synthesise of MPC controller for uncertain systems subject to input and output constraints: application to anthropomorphic robot arm," *International Journal of Automation and Control*, vol. 14, no. 1, pp. 80–97, 2020.

[13] P. Liu, H. Yu, and S. Cang, "Adaptive neural network tracking control for underactuated systems with matched and mismatched disturbances," *Nonlinear Dynamics*, vol. 98, no. 2, pp. 1447–1464, 2019.

[14] J. Wang, Z. Liu, Y. Zhang, C. L. P. Chen, and G Lai, "Adaptive neural control of a class of stochastic nonlinear uncertain systems with guaranteed transient performance," *IEEE Transactions on Cybernetics*, vol. 50, no. 7, pp. 2971–2981, 2020.

[15] H. Nagano, H. Takenouchi, N. Cao, M. Konyo, and S Tadokoro, "Tactile feedback system of high-frequency vibration signals for supporting delicate teleoperation of construction robots," *Advanced Robotics*, vol. 34, no. 11, pp. 730–743, 2020.

[16] H. Mo and G. Farid, "Nonlinear and adaptive intelligent control techniques for quadrotor uav – a survey," *Asian Journal of Control*, vol. 21, no. 2, pp. 989–1008, 2019.

[17] J. Fei and Y. Chu, "Double hidden layer output feedback neural adaptive global sliding mode control of active power filter," *IEEE Transactions on Power Electronics*, vol. 35, no. 3, pp. 3069–3084, 2020.

[18] Z. Wang, H. Li, and X. Zhang, "Construction waste recycling robot for nails and screws: computer vision technology and neural network approach," *Automation in Construction*, vol. 97, pp. 220–228, 2019.

[19] J. Yu, C. Hu, H. Jiang, and L Wang, "Exponential and adaptive synchronization of inertial complex-valued neural networks: a non-reduced order and non-separation approach," *Neural Networks*, vol. 124, pp. 50–59, 2020.

[20] Q. Shen, P. Shi, J. Zhu, S. Wang, and Y Shi, "Neural networks-based distributed adaptive control of nonlinear multiagent systems," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 31, no. 3, pp. 1010–1021, 2020.

Hindawi
Computational Intelligence and Neuroscience
Volume 2022, Article ID 1926227, 11 pages
https://doi.org/10.1155/2022/1926227

Research Article

Neural Network Model Design for Landscape Ecological Planning Assessment Based on Hierarchical Analysis

Jing Liu [1] and Xudan Zhou [2]

[1] *School of Xiamen University of Technology, Xiamen, Fujian 361024, China*
[2] *College of Forestry and Grassland of Jilin Agriculture University, Changchun, Jilin 130118, China*

Correspondence should be addressed to Jing Liu; 2013110905@xmut.edu.cn

Received 6 July 2022; Revised 5 August 2022; Accepted 9 August 2022; Published 29 August 2022

Academic Editor: Ning Cao

In this paper, an in-depth study and analysis of landscape ecological planning and evaluation are carried out using the analytic hierarchy process (AHP) algorithm that integrates neural networks. The application of AHP in the field of tree species planning and the introduction of quantitative analysis methods can effectively change the subjectivity of previous qualitative analysis in tree species selection and make it objective, scientific, and reasonable. The research can provide a reference for other urban tree species planning. From the connotation of landscape ecological service process and ecological space structure, the analysis of landscape ecological service process involves service supply area and service association area, which correspond to different key components of ecological space structure. With the help of the platform, based on the identification and identification methods and theories of ecological spatial structure, the key components of ecological spatial structure in different environments are identified and extracted by using the representation model, binary suitability model, weighted suitability model, and process model. The type of service is based on the different service processes supported by the key components of the ecological spatial structure, forming the ecological spatial structure under different service types. Spatial structure; on this basis, the basic characteristics of the key components of the ecological spatial structure are analyzed, and the correlation characteristics of the ecological spatial structure are analyzed based on the correlation classification system of ecological spatial structure. A backpropagation (BP) neural network-based state assessment method of the grid structure is established. The method takes the parameters of the autoregressive model constructed by the acceleration signals of different working conditions as the feature quantity and the results of the fuzzy hierarchical analysis method as the labels, divides the data set into a training set and a test set, and uses the BP neural network learning method and the training set to supervised train the BP neural network learning assessment model. The test set is used to test the effectiveness and accuracy of the BP neural network-based learning method. The study shows that the evaluation system established by the BP neural network structure is fast and accurate and can substantially reduce the cost of manual testing.

1. Introduction

With the deepening of theoretical research in ecology and other disciplines, the theoretical and practical methods of landscape ecological planning have been developed in various countries. The early landscape ecological planning was short-sided and relied more on subjective materialistic judgments for design, but with the access and role of these integrated disciplines, the later design increasingly relies on the integrated consideration of multiple disciplines for more reasonable planning and design. In the context of multi-planning, the study of cognitive frameworks provides an opportunity to unify the cognition of more decision-makers, provide methods for clear cognition of sites, and coordinate multiple decision-making to propose suitable planning methods. The practice of landscape ecological planning started late, and the translation of multidisciplinary knowledge into tools to guide space and methods to guide planning practice belong to research gaps. Nowadays, increasingly ecological and environmental problems threaten human life, and the need for new holistic approaches is more urgent [1]. As a disciplinary direction that aims to integrate

multiple disciplines to coordinate the relationship between humans and nature, it is necessary to organize the existing landscape cognitive view and research methods to contribute to environmental issues. With the impact of urban construction and public infrastructure in recent years, the surrounding ecological environment has been destroyed, making the ecosystem service function gradually weakened. The hilly and mountainous region of Jiangnan is rich in forest resources, lakes and rivers, and fewer plains, but most of the human socio-economic development is concentrated in the water-rich plains, thus forming a unique landscape unit with complex land use and fragile ecological environment [2].

This paper combines the research methods of cognitive psychology with the practice of landscape ecological planning and composes the historical process of multidisciplinary intervention in landscape ecological planning and design through historical research [3]. People have invested a lot of human and financial resources in the scientific planning of urban wetland parks to achieve the maximum ecological value, but the path of ecological construction of urban wetland parks, rational development, scientific planning, and sustainable development is still being explored. Landscape ecological planning is a scientific planning method that combines the principle of ecological nature with landscape planning, by studying the landscape pattern—the ecological process and the interaction between human activities and the landscape, based on landscape ecological analysis and comprehensive evaluation, the optimal solution and the proposed landscape planning approach, the connotation of which lies in respecting the ecological process and ecological patterns, respecting species diversity, reducing deprivation of resources, maintaining nutrient and water cycles, maintaining the quality of plant habitats and plant and animal habitats, any form of design that minimizes the impact of damage to the environment is called ecological design. To make full use of wetland resources, protect the ecological environment, promote urban-rural integration, and maintain the sustainable development of the city and the environment, landscape ecological planning for urban wetland parks, through the combination of economic planning, environmental planning, and landscape design, so that regional development, resource utilization, and ecological protection are linked to achieving a high degree of unity of economic, social and ecological benefits, and the overall optimization [4]. This is the main way to build a high-quality, harmonious, and livable city and is the way to meet the growing demand of the people for a better life.

The use of hierarchical analysis to plan the garden tree species in Kunming combines quantitative analysis with qualitative analysis, which can achieve scientific and objective tree species planning. The scientific and reasonable planning of urban garden tree species can realize the sustainable development of urban landscaping and make the urban green space system planning effectively play the role of guiding the construction of urban green space and improving the benefits of urban greening. This research result realizes the application of hierarchical analysis to tree species planning, which can verify the rationality of tree species planning in the existing green space system planning and provide a reference basis for the garden tree species planning in other cities. The principles and strategies and methods of landscape planning and design for sponge campus with regional adaptation are proposed, and the proposed strategies and rainwater landscape creation methods are applied to design practice. The paper tries to see the big picture in a small way through this site feature of campus landscape planning and design so that the sponge campus landscape planning and design based on regional adaptation can provide a reference for the planning and design of stormwater landscape in the same type of site in the future.

2. Related Work

With the rapid development of the third information technology revolution and the large-scale application of computer technology, the simulation and prediction of stormwater processes using various stormwater simulation software have given rise to the introduction of relevant policies and industry standards for stormwater management [5]. Compared with the traditional stormwater treatment, the new research is more focused on coordinated development and is a forward-looking vision. However, the planning and design of stormwater control in the pavilion have failed to meet expectations in the process of practice because of the lack of relevant laws and regulations to guide and various supporting and coordinating facilities [6]. With the rapid development of the industrial industry in Britain, many laborers moved from the countryside to the cities, and urban housing problems, traffic problems, and environmental problems came to a head. As urban problems became increasingly serious, there was a movement of people moving from big cities to suburban areas to live in idyllic cities, and people took practical actions to pursue a suitable living environment [7]. He argues that high-income residents have a personal stake in urban development and that the gap between rich and poor urban residents can affect the overall livability of cities, so he suggests that the government should reduce the gap between rich and poor urban residents by effectively strengthening control over economic units. Kang has done a lot of research on the theory of livable cities and the human living environment, and he proposes that the construction of livable cities should highlight the concept of being "people-oriented" [8].

The landscape pattern refers to the characteristics and spatial pattern of the landscape component units. Different configurations and combinations of landscape elements form different landscape patterns, and the interaction of different landscape structures and the circulation of different components of material and energy flows in the landscape, i.e., ecological processes, will make the landscape pattern differentiated [9]. General landscape spatial structure characteristics are regular, this law created the landscape pattern of expression. Landscape patterns can reflect the landscape productivity, habitat quality control factors, and ecological stability, according to the pattern of the law of effective prediction of landscape dynamics, in the later stages

of the establishment of landscape planning, design, and management objectives [10]. The landscape pattern of wetlands is a special configuration and combination of different landscape elements in the wetland space and is also the result of long interaction between wetland ecological processes and wetland landscape structure [11]. The landscape pattern of artificial wetlands is the principle of human planning and design, respecting the standards of landscape ecological planning, respecting the laws of nature, respecting the normal exchange of materials and energy, and reducing to a minimum all designs that may damage the ecological environment and natural resources to the wetland park planning [12]. The analysis of the landscape pattern can reveal the stability of the ecological environment of urban wetland parks and potential security risks and can also determine the root cause of the generation and control of the landscape pattern [13].

In the studies of small-scale landscape pattern analysis, most scholars only analyze the landscape pattern of a single park, resulting in simple data levels and a lack of comparability, and since there is no standard basis for the meaning of index values, the conclusions of most studies are too subjective and not highly referential, lacking guiding suggestions for park planning and construction in a city, not to mention the geographical environment and urban development conditions different other regions. Using hierarchical analysis to plan garden tree species, combining quantitative analysis with qualitative analysis, can realize the scientific and objective nature of tree species planning. Scientific and reasonable planning of urban garden tree species can realize the sustainable development of urban landscaping and make the urban green space system planning effectively play the role of guiding the construction of urban green space and improving the benefits of urban greening.

3. Hierarchical Analysis of Landscape Ecological Planning Assessment Neural Network Model Design

3.1. Hierarchical Analysis Fusion Neural Network Model Design. The origin of the artificial neural network is from the neural system in biology class, which is a physical mechanism based on the knowledge of network topology, and simulates the way of human brain processes things by many computing units, and then carries out the mathematical model of distributed transmission. It not only has the ability of self-repeating learning but also imitates the thinking, information processing, and intelligent functions such as storage, recognition, and classification of filtering useless knowledge of the human brain's nervous system to a certain extent claim. The neural network has the following advantages over the human brain system: functions that can be realized by hardware and software together, nonlinear mapping processing capability, non-limitation that can mimic the brain's associative memory, avoidance of complex mathematical models requiring only input and output network topology knowledge, easy distributed parallel computing, fault tolerance and storage capacity, etc.

The structure of artificial neurons is like that of biological neurons, and it is also an abstraction and simulation of the biological nervous system, the so-called abstraction is reflected in the mathematical perspective, and the simulation is reflected in the functional perspective of the structure. From the characteristics and functions of the human brain, a neuron is a signal processing unit. All neurons can have the ability to simulate computer processing information, and their processing steps are divided into three steps: input processing, activation processing, and output processing [14]. Clear awareness of the site, coordination of multi-party decision-making, and proposing appropriate planning methods provide the means. The analysis process takes the product of each input activity and its weight on the neural link, also takes the sum of all weighted inputs to get a total value, and finally passes the transfer function and converts the total value into an output activity. When applying the neural network approach to spatial grid structure state assessment problems, its neuron model often determines the applicable transfer function based on the nonlinear system of the actual spatial grid structure. The information transfer between neurons is also operated by the transfer function.

The hidden layers of a BP neural network generally contain one or more nodes, and there is no linkage between the same layers. The input signal needs to pass through all the nodes in the hidden layer to reach the output node, and the output signal can only affect the output of the nodes in the next layer. As the network processes the data, the neurons are activated and the signals are passed from the input layer to the output layer through the intermediate layers, thus stimulating the neurons in the output layer to produce the input results. The next step corrects the direction of the target output and the actual output by allowing the signal to be passed back from the output layer to the input layer, correcting the connection weights of each layer separately, and this reverse transfer method reduces most of the errors in the input data.

$$z_k = f_1\left(\sum_{l=0}^{n} v_{ki}^2 x_l^2\right), \quad k \in Q. \tag{1}$$

The output of the output layer node is:

$$y_i = f_2\left(\sum_{k=0}^{q} w_{ki}^2 Z_l^2\right), \quad k, i \in M. \tag{2}$$

The hierarchical analysis is a decision-making method that decomposes various types of factors affecting the research objectives into different layers, compares the importance of each layer, and translates the comparison results into weight values. It has the advantages of being concise, practical, and easy to operate. In the calculation process of hierarchical analysis, it is necessary to choose suitable weight vector calculation methods, such as the arithmetic mean method, geometric evaluation method, least-squares method, and eigenvector method. The hierarchical analysis is suitable for problems with hierarchical interlaced

evaluation indexes and difficult to quantitatively describe the target value of the system [15]. Some scholars often use the hierarchical analysis method in conjunction with the fuzzy integrated evaluation method to study a wider range of problems.

The study of urban livability is a small branch in the field of environmental science, and the establishment of an urban livability index system in Hebei province is a target system problem with interlaced evaluation indexes, and the target values are difficult to be described quantitatively, which is within the scope of application of hierarchical analysis.

According to the judgment matrix established in the previous step, the maximum eigen root max and the corresponding eigenvector W are calculated, and then the consistency test is performed, and if it passes, it is determined as the eigenvector for the next calculation; if it does not pass, the judgment matrix is incorrect due to inconsistency in the judging criteria, and the scaling scoring needs to be performed again to construct a new judgment matrix, and the square root method calculation formula is as follows:

$$M_i = \prod_{j=1}^{m} a_{ij}^{m}, \quad i = 1, 2, \ldots, m. \tag{3}$$

People use the things they know, through the previous common sense of life and previous research to confirm the affiliation function. With a view to contributing to environmental issues. With the impact of urban construction and public infrastructure construction in recent years, the surrounding ecological environment has been destroyed, and the ecosystem service function has been gradually weakened. This strategy is often used for those things and targets that are already well understood and familiar, or for scenarios where it is not easy to get actual data. The strategy is to count the relationship between the importance of different elements in comparison to each other and then rank the elements to determine the affiliation function. This strategy can work better in situations where certain features require many professionals to delineate the boundaries, as shown in Figure 1.

When building an ANN model, the number of nodes in the input layer is related to the number of dimensions one wants to input, and the number of nodes in the output layer is the same as the number of properties in the output, but the number of nodes in the middle layer can be adaptively adjusted according to the actual situation; in the ANN three-layer model shown above, the arrow is the direction of data propagation, and the circle is the neuron node [16]. The calculation process propagated from the bottom input layer to the top output layer, and the weight value of each link is updated through continuous calculation, and the error is back-propagated to make the weight value of each link reach the required value, which finally makes the whole model achieve the optimal prediction effect.

$$\phi = \left(\left\|X_i - C_i^2\right\|\right). \tag{4}$$

Radial basis function ANN model: The radial basis function is used as the basis of the hidden layer, and the vector of the input layer of the network is computationally transformed into each neuron node of the hidden layer. When the centroids of each node in the hidden layer are computed using the k-means method, then all the computational processes from the input layer to the intermediate layer are fixed. In summary, the process from the input layer to the intermediate layer cannot be expressed using a one-time function, however, the results from the intermediate layer to the output layer are obtained by weighting the results of each hidden layer node and then accumulating all the results, where the weights vary with different data samples or learning times. In this whole process of transformation, it is possible to map the low-dimensional problem to a higher dimension, making the process describable using the network. So, when the output of the hidden layer is known and the expected result, we can answer the weights of each hidden layer neuron by uniting the whole set of linear equations.

By default, a Gaussian function is generally used as the network driving function as shown in equation (5). Its connotation lies in respecting the ecological process and ecological pattern, respecting species diversity, reducing the deprivation of resources, maintaining the nutrient and water cycle, and maintaining the quality of plant habitats and animal and plant habitats.

$$\phi = \exp\left(\left\|X_i - C_i^2\right\| - \frac{e^2}{2s^2}\right). \tag{5}$$

The classification system of ecological spatial structure composition relatedness is the basis for the analysis of ecological spatial structure relatedness characteristics. The classification system of ecological spatial structure composition relatedness is based on landscape ecological service mode and landscape ecological service range and classifies the spatial location relationship among service supply area, service-related area, and service demand area [17]. So far, spatial correlation classification research is relatively sparse, and the cognition of spatial correlation relies heavily on the division of spatial relationship categories between service supply area and service demand area.

It is known from the connotation of ecological spatial structure that the landscape ecological service process is the internal correlation mechanism of ecological spatial structure formation, and different types of landscape ecological service processes need different ecological spatial structure characteristic attributes to support. Therefore, to identify the ecological spatial structure, we need to clarify the key components of the ecological spatial structure under different landscape ecological service types to support the ecological service process, and then summarize the key attributes of the ecological spatial structure associated with them, and finally form the key elements to be considered for the identification of ecological spatial structure under various landscape ecological service types to support the corresponding landscape ecological service process, as shown in Figure 2.

The network layer, as the lowermost layer of the model proposed in this paper, is the cornerstone of the whole assessment system and plays a crucial role. The selection of

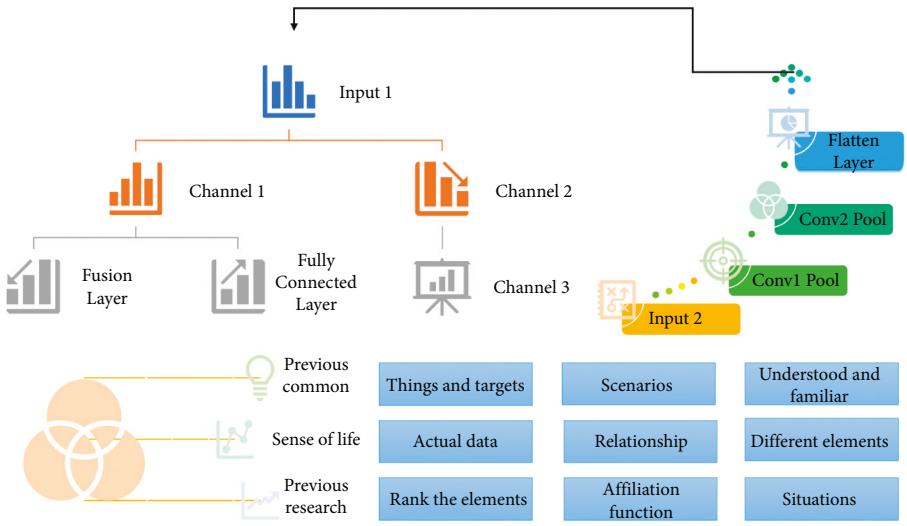


Figure 1: Hierarchical analysis fusion neural network model.

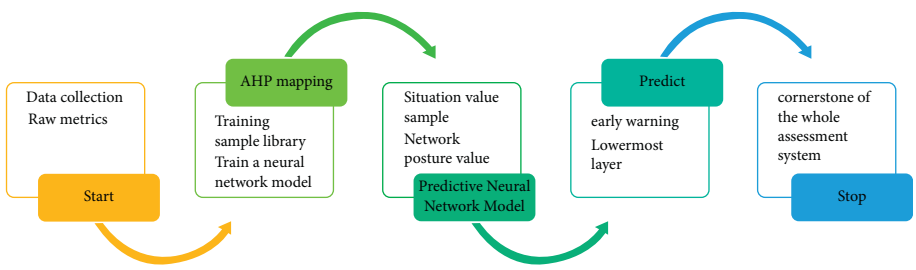


Figure 2: Prediction model.

network layer indicators directly determines the reasonableness and accuracy of the whole model in assessing the network operation, so the selection of network layer indicators needs to be carefully combined with various factors. Therefore, this paper first describes how to select the network layer metrics.

Since bandwidth, traffic parameters of the network, and throughput metrics represent similar attributes and are relatively redundant metrics, only one of them needs to be selected to represent the size of the transmitted data volume in the network [18]. Any design form that minimizes the damage to the environment is called designed for ecology. Also, since the false packet rate is not mentioned and not used in previous literature, this paper discards this metric. In summary, this paper selects four metrics, namely, delay, delay jitter, packet loss rate, and throughput, as the metrics of the network layer.

3.2. Experimental Design for Landscape Ecological Planning Assessment. The impact of land use and landscape pattern on ecosystem services is mainly expressed in the change of ecosystem service value. This method analyzes the relationship between biotic and abiotic factors in time from a vertical perspective and uses the "lasagna" model to synthesize and filter this knowledge and results to solve the problem. He believes that nature is an evolutionary process, a vertical process in which physical and biological factors are constantly interacting and changing with each other under the action of time, and evolving in concert. It forms a certain value system that can provide certain services and constraints for human life. He suggests combining landscape ecological planning with natural processes and calls on landscape planners to use ecological methods to understand nature: and describe its processes to guide the development and utilization of land.

He sees the landscape as a vertical structure formed by the superposition and interaction of elements and considers the landscape as a vertical unit structure reflecting the intrinsic value of the land. In other words, it is necessary to study in detail the evolutionary laws of natural existence and to have a deep understanding of the characteristics of the landscape to find the appropriate value of the land itself and to realize the most suitable use of the land.

Different habitat units cause different distributions of organisms, and each distribution is the result of mutual matching between organisms and the environment [18]. Organisms are influenced by environmental changes and actively select or adapt to different types of habitat units. The distribution of organisms is best matched to the habitat unit, and the species preserved through adaptation is the result of the best match between ecological factors and the traits of the organisms or the best match between the characteristics

and traits of the organisms and the environment in which they live. This match must be made for the organism to survive, as shown in Figure 3.

By investigating and analyzing the natural resources of the area, determining the characteristics of the resources, and judging the ecological performance, analyzing, and evaluating the ability of the land in the area to adapt to development, and thus clarifying the future role and use of the land, a state formed in which the behavioral activities of human modification of nature are consistent and harmonized with the characteristics of the site and natural processes.

The landscape area TA is an index describing the total area of the park, the number of patches NP refers to the number of all patches in the park, and, the density of patches PD refers to the ratio of NP to TA; the ratio of the largest patches to the total area of the park LPI refers to the ratio of the largest patches to the total area of the park and is used to describe the most dominant patches in the park. The size of the area at the landscape level has a great influence on the ecological stability of the whole park, and a park with a large area can play a higher ecological benefit and its ecological stability is also better. The density index, on the other hand, reflects the degree of fragmentation of the landscape, and on the other hand, the density index also reflects the situation of landscape diversity and landscape richness.

Make urban green space system planning effectively play the role of guiding urban green space construction and improving urban greening benefits. The results of this research realize that the analytic hierarchy process is applied to the field of tree species planning, which can verify the rationality of tree species planning in the existing green space system planning. Although the situation faced in different regional scales is complex and varied, because the laws of surface runoff and hydrological cycle are universal, rainwater control methods based on the concept of low-impact development are generally accepted. But at the same time, sponge cities are not equivalent to certain kinds of fixed low-impact development facilities, and specific stormwater site environments should use different solutions and strategies, as shown in Figure 4.

The multi-factor superposition evaluation method is the most basic method among multi-factor comprehensive evaluation methods. This method does not require complex model calculation, directly superimposes each factor affecting the ecological service process of the landscape, uses evaluation criteria to screen and limit the range of key composition areas of ecological spatial structure, and requires high data accuracy because this method directly uses spatial data or relevant map pieces for spatial evaluation analysis. The combination of multiple indicators and data sources is used and the minimum homogeneous geometric area unit is generated by the GIS overlay function, and this method is also the main method for service supply area (SPA) or service demand area (SDA) identification. In addition, regional units are extracted based on quantitative index evaluation results, and most of the index selection involves the depiction of ecosystem conditions and current land use characteristics.

The logical framework of the evaluation of the service performance of the ecological spatial structure shows that the discrimination of the natural and pattern attributes required to support the service supply area (SPA) and service association area (SCA) of different landscape ecological service processes, and the determination of the evaluation benchmark in this way is one of the cores of the research on the evaluation of the service performance of an ecological spatial structure [19–21]. The evaluation of the performance of ecological spatial structure services is based on the natural attributes and pattern attributes. To summarize the service performance evaluation benchmarks of key components of the ecological spatial structure under various landscape ecological service types.

4. Results Analysis

4.1. Hierarchical Analysis Fusion Neural Network Model Performance Results. In the whole grid structure health state evaluation system, the calculation and rating of the third level performance layer, the second level system layer, and the first level target layer draw on the relevant specifications, and deterministic evaluation. While the third level performance layer indicators are influenced by the fourth level location, stress ratio, and damage degree on the structural state, so combined with this test then the fuzzy comprehensive evaluation method is used to assess the fourth level indicator layer.

The distribution form of the fuzzy problem of a comprehensive evaluation of this test is the same as most of the actual engineering, and it is handled by trapezoidal and approximate triangular distribution, which is combined with common trapezoidal distribution to derive its affiliation function as shown in Figure 5.

The process of the damage state of each component of the whole grid is considered comprehensively in the spatial grid structural health state assessment system, and the influence of location, stress ratio, and damage degree on the structural state is calculated in this paper based on the state assessment method of hierarchical analysis. Then a mathematical model of distributed transfer is performed. It not only has the ability of self-repetitive learning, but also imitates the human brain nervous system's thinking, information processing, storage, identification, and classification of the intelligent functions of filtering useless knowledge to a certain extent and saying.

The least-squares errors for training and testing are shown separately, where the least-squares error for batch size 10 is less than 0.1. The trend of least-squares error throughout training and testing shows that the least-squares error decreases faster in the early stage of training, and increases when the batch size is larger than 10, instead of decreasing the training improvement of the neural network. This is related to the number of layers and the number of units per layer of the selected neural network. To investigate the effect of the number of layers and cells, the number of layers and cell nodes will be adjusted in the subsequent subsections to quantitatively analyze the effect of the number of layers and cells on the data of this experiment.

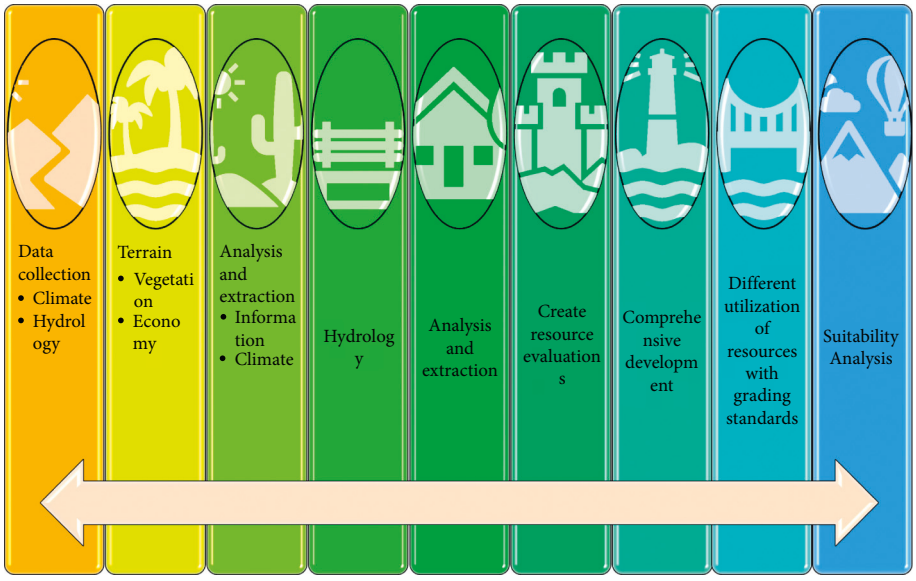


Figure 3: Ecological planning flow chart.

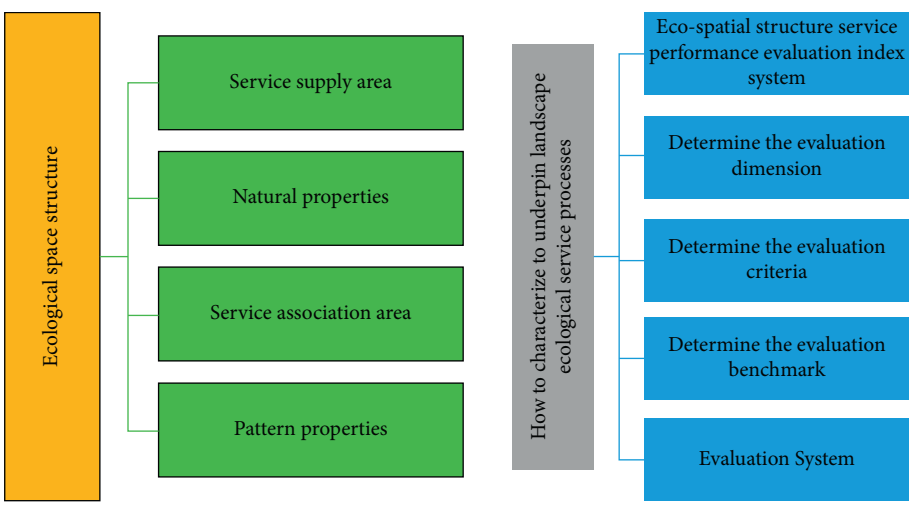


Figure 4: Ecological spatial structure evaluation logic.

As shown in Figure 6, the changes in gradient, Mu value, and the error rate of the neural network during the training process are demonstrated. From the gradient graph, it is easy to conclude that the gradient value will slowly decay to 0.010422 as the number of batch layers increases, the gradient decreases continuously as the training progresses, and the gradient is already less than 0.01 when the batch size is 10, which corresponds to the neural network training effect test graph. The Mu value determines whether the model is trained according to the Newton method or the gradient method. As the Mu value increases, the learning process is trained mainly based on gradient descent, and when the iteration makes the error increase, the Mu value increases until the error no longer increases, but if the Mu value is too large, it will cause the learning to stop when the minimum error has been found, which is why the learning is stopped when the Mu value reaches the maximum value. The Mu value represents the weight error tuning. The Mu value represents the area where the range can affect the number of outputs. From Figure 6, the Mu value fluctuates in the first three batch sizes and remains constant at 0.0001 when it is greater than three. When applying the method of the neural network to evaluate the state of the spatial grid structure, the neuron model often determines the applicable transfer function according to the nonlinear system of the actual spatial grid structure. The change in the error rate is consistent with the neural network training effect test plot, where the error rate increases after the batch size of 10.

The regression analysis system in the network, as shown in Figure 7, gives a precise linear relationship between the target value and the actual output with $R = 1$. This layout (topology) of the BP neural network is satisfactory.

The overall error of state evaluation using a neural network is relatively uniform, and few local errors very

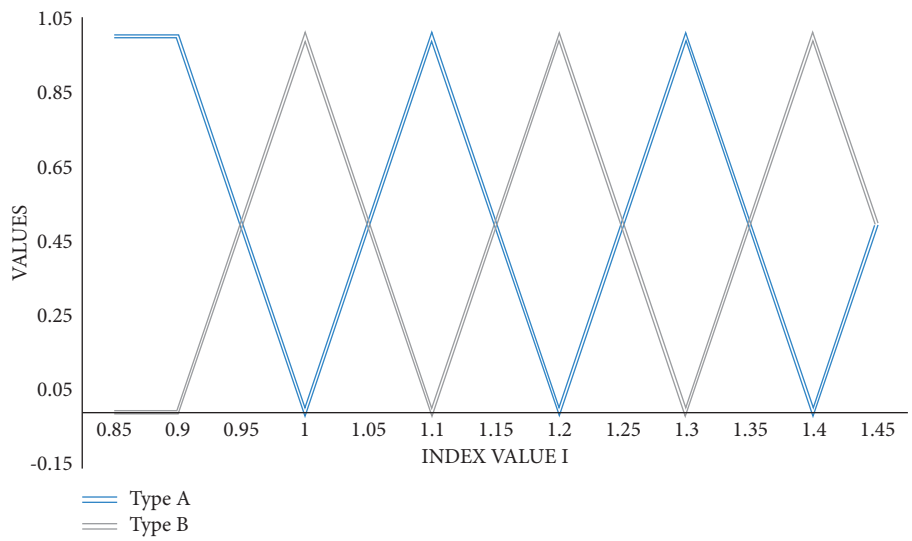


Figure 5: Affiliation function diagram.

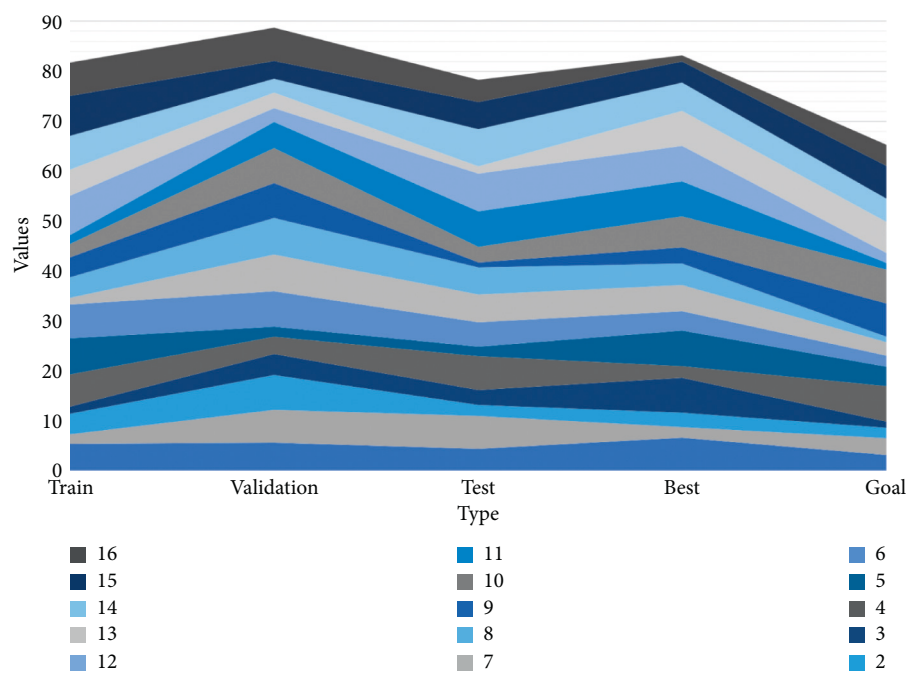


Figure 6: Plot of the inspection of the effect of training via the network.

particularly widely, and the number of positive and negative errors is the same, except for some special points where the error is greater than [−1, 1], these special points may be an artificial collision or sudden external noise interference during the test, which are uncontrollable factors, while the rest of the points are within this range. For a more uniform analysis of the error performance, Figure 7 illustrates the percentage distribution of the BP neural network state evaluation error. From the figure, the maximum percentage error is less than 15%, most of the errors are less than 10%, and the errors are evenly distributed. The error plot analysis shows that for the practical application of the BP neural network for state assessment research, engineers, and researchers must consider the existence of assessment errors. Although the neural network-based condition assessment achieves easy and fast identification, before the result is identified, other means including local detection, human observation, etc. must be used to reconfirm the degree of damage and analyze the condition level to prevent the impact of these two types of errors on human life and property.

4.2. Experimental Results of Landscape Ecological Planning Assessment. Ecosystem service values showed a weak correlation with patch density (PD), patch average area (AREA-

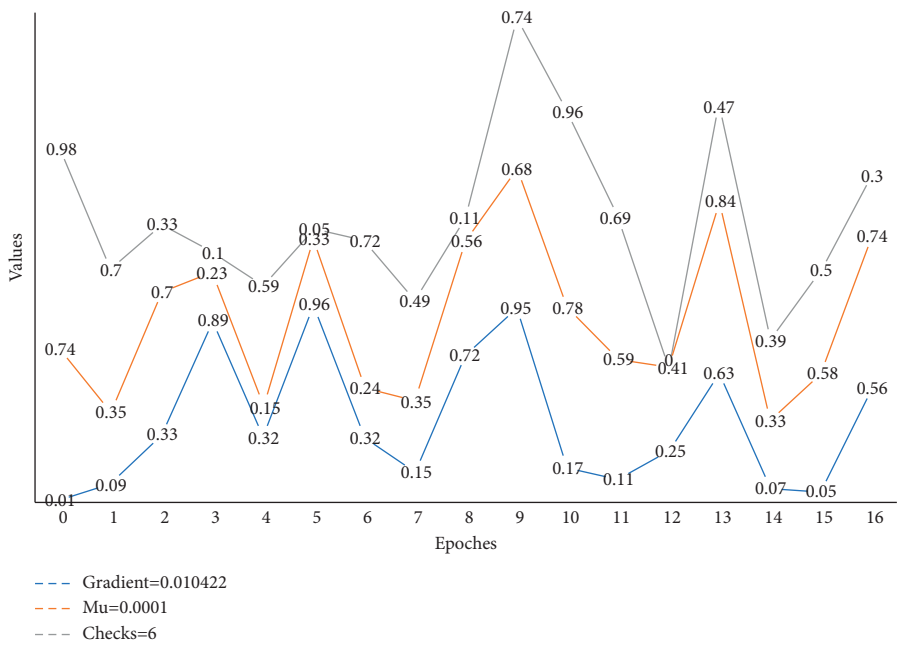


Figure 7: Neural network training state effect graph.

MN), and aggregation (AI), with correlation coefficients of 0.261, 0.442, and 0.665, respectively, at the 0.01 level. The average area of the patches responded to the fragmentation complexity of the landscape pattern, and its fragmentation degree will directly affect the size of the ecosystem service value of the whole study area. The fragmentation of the landscape pattern is related to the land use type, and the fragmentation of woodland and water area is increasing year by year, and the area of woodland and water area is decreasing year by year due to human interference and urbanization, which makes the value of ecosystem services decrease year by year. Therefore, to achieve sustainable development, it should be helpful to protect woodland and watershed ecosystems so that their fragmentation can be reduced and the ecosystem service value of the whole study area can be increased. The value of ecosystem services was not significantly correlated with the correlation coefficient of 0.032 for the walk and juxtaposition index, which describes and calculates the distribution characteristics of ecosystems among various types of blocks. The correlation coefficient indicates that there is no significant correlation between the walk and juxtaposition indices for each tessellation.

There is a strong correlation between ecosystem service values and land use landscape patterns. Changes in land use directly affect changes in landscape patterns and thus indirectly affect changes in ecosystem service values. To improve the value of regional ecosystem services, it is necessary to promote and implement a sustainable economic policy aimed at protecting the environment, realize a social-ecological-economic circular ecological-economic model, and promote the rapid development of the regional economy and society, as shown in Figure 8.

The overall trend of carbon storage in woodlands, water bodies, and built-up areas is increasing. Due to the policy of "two rounds" of returning farmland to forest and the increase of abandonment of rural sloping land, the landscape type has changed more obviously, which has prompted the recovery of woodlands and the advantage of forests in carbon sequestration, resulting in the continuous increase of carbon sequestration in woodlands in the reservoir area. Of course, the increase in carbon sequestration in water bodies does not mean an increase in their carbon sequestration function, but the increase in water bodies in the reservoir area, which floods the original arable land, forest land, and grassland. The built-up area itself has a low carbon storage capacity, and although it has increased to a certain extent during the study period, the expansion of the built-up area inevitably leads to the loss of carbon storage function of other land types from the perspective of the whole study area. In contrast, the total carbon storage in paddy fields, drylands, grasslands, and unused lands tends to decrease. In addition to the decrease in area, the decrease in carbon storage in these three landscape types is also due to the declining carbon sequestration capacity in the conversion of landscape types. It is mainly used in the fields of safety production science and environmental science. Some scholars often combine AHP and fuzzy comprehensive evaluation to study problems in more fields.

From the current situation of wetland parks, woodland and water bodies are the dominant landscape in wetland parks, and the distribution of water bodies concentrated and woodland is relatively balanced, the distribution of water bodies and woodland is the most obvious; the area of scrub and construction land is generally lower, and their edge effect is also lower, which is caused by their small area or more homogeneous shape. The edges of woodland and roads are generally more complex in wetland parks, and the edges of Belus Bay Wetland Park and the edges of roads in Jincheng Lake Wetland Park are the most complex, as shown in Figure 9.

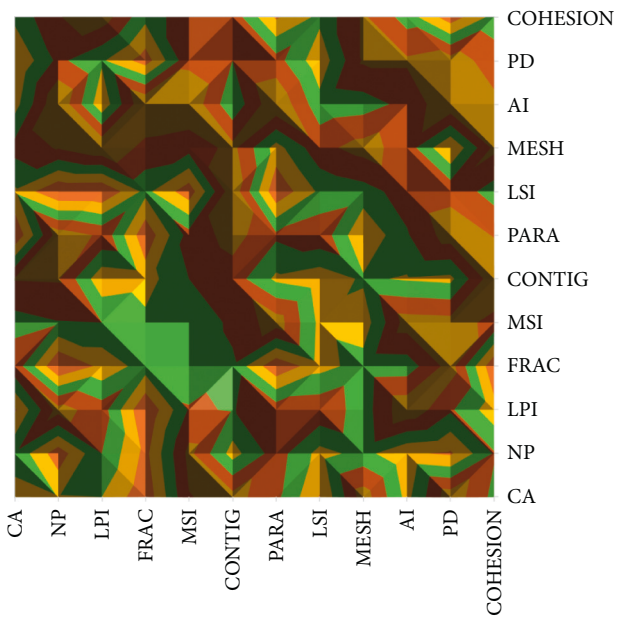


Figure 8: Results of Pearson correlation analysis of landscape type indicators.

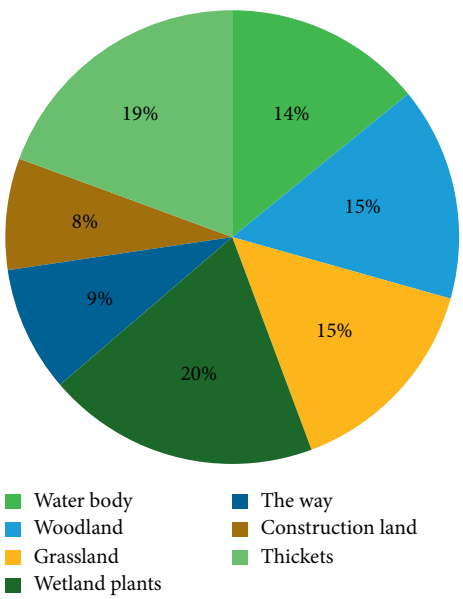


Figure 9: Reference value of patch type area share.

According to the analysis of the indices of patch type area and density, in addition to roads, the accessibility of its woodland, grassland, and scrub is also higher, and the modularity of roads and construction land is obvious, with roads accounting for 17.5% of the landscape area, which is greater than the average of the three parks' road area ratio of 9%, and construction land accounting for 9.3% of the landscape area ratio, which is higher than the average of the three parks' 4.8% nearly twice, which indicates that its service facilities. This indicates that its service facilities and recreation sites are better and can provide more public service space for the public, but it also increases the degree of landscape fragmentation in the park, with a patch density of 2.5 patches per hectare, which is higher than the average of 1.6 patches per hectare in the three wetland parks. In addition to roads and construction land, the park is not high in terms of area, aggregation, and combination of the remaining landscape types, which are more evenly distributed, and its landscape level has better landscape diversity, higher landscape uniformity, and even distribution of landscape pattern, and richer and more diverse landscape style.

5. Conclusion

This paper is guided by the main ideas and theories of garden arboriculture, urban green space system planning, plant landscape planning and design, plant community science, landscape ecology, operations research, and other disciplines. The water-adapted landscape is a traditional ecological heritage that cannot be ignored in building a spongy country and is a spongy landscape with vernacular regional characteristics. The water-adapted landscape, because of its close integration with the natural and human characteristics of the region, has a high regional human connotation and contains the traditional sponge experience and wisdom of adapting to regional hydrological changes. These have important implications for the construction of modern sponge cities, especially when the development site is closely linked to the regional water-adapted landscape, which should be considered in the planning and design of rainwater. The main logic of the framework for evaluating the service performance of the ecological spatial structure is to study whether the ecological spatial structure can efficiently respond to and support the necessary landscape ecological service process, to evaluate the key components of the ecological spatial structure, including the service supply area and the service association area, and to analyze whether the natural and pattern properties of the two parts can effectively support the landscape ecological service process. The service process is the criterion to judge whether the ecological spatial structure can efficiently and sustainably supply the landscape ecological services, and then reflect the service performance level of the ecological spatial structure.

Data Availability

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

Conflicts of Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work was supported by the School of Xiamen University of Technology.

References

[1] L. Mei and L. Qi, "Central urban open space system and green economy planning based on spatial clustering algorithms and AHP model," *Journal of Intelligent and Fuzzy Systems*, vol. 39, no. 4, pp. 5785–5795, 2020.

[2] E. Shirgir, R. Kheyroddin, and M. Behzadfar, "Defining urban green infrastructure role in analysis of climate resiliency in cities based on landscape ecology theories," *TeMA Journal of Land Use, Mobility and Environment*, vol. 12, no. 2, pp. 227–247, 2019.

[3] T. Xu, J. Gao, and G. Coco, "Simulation of urban expansion via integrating artificial neural network with Markov chain – cellular automata," *International Journal of Geographical Information Science*, vol. 33, no. 10, pp. 1960–1983, 2019.

[4] H. Siroosi, G. Heshmati, and A. Salmanmahiny, "Can empirically based model results be fed into mathematical models?" *Environment, Development and Sustainability*, vol. 22, no. 4, pp. 3701–3722, 2020.

[5] S. Chen, H. Zhao, Z. Xu, W. Xie, L. Liu, and Q. Li, "Landslide risk assessment in Nanping City based on artificial neural networks model," *Chinese Journal of Geological Hazard and Control*, vol. 33, no. 2, pp. 133–140, 2022.

[6] W. Zhang, "A method for monitoring the spatial distribution of nutrients from the perspective of landscape ecology," *Journal of Coastal Research*, vol. 95, no. sp1, pp. 182–186, 2020.

[7] R. Huang, Y. Nie, L. Duo, X. Zhang, Z. Wu, and J. Xiong, "Construction land suitability assessment in rapid urbanizing cities for promoting the implementation of United Nations sustainable development goals: a case study of Nanchang, China," *Environmental Science and Pollution Research*, vol. 28, no. 20, pp. 25650–25663, 2021.

[8] Z. Kang, S. Wang, L. Xu, F. Yang, and S. Zhang, "Suitability assessment of urban land use in Dalian, China using PNN and GIS," *Natural Hazards*, vol. 106, no. 1, pp. 913–936, 2021.

[9] M. Talebi, B. Majnounian, M. Makhdoum, E. Abdi, and M. Omid, "Predicting areas with ecotourism capability using artificial neural networks and linear discriminant analysis," *Environment, Development and Sustainability*, vol. 23, no. 6, pp. 8272–8287, 2021.

[10] A. Tiwari, M. Shoab, and A. Dixit, "GIS-based forest fire susceptibility modeling in Pauri Garhwal, India: a comparative assessment of frequency ratio, analytic hierarchy process and fuzzy modeling techniques," *Natural Hazards*, vol. 105, no. 2, pp. 1189–1230, 2021.

[11] E. Ustaoglu and A. C. Aydınoglu, "Land suitability assessment of green infrastructure development," *TeMA-Journal of Land Use, Mobility and Environment*, vol. 12, no. 2, pp. 165–178, 2019.

[12] X. Liang, X. Liu, G. Chen, J. Leng, Y. Wen, and G. Chen, "Coupling fuzzy clustering and cellular automata based on local maxima of development potential to model urban emergence and expansion in economic development zones," *International Journal of Geographical Information Science*, vol. 34, no. 10, pp. 1930–1952, 2020.

[13] J. Zhou and P. Wang, "Image simulation of urban landscape in coastal areas based on geographic information system and machine learning," *Neural Computing & Applications*, vol. 34, no. 12, pp. 9397–9411, 2022.

[14] N. Nuthammachot and D. Stratoulias, "Multi-criteria decision analysis for forest fire risk assessment by coupling AHP and GIS: method and case study," *Environment, Development and Sustainability*, vol. 23, no. 12, pp. 17443–17458, 2021.

[15] C. Wang, Y. Tang, M. A. Kassem, H. Li, and C. Yan, "BP neural network for typhoon warning and prevention for historical buildings," *Arabian Journal for Science and Engineering*, vol. 47, no. 4, pp. 5237–5254, 2022.

[16] N. Erdenejargal, B. Dorjsuren, L. Choijinjav et al., "Evaluation of the natural landscape aesthetic: a case study of Uvs province, Mongolia," *Polish Journal of Environmental Studies*, vol. 30, no. 5, pp. 4497–4509, 2021.

[17] Q. Niu, L. Yu, Q. Jie, and X. Li, "An urban eco-environmental sensitive areas assessment method based on variable weights combination," *Environment, Development and Sustainability*, vol. 22, no. 3, pp. 2069–2085, 2020.

[18] Z. Khorrami, T. Ye, A. Sadatmoosavi, M. Mirzaee, M. M. Fadakar Davarani, and N. Khanjani, "The indicators and methods used for measuring urban liveability: a scoping review," *Reviews on Environmental Health*, vol. 36, no. 3, pp. 397–441, 2021.

[19] L. Singh, S. Saravanan, J. J. Jennifer, and D. Abijith, "Application of multi-influence factor (MIF) technique for the identification of suitable sites for urban settlement in Tiruchirappalli City, Tamil Nadu, India," *Asia-Pacific Journal of Regional Science*, vol. 5, no. 3, pp. 797–823, 2021.

[20] G. Xu, X. Zhu, and N. Tapper, "Using convolutional neural networks incorporating hierarchical active learning for target-searching in large-scale remote sensing images," *International Journal of Remote Sensing*, vol. 41, no. 11, pp. 4057–4079, 2020.

[21] D. Verma, A. Jana, and K. Ramamritham, "Classification and mapping of sound sources in local urban streets through AudioSet data and Bayesian optimized Neural Networks," *Noise Mapping*, vol. 6, no. 1, pp. 52–71, 2019.

Hindawi
Computational Intelligence and Neuroscience
Volume 2022, Article ID 1158509, 10 pages
https://doi.org/10.1155/2022/1158509

Research Article

A Comprehensive Quantitative and Biological Neural Network Optimization Model of Sports Industry Structure Based on Knowledge Mapping

Zhaohong Wang,[1] Yang Gao ,[1] and Wenge Li[2]

[1]*College of Physical Education and Sports, Beijing Normal University, Haidian District, Beijing 100875, China*
[2]*Kehua School, Nanshan Foreign Language School, Shenzhen, Guangzhou 518057, China*

Correspondence should be addressed to Yang Gao; 201621070009@mail.bnu.edu.cn

Received 12 July 2022; Revised 1 August 2022; Accepted 9 August 2022; Published 29 August 2022

Academic Editor: Ning Cao

In this paper, a comprehensive quantitative and biological neural network optimization model of sports industry structure is thoroughly studied and analyzed using knowledge graphs. To address the problems of poor performance interpretability deficiency of knowledge graph-based recommendation methods in the face of relational sparse graphs, a pretraining-based implicit characterization algorithm strategy is proposed for the recall stage, which can solve the problems of difficulty in going online and high delay in the recall stage of the recommendation system while improving the accuracy, and not only this can be applied in the recall stage, but also the sorting and postsorting modules can be used as features. To study the relationship between signaling activity and energy metabolism of pyramidal neurons, an empirical model of the synaptic vesicle cycle is proposed to simulate the synaptic transmission process, the role played by energy metabolism in synaptic transmission is studied from the perspective of feedback control, and the quantitative relationship between neuronal pulse discharge frequency, energy consumption, and information quantity in dendritic integration is analyzed using the cable theory and atrial chamber model. It was found that, when $0 \leq \varepsilon \leq 0.6$, the chaotic region shrinks and eventually disappears with the increase of the memory factor ε; however, when $0.6 \leq \varepsilon \leq 1$ is used, chaos is recreated and the chaotic area gradually increases with the increase of the memory factor ε. This paper conducts comparative experiments on data sets in the recommendation domain and verifies that the proposed model and the feature intersection module can effectively perform feature interaction between items and entities, thus enhancing the recommendation effect.

1. Introduction

The existing studies often have poor recommendation performance when facing graphs with sparse relationships and lack certain interpretability. Meanwhile, as the link between knowledge utilization and creation, the academic field is flooded with a large amount of knowledge and recommendation systems need to be introduced to solve the information overload problem. Increased platforms are providing services for human beings through the Internet, almost all over the clothing, food, and housing of people's lives [1]. Since the birth of the Internet, the development of related technologies has been advancing all the time, and people can access more information by going online to fully enjoy the convenience of the Internet. Internet technology has brought unprecedented changes to all aspects of people's lives [2]. The way people buy goods has changed from offline to online without leaving home, the way they taste food has changed from dining out to delivery, and the sales method of each merchant has changed from local sales in physical stores to online sales in multiple ways; but, at the same time, people are also in an era of hyper-information, and it is difficult for them to choose what they want from the huge amount of information. In recent years, many researchers have fused recommendation systems with relevant auxiliary information to improve the

effectiveness of recommendations, and knowledge graphs are one of the most popular types of structural information. Knowledge graphs cover different domains, such as academics, films and television, music, etc., which makes it very convenient to use domain knowledge graphs to assist recommendation systems.

In recent years, with the rapid development of neuroscience, artificial intelligence, and other technologies, the great potential of brain science has been highly valued, and the China Brain Project, which has been in the making for many years, has been launched. The development of brain science not only contributes to the prevention and treatment of brain diseases but also promotes the development of brain-like intelligence [3]. In the field of brain science, the simulation modeling of biological neurons and their networks occupies an important position and is a popular research direction in the field of brain science. On the one hand, simulation modeling of biological neurons and their networks helps to understand the working principle of the nervous system from a global perspective and to recognize the role played by various neural mechanisms in the operation of the brain; specifically, due to the limitations of experimental conditions and observation techniques, the physiological activities of a small number of neurons can only be observed simultaneously at mesoscopic or microscopic scales, and it is difficult to observe the physiological activities of large-scale biological neural networks at the same time. In addition, simulation modeling can study the effect of certain factors on neural network activity by adjusting parameters, whereas biological experiments require the creation of many control and experimental groups to accomplish this task, which is more flexible and convenient [4]. On the other hand, simulation modeling of biological neurons and their networks helps artificial intelligence technologies break through bottlenecks and move toward strong artificial intelligence. The goal of artificial intelligence technology is to produce machines that can simulate human thought processes and intelligent behaviors, or even surpass human intelligence. A prerequisite for achieving this goal is to master how the brain works and then simulate the brain's cognitive functions. The success of current AI technologies is largely due to mathematical optimization and is far from the workings of the real brain. Although AI technologies have been widely used in various fields, their level of intelligence is still far from human intelligence. The focus of simulation modeling of biological neurons and their networks is to explore and learn from the working principles of the brain rather than mathematical optimization. Therefore, simulation modeling of biological neurons and their networks can help current artificial intelligence break the shackles and improve the level of intelligence.

With the steady and rapid development of the economic environment, the sports industry is receiving higher and higher attention, both in terms of scale and quality, which is most directly reflected in the number of relevant employees and added value. The value-added created reached 325.59 billion yuan, accounting for 0.62% of the total national GDP in that year, maintaining an increased rate of more than 10% for ten consecutive years, which is significantly higher than the growth rate of the overall national GDP. The sports industry has grown into a new hotspot of economic growth, especially in some developed eastern provinces and cities, the scale of the industry and its ability to absorb employment has been no different from the level of medium developed countries in Europe and the United States at the end of the last century [5]. The analysis and research of industrial development based on the theory of industrial structure is one of the research hotspots of global scholars in recent years. At the same time, with the development of society, the sports industry is facing new development opportunities and challenges, among which the industrial structure is one of the bottlenecks. Objectively speaking, compared with Western countries, our research in this field has started relatively late and we have not systematically evaluated and analyzed the current situation of sports industry structure development from a macro perspective [6]. With this background, the author firstly composes the existing theoretical concepts related to the optimization of sports industry structure, then makes a more comprehensive and objective evaluation of them, lists the objective problems based on the in-depth understanding of the development process and change trend of the sports industry, and finally proposes feasible strategies for the optimization of sports industry structure from a practical perspective. The final proposal is a feasible strategy for optimizing the structure of the sports industry from a practical perspective. On the one hand, it provides the necessary theoretical basis for the government to formulate the development plan for the sports industry, and on the other hand, it makes an important contribution to the realization of the dream of a "strong sports nation."

2. Related Works

With the development of knowledge graphs, it has become much more than a simple semantic network, but a technical system containing entities, attributes, concepts, and various rich semantic relationships. Compared with traditional semantic networks, knowledge graphs have a huge scale, rich semantic information, refined quality, and a more friendly structure [7]. The knowledge graph is a large-scale semantic network that generates new knowledge by acquiring and integrating information into a knowledge base and then reasoning about many entities, attributes, and semantic information between entities. As a technical system, knowledge graphs have a wide range of applications in cognitive intelligence fields such as data analysis, search engines, recommendation systems, and human-computer interaction, and corresponding research projects related to knowledge graphs have been conducted in domestic commercial fields as well as academic fields. The knowledge graph is regarded as a third-order tensor, and low-dimensional entity embedding is used, and low-dimensional entity embedding and relationship embedding are used to reduce the third-order tensor. Neural network-based and graph neural network-based approaches, such as R-GCN, fuse entities with information about other entities associated with

them through graph convolutional neural networks to obtain a semantic representation of the target entity [8]. Embedding-based approaches use the translational distance model to implement entity embedding, obtain the corresponding entity representation through embedding, and use the entity representation to achieve recommendation. Yang et al. proposed an embedding-based collaborative knowledge recommendation approach (CKE) in 2016, which extracts structured content, textual content, and visual content from the knowledge graph and uses the TransR model, noise reduction autoencoders, and convolutional autoencoders to embed these three types of content representations [9]. Qingwen et al. in 2019 proposed a multitask feature learning recommendation model (MKR), which designs a cross-compression unit that enables knowledge sharing between the embedding part and the recommendation part and implements recommendations accordingly [10]. The fusion-based approach takes the entity relationships of the knowledge graph and optimizes the item or user vector in the recommendation algorithm by convolutional neural networks, graph neural networks, etc. Baskonus et al. simulated the propagation process of user interests on the knowledge graph and portrayed the user preferences in more detail using the relational information within the knowledge graph [11]. The recommendation methods based on knowledge graphs have different advantages and disadvantages, but they generally have the problem of unsatisfactory recommendation effect when dealing with knowledge graphs of sparse relationships.

As a sunrise industry supporting the rapid development of the national economy, the internal structure of the sports industry is also inextricably linked. It is especially important to follow the rules of economic development and sports industry development to have a comprehensive understanding and grasp of its industrial structure. As an indispensable and important part of the sports industry, the importance of the structure of the sports industry has become increasingly prominent. After years of development and precipitation, the academic research results on the current situation of sports industry structure are richer, with the foothold of research rising from the perspective of the whole country, as well as combining the actual situation of field investigation in provinces, cities, and autonomous regions, and comprehensive analysis of the sports industry from the overall perspective, as well as a study on its composition and development trend from a micro perspective. According to Gupta and Chandra, the structure of the sports industry refers to the technical and economic links and quantitative proportional relationships among the production sectors in the industry [12]. On the one hand, it can comprehensively reflect the interdependence and mutual constraints in production technology among the production sectors of sports goods and sports services, and on the other hand, it can reflect the current allocation of various economic resources in each sector and the specific distribution of the total output value of the sports industry in each sector. There are still some scholars who define the structure of the sports industry from different disciplinary perspectives. Novak et al., from the perspective of sports consumption, believe that the structure of the sports industry is the arrangement, elements, and configurations of various sectors in the context of the sports industry, which needs to optimize the allocation of sports resources in a market-oriented way to meet the objective needs of the public as much as possible [13]. From a macro perspective, some scholars believe that the structure of the sports industry is "the linkages between sports industries and the contacts they form." It is known that there are also structural links in the internal structure of the sports industry. It should be noted that, when studying the internal links of the structure of the sports industry, it should be assumed that the external conditions remain unchanged and the ratio of inputs and outputs within the sports industry should be considered as the focus of research. Ren et al. analyzed the contradiction between supply and demand in the structure of China's sports industry and points out that the core part of China's sports industry accounts for less than 20% and is still in the primary development stage, and the main problem facing the supply and demand of China's sports industry structure is the insufficient supply of sports products and the ineffective supply of some sports goods [14].

Neurons are the structural units of the nervous system, which is made up of many interconnected neurons. Neural behavior is not determined by individual neurons but is the result of the interaction of neuronal populations. Therefore, studying the topological properties and structural evolution of the nervous system, clarifying how neurons establish synaptic connections, how synaptic connections change, and other fundamental issues are necessary prerequisites for mastering the operating mechanism of the brain. In the research process, biological neural network simulation modeling is an indispensable technical tool. In terms of identification methods, two main types of methods are currently used for the identification of biological neural network systems [15]. The first is to use adaptive state observer for identification. By constructing a following system with the same structure and unknown parameters to be identified, and based on Liapunov stability theory, a parameter adaptive adjustment equation is constructed, so that the parameters to be identified converge to the actual parameter value when the time tends to infinity.The second is to identify modern optimization algorithms based on strong optimization algorithms. The identification problem of biological neural network system is transformed into an optimization problem by constructing an appropriate objective function. Modern optimization algorithms are used to find the best solution, so as to obtain a more satisfactory solution through continuous iterative search, that is, to identify unknown parameters or unknown topologies. The simulation modeling of biological neural networks refers to the creation of biologically interpretable network models that can elaborate certain properties or functions of the nervous system based on the current knowledge of the nervous system, such as the study of synaptic plasticity, topological properties, and structural evolution of the nervous system. Jinfeng and Bo confirmed the validity of Hebb's theory [16]. Chen et al. investigated the molecular mechanisms underlying LTP and LTD and showed that most glutamatergic

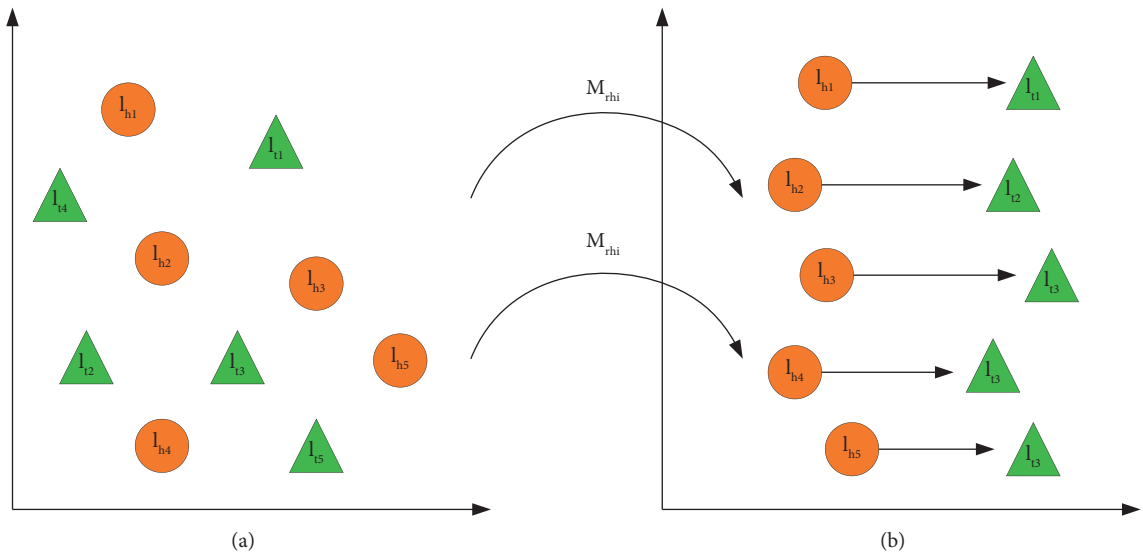


Figure 1: TransR mapping relationship diagram: (a) entity space and (b) relative space of r.

synapses exhibit LTP and LTD concerning the activation of NMDA receptors on the postsynaptic membrane and calcium ion concentration [17].

3. Construction of a Comprehensive Quantitative and Biological Neural Network Optimization Model of Sports Industry Structure Based on Knowledge Mapping

3.1. Knowledge Graph Model Design. Knowledge graph construction technology is a complex technical system that requires knowledge extraction, knowledge fusion, knowledge processing, and other techniques to support it. Starting from heterogeneous data and ending with the formation of a knowledge graph, many key technologies are involved in the middle, including knowledge acquisition, processing, and structured representation [18]. Knowledge extraction refers to extracting valuable knowledge from multiple heterogeneous information sources. Knowledge fusion is to resolve the ambiguity of knowledge by using related technologies and to form a standard knowledge base by fusing multiple knowledge bases. Knowledge processing is the process of inferential representation of knowledge, which determines the final quality of the knowledge graph. The stability of the single-task model is insufficient. ESMM considers everything from exposure to click to conversion.

$$I = \{u, U_s, S\}. \tag{1}$$

Knowledge fusion is mainly used to solve the problem of the heterogeneous knowledge graph, and it is mainly realized through entity alignment. According to the scholar's information obtained from unstructured data and semistructured data, it is necessary to clean them first, then construct triples according to the cleaned data, and finally align the two groups of triples data through entity alignment. The data obtained from unstructured data are extracted by using a relatively mature algorithm, and the accuracy of entity recognition is also high. Therefore, the main object of data cleaning is the data obtained from semistructured data. As shown in Figure 1, it is indicated that TransR uses the relation-specific matrix. Maps$M^r h$and$t$to subspaceR^kas follows:

$$V_h^r = v_h M^r, v_t M^r. \tag{2}$$

The EL is debugged through predefined entities of user reviews and knotted food for a specific product, using DBpedia as the reference set . First, the NER system was used to distinguish between structured and unstructured texts and filter the DBpedia for irrelevant entities. Of course, many publicly available entity identification tools can also be used. The scoring functions of the embedded technology reference are classified into the translation distance model based on the distance scoring function and the semantic matching model based on the similarity scoring function [19]. The process of knowledge graph construction is shown in Figure 2.

For the semistructured data of Scholars.com, certain extraction rules need to be designed to extract the data from the Web pages automatically; while for the unstructured data, the corresponding algorithm needs to be used to extract them and the BERT-BiLSTM-CRF model with excellent stability and accuracy in recent years is chosen here. After obtaining the two data, we need to clean the data, then construct the triad data. After the knowledge fusion is completed, certain id mapping relationships are set for the triad and the entity. Finally, the entities and relationships are stored in Neo4j in bulk using the import method. At this point, the nodes, relationships, and the knowledge graph of scholars formed by them can be viewed in the graph database. Due to the differences in the structured of semistructured scholar data, it is obvious that extraction using algorithms is not easy to implement. Therefore, it is necessary to design the corresponding entity extraction rules by

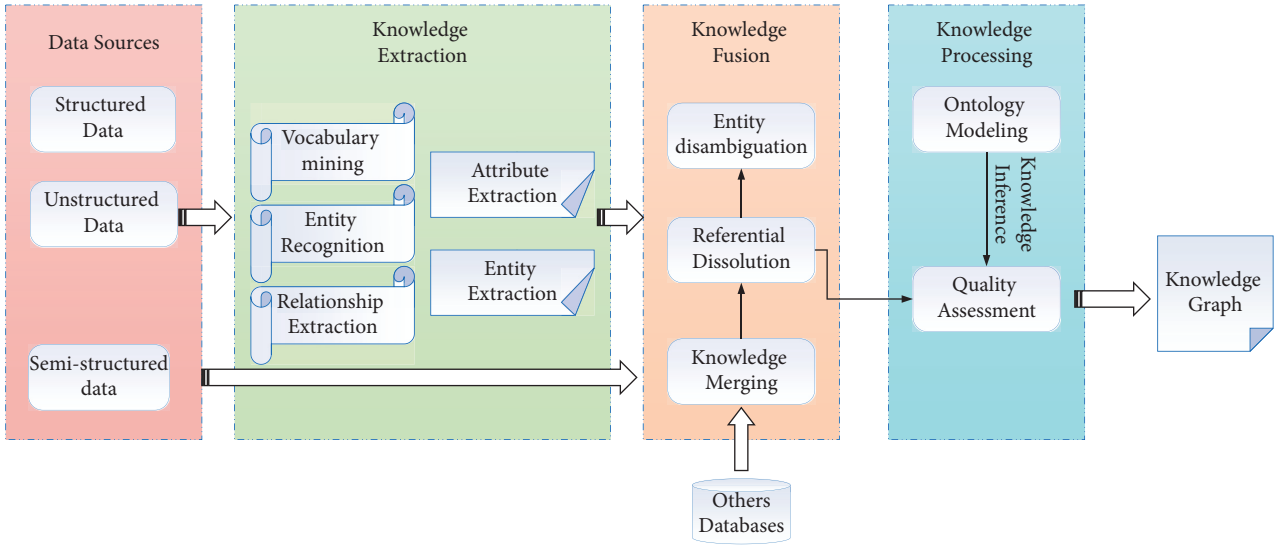


Figure 2: Flow chart of knowledge graph construction.

hand and use the rules to extract the semistructured data from the Web pages and transform them into structured scholarly data.

3.2. Comprehensive Quantitative and Biological Neural Network Optimization Model Construction for Sports Industry Structure. Considering the complexity of the sports industry, to fully analyze the structure of the sports industry, we should recognize the links between the sports industry and other industries, as well as the links between the elements of interest within the sports industry. On the one hand, the direct structural size of different elements of the sports industry in terms of the scale of "input-output" specifically expresses the characteristics of the structural volume of the sports industry; and on the other hand, the arrangement of different elements in the industrial content has different structural characteristics. The structural form of the sports industry is specifically expressed in the different levels of the industrial structure [20]. Considering the problem of the extensiveness and complexity of different industrial links, we should analyze and study the structure of the sports industry from multiple levels and perspectives in the research process. Accordingly, we can also analyze the structural form of the sports industry from the following aspects: industry structure, hierarchical structure, organizational structure, ownership structure, market structure, regional structure, etc.

The sports service industry includes other sports industries except for sports goods and manufacturing and sports venue facilities construction, including a total of seven categories; the proportion of each category is relatively small, and the quality of supply and service efficiency of the sports service industry are not high enough. Therefore, after more than a decade of growth and rapid development, China's sports industry is growing and maturing, with rapid growth in scale, continuous improvement of the industrial system, and increasing enrichment of industrial categories. China's sports industry is

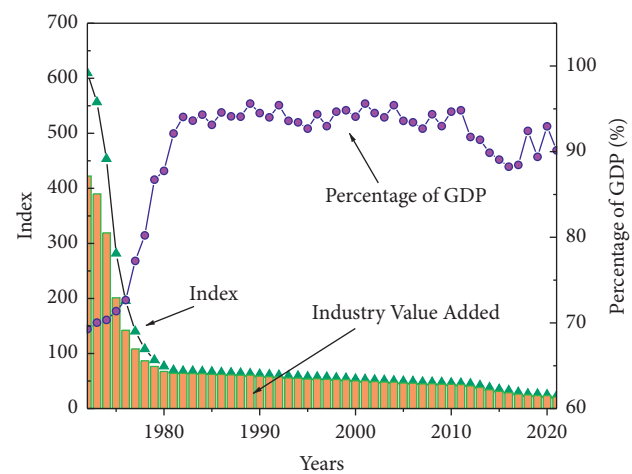


Figure 3: Value added of the national sports industry and proportion of GDP.

growing rapidly and significantly higher than the growth rate of the macro economy in the same period, with a total scale of more than 1 trillion yuan, and the proportion of China's GDP is also increasing, accounting for about 1.14% of GDP in 2019. As shown in Figure 3, it reflects the value added of China's sports industry and its share in GDP in recent years.

Based on the continuous optimization and adjustment of the sports industry structure, we have transformed the quantitative and economic-technical links between sports industries from the previous uncoordinated state to the current coordinated state. In other words, the restructuring of the sports industry involves two different aspects: market adjustment and government regulation. Among them, the market adjustment belongs to the category of natural adjustment, while government regulation belongs to the category of artificial adjustment. In the current market economy, we mainly do the optimal allocation of market resources among sports industries under the role of market value law; at the same time, from the government level, the

government mainly adjusts industrial policies and other macro-control methods and then can provide institutional support for the healthy development of sports industry from the institutional level and effectively promote the rationalization of sports industry development level. This means that changes in cell membrane potential do not occur synchronously, and local membrane potential changes can cause a chain reaction across the cell membrane.

This ionic osmotic pressure together with the unique ionic permeability of the cell membrane leads to the generation of the cell membrane potential, which in turn indirectly changes the ionic permeability of the cell membrane and the ionic osmotic pressure inside and outside the membrane by affecting the opening and closing state of voltage-gated ion channels [21]. The cell membrane potential can be divided into resting and action potentials, as shown in Figure 4. When a biological neuron is not stimulated, the ionic osmotic pressure inside and outside the cell membrane, the cell membrane potential, and the ionic permeability of the cell membrane are in equilibrium, and the cell membrane potential is in a stable state at this moment is called the resting potential. It is found that the potassium channels in the cell membrane are open and the sodium channels are closed, the high potassium ion osmolarity inside the cell membrane and the positive voltage outside the cell membrane cancel each other out, and the cell membrane potential is stable at about 65 mV.

Biological neurons have a complex morphological structure, and their dendrites, cell bodies, and axons may be distributed far apart in space, and stimulation of the cell membrane often occurs in local areas, which means that changes in the cell membrane potential do not occur synchronously, and local membrane potential changes can cause a chain reaction in the cell membrane. For example, a stimulus signal causes a change in membrane potential in a local region of the dendrites of a biological neuron, which in turn disrupts the equilibrium in the neighboring region, which in turn causes a change in membrane potential in the neighboring region. Because of the spatial differences in the morphological structure of biological neurons and the ion permeability of cell membranes, and because biological neurons usually receive many stimulus signals simultaneously, these stimulus signals are constantly scaled, filtered, fused, and produce other complex nonlinear changes when they are transmitted within biological neurons. Static neurons are neurons whose state does not change over time.

$$f(x) = \sum_{i=1} w_{ij} + f(x_{\mathrm{i}}) - \theta_j, \tag{3}$$

where $x_i (i = 1, 2, \ldots, m)$ is the incoming signal of other neurons, θ_j is the bias of neuron, w_{ij} is the connection weight of the i-th neuron and j-th neuron, f is the activation function, the sigmoid function is one of the commonly used activation functions, and $f(x)$ is the output of the neuron. The structure of the static neuron model is very simple, the weighted summation of the input stimulus simulates the information integration function of the cell body of the biological neuron, the bias θ_j simulates the threshold potential of the biological neuron, and the nonlinear activation function simulates the nonlinear properties of the biological neuron.

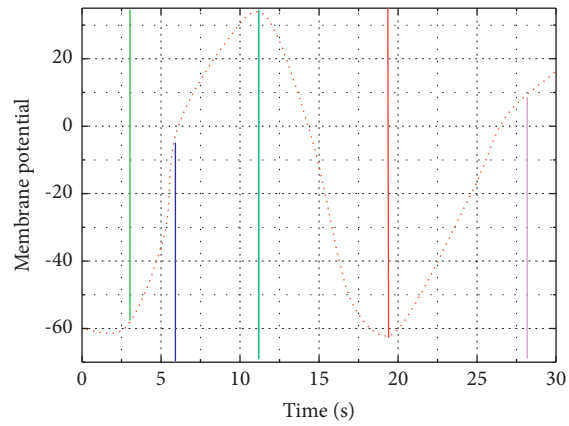


Figure 4: Mechanism of action potential generation.

4. Analysis of Results

4.1. Knowledge Graph Model Results. Compared with traditional semantic network, knowledge graphs have the characteristics of huge scale, rich semantic information, excellent quality, and more friendly structure. The core of the advanced industrial structure is the innovative use of high-level production factors such as science and technology and the optimization of the transformation capacity of the industrial structure. The process of the heightening of the sports industry structure is due to the different degrees of productivity improvement in various sports industries or the different income elasticity of demand for sports products, resulting in the differentiation of the development speed of various sports industries and changes in their dominant positions, resulting in a new adjustment of the sports industry structure. In addition, the best entity embedding dimension will be selected $\{32, 64, 128, 256, 512\}d$; the best depth of the propagation layer will be selected $\{1, 2, 3, 4, 5\}l$. Each set of experiments will be repeated 10 times, and the average of the best results will be taken. As shown in Figure 5, HoPKG also performs best in Rec@K and Pre@K compared with other methods. The reason PER and MCRec do not perform as good as other methods is that they require manual construction of meta-paths, which will lead to some uncertainty factors. Therefore, the performance is more general when the data set is sparse. In addition, different aggregation methods can produce different results. The dual aggregation method BiPart proposed in this paper has some improvement over GraphSAGE in both AUC and F1 since the HoPKG model combines entity representations aggregated in two different ways and can propagate entity information more effectively.

MF can only characterize features and end targets and has poor performance in constructing contexts. nFM treats user-goods interactions as features and models them with linear neural networks and nonlinear multilayer perceptrons, thus enabling better predictions than FM [22]. However, due to single-task modeling, there are many

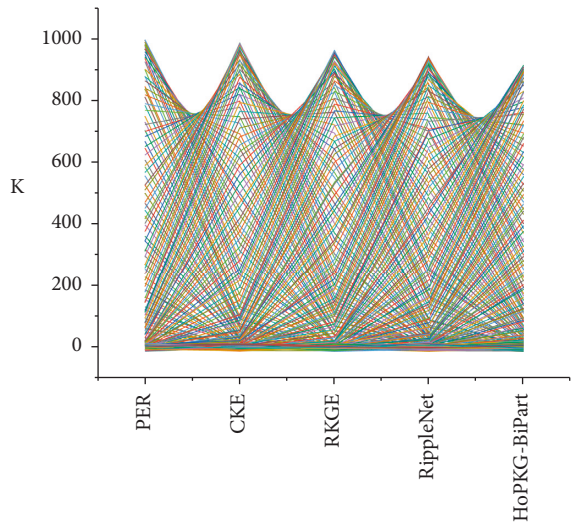


Figure 5: Performance comparison results of the model.

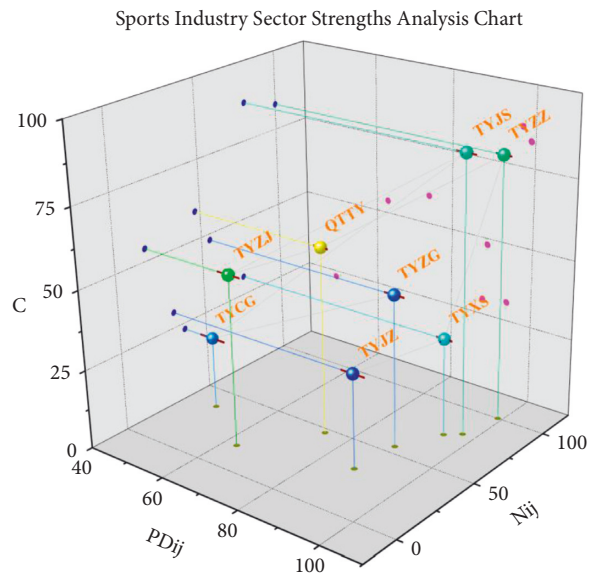


Figure 6: Analysis of the advantages of the sports industry sector.

possibilities for the distribution in the links, resulting in sparse samples of depth events and large fluctuations in the distribution of depth events when performing single-task modeling of deep links. The stability of the single-task model is not sufficient. eSMM considers the transition from bursts to clicks and back again. Modeling is only applicable to the e-commerce case. The set-up assumption is $ctcvr = ctr \times cvr$, but this may not be the case in real situations. The results suggest that appropriate neural networks can improve feature learning in the user interaction information. The performance of ESMMV2 is mainly due to its deeper modeling links and the addition of more mid-question states to assist in modeling. Again, the modeling still depends on the strong correlation between the deep and shallow tasks that complete the modeling. The changing correlations of user intent are not fixed, and the performance in terms of metrics is not satisfactory.

4.2. Comprehensive Quantitative and Biological Neural Network Optimization Model Performance Test of Sports Industry Structure. The sectoral analysis chart is a system of coordinates with the deviation component PD_{ij} on the horizontal axis and the share component N_{ij} on the vertical axis, in which the scatter points representing each industry sector are marked. The greater the value, the greater the ability of the region's industrial sector to contribute to the total volume of the sports industry. Nissl bodies are where proteins are synthesized and are involved in the synthesis of the neurotransmitter acetylcholine. The essence of neurofibrils is the aggregation of neurofilaments and nerve microtubules in neurons when they are fixed, and their functions are related to material transport and axonal growth.

Among the eight sectors of the sports industry, the competitive weight of "sports goods, clothing, footwear, and hat manufacturing," "sports construction activities," "sports goods, clothing, footwear, and hat sales," "sports stadium management activities," and "sports intermediary activities" is greater than zero, indicating that the growth rate of these five sectors is greater than the average growth rate of the national sports industry. In particular, the growth rate of "sporting goods, clothing, footwear, and hat manufacturing" and "sports construction activities" is higher than the average growth rate of the national sports industry; the competitive components of the other three sectors are less than 0, indicating that their growth rates are smaller than the average growth rate of the national sports industry [23]. The growth rate of the other three sectors is less than 0, indicating that their growth rate is smaller than the national average growth rate of the sports industry. When the competition component >0, it means that the growth rate is greater than the average growth rate of the national sports industry. When the competition component is less than 0, it means that its growth rate is less than the average growth rate of the national sports industry.

As shown in Figure 6, from the deviation component, the values of "sales of sports goods, apparel, footwear, and hats," "sports construction activities," and "manufacturing of sports goods, apparel, footwear, and hats" are larger, which means that these three sectors have more obvious growth rates. This indicates that these three sectors have more obvious sectoral advantages. In terms of share, "sporting goods, apparel, footwear, and hat manufacturing" has the largest contribution to the total sports industry.

Knowledge subgraphs were constructed for the experiments using DBpedia, and all models were executed in Python. For each item, entities were extracted from their reviews and extended to 3-hop relevant entities. 25% were used as a test set, 25% as a validation set, and the remaining 50% of the original data were used as a training set to obtain average results for five random groupings. The Adam optimizer was used for all models. The learning speed was set to 0.001, and the embedded item dimension and the users and entities were fixed at 128. To be fair, the

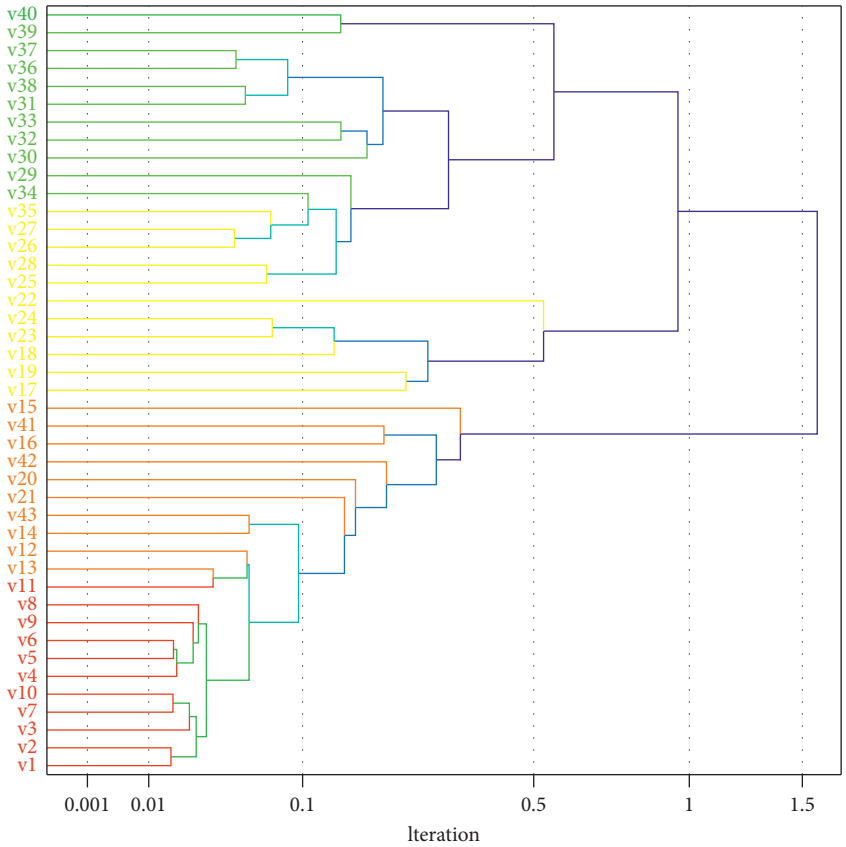


Figure 7: Topology identification results.

baseline was also set to the same size. BatchNorm was used to achieve faster and more stable training; research the topological structure characteristics and structural evolution law of the nervous system; and clarify the basic issues such as how neurons establish synaptic connections and how synaptic connections change.

As shown in Figure 7, there are errors in both parameters and topology at the beginning, and the errors decrease gradually as the iterations proceed. At the end of the iteration, there is a small error in the parameter identification, while the topology identification error is 0, indicating that the topology is correctly identified. The identification result is as follows: $a_1 = 1.32, a_2 = 1.59, a_3 = 2.14, a_4 = 1.64, a_5 = 2.83$. The initial solution is generated randomly. As the number of iterations increases, the parameter identification error and topology identification error decrease gradually. The final topology recognition error is 0, which means that all topologies are correctly recognized, while there is still a small error in parameter recognition. Since the solution space is larger for all neurons together, all topologies are identified, while there is still a small error in parameter identification.

5. Conclusion

This paper builds a comprehensive quantitative and biological neural network optimization model of the sports industry structure with the help of knowledge graphs and conducts in-depth research and analysis on the comprehensive quantitative and biological neural network optimization model of the sports industry structure. Based on the constructed knowledge graph and user-scholar interaction history, we investigate more effective and interpretable recommendation methods. First, for the knowledge graph, the TranSparse method is used to embed the corresponding vector representation, and then according to the characteristics of the knowledge graph, an attention mechanism is introduced to calculate the score between the target node and the neighbor nodes. Second, using the idea of high-order propagation, the entity representation is obtained in the knowledge graph in a hierarchical manner, and a dual aggregation method is proposed to aggregate entity information from two aspects, so that a richer entity representation can be

obtained. Finally, experiments are carried out on the data set, and the experimental results show that the method proposed in this paper has certain advantages in recommendation effect and interpretability. The connection strength identification of the Hindmarsh–Rose neural network model is carried out using a real-coded genetic algorithm. The network model identification problem is transformed into an optimization problem, and according to the characteristics of multiple optimization variables of the neural network, the implementation process of the genetic algorithm using real coding is described in this paper, and the algorithm is optimized and improved for the high-dimensional identification problem. On this basis, combined with the actual situation, it should be optimized and adjusted to effectively promote the systematic and scientific development of the sports industry structure. In turn, it can promote the transformation of the sports industry structure from a lower level to a higher level. Although some achievements have been obtained in this paper, there are still shortcomings. The current research uses typical values to fix these parameters in synaptic coupling, but more accurate model identification problems need to take all parameters into consideration.

Data Availability

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

Conflicts of Interest

The authors declare that they have no conflicts of interest or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work was funded by the 2020 Research Topics of Major Social Development Policies and Situations of the National Development and Reform Commission (grant/award number: 2020-10) and the 2019 State Exercise General Administration Decision Consulting Research Project (grant/award number: 2019-C-09).

References

[1] L. Bai, K. Zheng, Z. Wang, and J. Liu, "Service provider portfolio selection for project management using a BP neural network," *Annals of Operations Research*, vol. 308, no. 1-2, pp. 41–62, 2022.

[2] L. Jiang, T. Zhang, and Y. Feng, "Identifying the critical factors of sustainable manufacturing using the fuzzy DEMATEL method," *Applied Mathematics and Nonlinear Sciences*, vol. 5, no. 2, pp. 391–404, 2020.

[3] R. Zhang, "Analyzing body changes of high-level dance movements through biological image visualization technology by convolutional neural network," *The Journal of Supercomputing*, vol. 78, no. 8, pp. 10521–10541, 2022.

[4] Y. Wang, Y. Liu, and Y. Sun, "A hybrid intelligence technique based on the Taguchi method for multi-objective process parameter optimization of the 3D additive screen printing of athletic shoes," *Textile Research Journal*, vol. 90, no. 9-10, pp. 1067–1083, 2020.

[5] M. Kazemi Garajeh, T. Blaschke, V. Hossein Haghi, Q. Weng, K. Valizadeh Kamran, and Z. Li, "A comparison between sentinel-2 and landsat 8 OLI satellite images for soil salinity distribution mapping using a deep learning convolutional neural network," *Canadian Journal of Remote Sensing*, vol. 48, no. 3, pp. 452–468, 2022.

[6] L. Ma and B. Sun, "Machine learning and AI in marketing – c," *International Journal of Research in Marketing*, vol. 37, no. 3, pp. 481–504, 2020.

[7] P. S. Varsha, S. Akter, and A. Kumar, "The impact of artificial intelligence on branding: a bibliometric analysis (1982-2019)," *Journal of Global Information Management*, vol. 29, no. 4, pp. 221–246, 2021.

[8] D. Memmert and J. Perl, "Game creativity analysis using neural networks," *Journal of Sports Sciences*, vol. 27, no. 2, pp. 139–149, 2009.

[9] D. Yang, Y. Du, H. Yao, and L. Bao, "Image semantic segmentation with hierarchical feature fusion based on deep neural network," *Connection Science*, vol. 34, no. 1, pp. 1772–1784, 2022.

[10] M. Qingwen, W. Bojie, and S. Yehong, "Progresses and perspectives of the resource evaluation related to agri-cultural heritage tourism," *Journal of Resources and Ecology*, vol. 13, no. 4, pp. 708–719, 2022.

[11] H. M. Baskonus, H. Bulut, and T. A. Sulaiman, "New complex hyperbolic structures to the lonngren-wave equation by using sine-gordon expansion method," *Applied Mathematics and Nonlinear Sciences*, vol. 4, no. 1, pp. 129–138, 2019.

[12] M. K. Gupta and P. Chandra, "A comprehensive survey of data mining," *International Journal of Information Technology*, vol. 12, no. 4, pp. 1243–1257, 2020.

[13] T. P. Novak, D. L. Hoffman, and Y. F. Yung, "Measuring the customer experience in online environments: a structural modeling approach," *Marketing Science*, vol. 19, no. 1, pp. 22–42, 2000.

[14] Y. Ren, T. Cheng, and Y. Zhang, "Deep spatio-temporal residual neural networks for road-network-based data modeling," *International Journal of Geographical Information Science*, vol. 33, no. 9, pp. 1894–1912, 2019.

[15] S. D. Pathak, J. M. Day, A. Nair, W. J. Sawaya, and M. M. Kristal, "Complexity and adaptivity in supply networks: building supply network theory using a complex adaptive systems perspective∗," *Decision Sciences*, vol. 38, no. 4, pp. 547–580, 2007.

[16] L. Jinfeng and Y. Bo, "Design of evaluation system of physical education based on machine learning algorithm and SVM," *Journal of Intelligent and Fuzzy Systems*, vol. 40, no. 4, pp. 7423–7434, 2021.

[17] W. Chen, X. Li, X. Chen, and Y. Xiong, "Research on influence mechanism of running clothing fatigue based on BP neural network," *Journal of Intelligent and Fuzzy Systems*, vol. 40, no. 4, pp. 7577–7587, 2021.

[18] Y. Ren, S. Rubaiee, A. Ahmed, A. M. Othman, and S. K. Arora, "Multi-objective optimization design of steel structure building energy consumption simulation based on genetic algorithm," *Nonlinear Engineering*, vol. 11, no. 1, pp. 20–28, 2022.

[19] I. Almarashdeh, H. Bouzkraoui, and A. Azouaoui, "An overview of technology evolution: investigating the factors

influencing non-bitcoins users to adopt bitcoins as online payment transaction method," *Journal of Theoretical and Applied Information Technology*, vol. 96, no. 13, pp. 3984–3993, 2018.

[20] W. Zhang, Q. M. J. Wu, and Y. Yang, "A width-growth model with subnetwork nodes and refinement structure for representation learning and image classification," *IEEE Transactions on Industrial Informatics*, vol. 17, no. 3, pp. 1562–1572, 2020.

[21] A. Arabameri, A. Seyed Danesh, M. Santosh et al., "Flood susceptibility mapping using meta-heuristic algorithms," *Geomatics, Natural Hazards and Risk*, vol. 13, no. 1, pp. 949–974, 2022.

[22] A. Mohamed, M. K. Najafabadi, Y. B. Wah, E. A. K. Zaman, and R Maskat, "The state of the art and taxonomy of big data analytics: view from new big data framework," *Artificial Intelligence Review*, vol. 53, no. 2, pp. 989–1037, 2020.

[23] S. T. Fonseca, T. R. Souza, E. Verhagen et al., "Sports injury forecasting and complexity: a synergetic approach," *Sports Medicine*, vol. 50, no. 10, pp. 1757–1770, 2020.

Hindawi
Computational Intelligence and Neuroscience
Volume 2022, Article ID 4674620, 11 pages
https://doi.org/10.1155/2022/4674620

Research Article

A Neural Network Model for Color Element Data Analysis for Urban Spatial Environment

Xiaotang Xia[1] and Tingyang Li[2]

[1]*School of Urban Construction, Wuhan University of Science and Technology, Wuhan, Hubei 430070, China*
[2]*Wuhan Business University, Wuhan, Hubei 430000, China*

Correspondence should be addressed to Tingyang Li; lisir715@whut.edu.cn

Received 2 July 2022; Revised 24 July 2022; Accepted 1 August 2022; Published 21 August 2022

Academic Editor: Ning Cao

In this paper, a CNN model for color element data analysis of the urban spatial environment is constructed through an in-depth study of color element data analysis. This paper investigates a high-order structure formed by a few nodes; it proposes a motif-based graph autoencoder MODEL, combining redefined first- and second-order similarities and perfectly integrating motif structure and autoencoder. The algorithm first proposes an efficient graph transformation method to add the influence of central nodes. It then offers a primary awareness mechanism to aggregate the information of noncentral neighbors. Cen GCN_D and Cen GCN_E outperform the latest algorithms in node classification, link prediction, node clustering, and network visualization. As the number of network layers increases, the advantages of these two variants become progressively more prominent. This paper uses a support vector machine to implement classification validation based on CNN. The experimental results show that when 450 images are randomly selected as training data, the classification accuracy obtained by using the features of different CNN output layers is distributed between 91.4% and 95.2%. When the training set of the experiment reaches more than 300, the accuracy can exceed 90%, and the experimental results corresponding to different training sets a more stable trend. Finally, the trained classifier model is obtained in this thesis, which achieves the purpose of fast classification prediction based on CNN for color element data analysis of urban spatial environments.

1. Introduction

With the continuous development of cities and the gradual speeding up of the urbanization process, the urban color landscape in the city is getting more and more attention, and urban color design has become a hot topic in urban planning and construction. Urban color design is an incomparably complex system engineering, including urban design, urban planning, architecture, landscape, management, color science, and other related theoretical knowledge; designers need to consider the role and influence of various factors in the city, combined with the city color management department, together to put forward urban color design direction and goals [1]. The urban color design aims to shape the characteristics of an urban color image, maintain the ideal urban landscape, build a design system with different control categories and degrees, and put the design content into practice. Since the 1990s, with the prominence of urban color contradictions, research institutions have gradually started research and related practical work for urban color to solve the current urban color problems, which has played a specific positive effect on the harmony, control, and management of the urban color landscape [2]. However, most cities (including those that have completed urban color design studies) are overwhelmed by the modern buildings and materials rising from the ground and are unable to form a coordinated, harmonious, and unified urban color landscape [3]. The failure and low operability of urban color control is a common and core issue that must be addressed in the current urban color design approach.

From the perspective of human visual cognition, the urban color element environment can be regarded as an environment with spatial cognitive benefit, spatial construction auxiliary shape, and spatial repair reconciliation

and has a perceptual emotional information transfer function that simple geometric forms do not have, and therefore it becomes one of the indispensable elements in the process of the urban landscape and spatial landscape construction, with an irreplaceable and essential role, highlighting the urban population. It also plays an irreplaceable and critical role in the construction of urban landscape and spatial landscape, highlighting the universal value pursuit and historical heritage of urban people [4]. In the era of traditional urban architecture, the material properties of the building, the level of technology, and the social, political, and economic conditions in which it is located have all imposed constraints on the use of urban color and formed the color application habits of urban architecture with individual characteristics and rich regional style, thus making the color landscape in the city present more harmonious and unified style characteristics [5]. With the level of science and technology in human society, the technical barriers to constructing modern urban space have become smaller and smaller. The new construction materials and cross-regional exchange of convenient, easy-to-achieve separation of architectural color and building materials significantly reduce the regional restrictions on selecting materials in the construction of urban space to shape the diverse urban color landscape [6]. It provides more practical conditions for shaping various urban color and landscape features and opens the door to the problem of urban color abuse.

Deep learning has made significant breakthroughs recently, providing new ideas to solve old and contemporary urban and rural planning problems. The technology represented by CNN has also made breakthroughs in color analysis and has been widely used in natural language processing [7]. Compared with traditional machine learning algorithms, CNN eliminates many preprocessing processes and performs much better processing audio, image, video, and other data. At the same time, deep learning has been a hot research topic in academia. New network architectures are emerging every year, and the effectiveness of deployment and implementation for various tasks is gradually improving [8]. For users, well-packaged deep learning libraries have steadily lowered the threshold of using deep learning technology. Using deep learning libraries, users can easily choose the matching network models according to data requirements and deploy and manage their deep learning networks. In the era of big data, the large and diverse data contain various helpful information that further drives academic research and business development [9]. In network data, each sample is no longer an isolated entity; there are multiple relationships between them. Neural network algorithms do not take advantage of the higher-order structural information inherently present in the network. Thus, these algorithms do not fully mine the network structure, resulting in many missing relationships. In the context of big data, networks become more complex, diverse, and large-scale. The similarity between shallow pairs of nodes can no longer adequately characterize the structure of existing networks [10]. Therefore, there is an urgent need for warp networks to incorporate more complex higher-order systems in the network and explore more relationships between nodes. In this paper, we will study the neural network model for color element data analysis of the urban spatial environment to improve the algorithm's performance in relevant applications.

2. Related Works

With the rapid development of big data technology, combining the protection and research of traditional culture with big data technology is a big trend. There are many studies based on big data technology for traditional culture, and they have got good results [11]. Color is the most intuitive and prominent feature that can distinguish things' surface characteristics. There are many methods of color extraction, including the color histogram method. The color histogram method reflects the proportion of color information of different images in the image color space and is an effective way to research color information. At the same time, the color histogram is based on other coordinate systems and color spaces [11]. It can be used to study color and spatial information by calculating RGB, CIEL $*a*b$, HSV, etc. The color aggregation vector method proposed by Nasiri et al. [12] mainly solves the problem that the color moment method and histogram method cannot express the spatial location of image colors clearly. This method is an evolution of the color histogram method; the core of the aggregation vector method is dividing the pixels belonging to each histogram bin into aggregated and nonaggregated pixels [12]. Eslami et al. [13] proposed that the purpose of color metrics is to represent the colors seen by everyday people objectively and numerically related to the requirements of a specific application. The concepts of color gamut and color space emerged. The commonly used color metric models in urban color research are closely related to the Munsell color representation system and the CIE color representation system [13]. In color planning, the HIS and HSV models under the Mimsell specimen system and the parallelized representation of RGB models based on primary color mixing are more commonly used [14]. Through these methods, we can digitally integrate and summarize colors and store the results in the form of text, maps, and diagrams in the database as necessary data support for planning.

From the 1990s to the present, research and practical work on urban color have been gradually carried out, and the color landscape of cities has been improved to a certain extent. In recent years, academic research on urban color design theory and practice has gradually expanded from the initial few scholars to an enormous scope. For a city, the more urgent, the more critical, and the more stringent control needs to be solved, and with the fastest results, the color issues we consider the most basic need to be solved [15]. This is because, from the planning manager's point of view, they often expect maximum results with minimal investment in the shortest possible time. At the same time, refinement can be carried out gradually at a later date. Urban color design's most urgent and primary task is to protect, restore, and remedy color pollution in the city's possible historic districts and the coordinated blocks around cultural preservation units [16]. Relatively higher and better pursuits,

such as creating a unique "color" of the city and creating a harmonious color atmosphere, are the icing on the cake. After that, it belongs to the overall city color design type, hoping to use the city color design to coordinate and control the city's color landscape. Accordingly, the current types of urban color design can be divided into three major categories: historical and cultural heritage and regional preservation, new urban areas and particular areas of urban color practice, and overall urban color design. Yang H et al. used computer vision technology to construct suitable convolutional neural networks (CNN) to identify the street green view rate, street functional attributes, and pedestrian attributes in public space and made a preliminary exploration of quantitative research on general urban areas based on big visual data [17]. Gu et al. [18] also proposed a model for urban imagery research based on web-based image data; constructed a research framework with urban imagery element composition, dominant direction, characteristic degree, and similarity of urban imagery as modules; and conducted an empirical study for 24 major cities [18].

Due to the significant differences in urban planning and construction patterns, foreign urban color design experiences and patterns cannot be directly copied and applied to Chinese urban color practice [16]. Regarding the practice and research of urban color design, based on foreign color theories and methods, we have been exploring ideas and techniques adapted to the current situation of cities from the beginning stage. We have developed a set of current urban color design research systems. The study of urban spatial environment identification mainly studies urban characteristics, urban planning, and urban design for shaping. So far, the more popular urban design theories and methods in the West focus on the functional design of public space, private space, and shared space from the urban environment landscape and logo design elements [17]. Since the foreign urban construction and urban planning system can reflect the will and vision of designers and planners more, in shaping the overall urban image, it can reflect the personality of the creator of the urban image as well the nature of the urban image more. Therefore, the urban spatial environment identification can often be reflected through urban planning and construction.

3. Neural Network Model Based on Data Analysis Construction

CNN is a mixture of convolutional operations and deep neural networks. Convolutional neural networks are often used for data feature selection and image processing applications because of their symbolic learning capability [19]. Compared to other multilayer network structures, convolutional neural networks have more powerful abilities; neurons are the basic units that make up a neural network. A neural network contains multiple structural layers, and each layer contains various neurons. In the map, the vectors x_1, x_2, and x_3 are the input signals to the neurons, and +1 represents their bias, usually represented by the symbol b. The neuron is described as follows:

$$f_{(w^t-x)} = \sum_{i=1} \frac{\sqrt{w_i x_i - b}}{w_i - x_i}, \tag{1}$$

where x represents the input to the neuron, W and b are the neuron's weight term and bias term, respectively, and the function f is a nonlinear activation function that represents the response to the input signal. The input signal is the output after a nonlinear operation through a specific activation function of a single neuron, which is calculated as follows:

$$h^x_{w-b} = \sum \frac{w^t - 1}{f\left(w^t - x\right)}. \tag{2}$$

Convolutional neural networks have a unique structure with a hierarchy consisting of an input layer, an output layer, and single or multiple hidden layers. $L1$ is the input layer, while x_1, x_2, and x_3 are the input signals and +1 bias terms. The middle of the network below is the hidden layer, while $L2$ is the final output layer used to output the final result. In this case, the network layers are interrelated, and the previous layer's output is used as input to the next layer and propagated backward. The vector dimensions of the information and hidden layers are usually set according to the problem under study, while the parameters of the hidden layers are determined experimentally. The neurons in each layer of the convolutional network receive the information output from a set of units in the previous layer and perform a series of convolutional operations. Usually, the computation of the convolutional layers can be represented by the following equation:

$$x_i = \sum x_i^{l+1} + w_{ij} - b\sqrt{(l+j)_{ie}^{mj}}. \tag{3}$$

In equation (3), l represents the number of convolutional layers. In contrast, w means the convolutional kernel, b is the bias term of the network, and Mj is denoted as the input feature image. Since the convolutional layers have the feature of weight sharing, the parameters will be significantly reduced. Now, suppose there is a model with 1000 hidden layer nodes and an input image with a pixel size of 1000×1000, while if a fully connected approach is used, the number of parameters will be $1000 \times 1000 \times 1000 = 109$; while if 100 convolutional kernels are used for processing, and the size of convolutional kernels is 10×10, there will be $10 \times 10 \times 100 = 104$ parameters, and it can be seen that the convolutional neural network makes a significant reduction in the number of parameters.

A general convolutional neural network can be broadly divided into three layers: an input layer, a hidden layer, and an output layer. Since neural networks are generally used to perform computer vision tasks, the input layer often accepts three-dimensional or one-dimensional arrays, that is, three-channel RGB images or single-channel grayscale images. In addition, the input layer is used to normalize the input data from the [0, 255] pixel value interval to the [0, 1] interval to improve the learning efficiency of the network since the network training generally requires learning and parameter updating using the gradient descent method. The output

layer is used to output the network results, and the output of the trained network is the final fitted result; for the network to be prepared, the difference between the predicted output and the natural result is calculated by the loss function during training, so that the weights in the previous layers can be updated by back propagation to make the network fit the result more in line with the demand [20]. This process is called the network learning process. The hidden layer is the most critical layer of a convolutional neural network and can be divided into three unique layers: a convolutional layer, a pooling layer, and a fully connected layer. A hidden layer can contain one or more convolutional and pooling layers and generally includes a fully connected layer. The input data can be abstracted by repeatedly performing convolution and pooling operations. This layer-by-layer abstraction represents the features of the input data at different degrees and scales. Therefore, this class of layers is the most critical layer for feature extraction and is the fundamental reason for the powerful performance of convolutional neural networks. Since the fully connected layer can be replaced by a convolutional layer of the same size as the input image, only the convolutional and pooling layers will be described in detail.

3.1. Convolutional Layer. A deep convolutional neural network's hidden layer generally contains many convolutional layers, which convolve the input image and then output the result and use it as the input of the next layer for forwarding propagation. The convolution operation can extract the deep features of the abstracted image data, and these removed features can make the final prediction result of the network more accurate. Hence, the convolution layer is the essential core layer of the convolutional neural network. Each convolutional kernel corresponds to a bias term in convolutional neural networks. After convolving each input, this bias term needs to be added and activated by the activation function, and the final result is the output of this convolutional layer. Adding a bias term makes the network fitting result more accurate, and the network convergence speeds faster. Therefore, the convolutional computation process after adding the bias term is as follows:

$$c_i^j = \sum_{m=1} w_{m,n} \times a_{m-i} + a_{n-j} - b, \tag{4}$$

where c is the convolution result, k is the convolution kernel size, w is the convolution kernel weight, a is the input image to be convolved, and b is the bias term. In the actual calculation, two other critical parameters for the convolution operation, namely, stride and padding, impact the output data size after convolution. The stride represents the distance that the convolution kernel slides on the original image data during each convolution, and the setting of this parameter can control the abstraction level of the output image after convolution; the more significant the stride is, the larger the abstraction level is, and the smaller the image size is. Generally, the step size is set to 2, which means that the convolution kernel extracts feature from every region of the image; the fill setting is related to the size of the convolution kernel, which is generally divided by two and rounded upward as the size of the image boundary fill area. The fill mode is usually chosen as zero fill. Therefore, for a step size of s, a filling of p, an input image size of $w * w * 4$, a convolutional kernel size of k, and several m, the output image size is as follows:

$$\mathrm{out}_w^k = \sum\left[\frac{w+k+m}{\sqrt{s-1}} - 2p\right] \times \left[\frac{w-k-m}{\sqrt{s+1}} - 2p\right]. \tag{5}$$

3.2. Pooling Layer. For most convolutional neural networks, there is often a pooling layer after each convolutional layer, and the role of pooling layer is mainly to abstract the input data further and also to perform feature dimensionality reduction to reduce the model parameters and improve the operation speed and prevent overfitting to a certain extent. Pooling can be divided into average pooling and maximum pooling, and their calculation formulas are expressed as follows:

$$\begin{aligned} b_{(i-j)} &= \sum_{n=1} \frac{\sqrt{a_{i-m} + a_{j-n}}}{n_h - n_w} \times (h - w), \\ b_{(i+j)} &= \sum \frac{a_{i-m} - a_{j+n}}{\max\left(a_{i-m} + a_{j+n}\right)}. \end{aligned} \tag{6}$$

Pooling is also not essential for convolutional neural networks. For example, a $2 * 2$ pooling operation can be replaced by a convolution with a step size of 2; thus, reducing the number of operations in a pooling layer. However, this approach is not stable and may give better results on sizeable deep network structures but may give worse results on some small networks. A convolutional neural network is a multilayer structure containing a feature extractor consisting of several convolutional layers, subsampling layers (pooling layers), and a connection layer in a convolutional layer; a neuron is connected to only some of its neighboring neurons. A convolutional layer in a CNN usually contains several feature maps, each consisting of several rectangularly arranged neurons, and the neurons in the same feature map share weights. The shared consequences are the convolutional kernel.

4. Color Element Data Analysis Model Construction for Urban Spatial Environment

Color is a necessary attribute of the urban spatial environment, and no urban landscape can exist independently without its color attributes. Urban color, in the general sense, can be regarded as the collection of color attributes of all visual objects in the urban spatial environment. It includes the color attributes of buildings and structures in the city, the colors of natural mountains and geography as the background of the urban spatial domain, and even the people and vehicles moving in the urban spatial entity [21]. From the perspective of color creation and planning design of the urban spatial environment, we often divide the color information covered in the urban spatial environment into fixed color information attributes, semifixed color

information attributes, and mobile color information attributes. The selected color information attributes refer to the relatively slow change of color attributes and their carriers, such as the urban spatial environment accounted for a relatively large number of buildings and structures belonging to the ranks of fixed color information attributes; and semifixed urban color information attributes refer to the urban spatial environment that requires a certain amount of time before the more significant changes in color attributes and their carriers, such as the common trees in the city landscape gardening and street landscape vignettes, flowing. The flowing color information attributes usually refer to the users of urban space entities and generally refer to the harmonious color attributes and their carriers, such as the flow of people and traffic in the city. The urban color planning covered in this thesis only refers to fixed and semifixed color information attributes in the urban spatial environment. Color planning mainly reflects the color characteristics and differences embodied in the different functional divisions of regional cities, which can be attributed to the overall positioning of specific city colors: responding to the impact of changes in the regional natural environment, cultural environment, and artificial built environment, as well as the color evolution and adaptation caused by the expansion of urban space, urban industrial structure adjustment, and the strengthening of spatial and temporal links between cities in the regional development environment. The matching is shown in Figure 1.

According to the composition mechanism of color landscape elements in the urban space environment, the color attributes can be divided into natural environment color and artificial environment color. Natural environment color refers to the natural color elements and their carriers, such as the sky, water system, and mountains, which are within reach of our visual perception. The artificial environment color refers to the color elements and their carriers reflected by the artificially constructed structures, including the systems and rigid pavements used daily by urban residents; all belong to the artificial environment color category. In the specific color planning and design implementation, we can divide the urban space environment color into background color, primary color, auxiliary color, and embellishment color. Background color refers to the natural geographic environment on which the urban spatial environment is based.

In contrast, the primary color refers to the main color information presented on the facade of the buildings in the urban spatial entity. The additional and embellishment colors are the color information elements that account for less on the buildings and play a supporting role or decoration. Through the color classification of the urban space environment, we can see that, from different perspectives, different urban space color environments can be divided; these color elements, based on differences but also have a specific internal connection, constitute the urban space we know color landscape style.

An urban spatial environmental system is a complex system composed of interacting and interdependent spatial elements with specific levels, structures, and functions in a particular social environment. It has the following characteristics: first, the urban spatial environment system is an essential subsystem of the extensive urban system, enabling the city to exist in material form as a comprehensive system. Secondly, the urban spatial environment system is the projection of the interaction of various subsystems of the urban social system on the urban land. The composition of modern urban space has become increasingly diversified, incorporating many traditional outdoor public activities into the indoor spatial environment, resulting in the emergence of microurban rooms in integrated buildings. This is a powerful complement to the complete connotation of the urban public space system, which meets the spatial needs of people for modern social life. Thus, it is measured regarding urban space's social and physical attributes. The analysis of urban space data is shown in Figure 2.

The urban spatial environment is, first of all, a material space whose core function is to carry out all kinds of social activities in the city and is the primary field for people to cognize and experience the city. The identification system of the urban spatial environment is complex in a particular social setting, composed of spatial identification elements that interact and depend on each other. It has a certain level, structure, and function. The urban spatial environment has many material and nonmaterial aspects [22]. It is necessary to organize and manage them effectively and orderly to make them an integral part of the urban spatial environment identification system. Building a structural model is feasible to contain these complicated elements in an orderly way. The structural model can analyze and categorize the ingredients in many aspects and highlight the characteristics of the urban spatial environment. Therefore, this chapter focuses on the urban spatial environment identification system, composed of the structural model of urban spatial environment identification and the identification of typical forms of urban spatial environment. The establishment and use of the urban spatial environment identification system are essentially the processes of discovering the characteristics of the urban spatial environment and its effective identification. The analysis of urban color design type data is shown in Figure 3.

The spatial environment created according to different design concepts will be different, and there are pronounced differences in characteristics between them, with prominent branding of the concept. Suppose the information on the features of the urban spatial environment that can reflect the history and culture of the city is refined and processed in a certain way. In that case, it can be associated with the corresponding elements so that the purpose of identification can be achieved. For example, suppose the information reflecting the characteristics of the urban spatial environment is refined and processed in a certain way so that it can be associated with the corresponding design concept. In that case, the purpose of identification from the perspective of the design concept can be realized. It is assumed that linear and point-like buildings, streets, and spatial forms are a set of information vectors reflecting a particular design concept in an urban spatial environment, and each information vector has multiple spatial environment characteristics. By

Figure 1: Urban color planning corresponds to the system process.

assigning each information vector to different categories according to the design concept through fuzzy clustering, a mapping can be established between the information of the urban spatial environment and the characteristics of the design concept signs. Through the analysis and understanding of people's urban imagery, we can grasp the environmental elements and their features that affect people's perception and classify these characteristics to make the urban spatial environment highlight its unique signs for identification. This kind of clustering is based on analyzing and generalizing different people's urban imagery. The title's premise is to evaluate and analyze a certain number of model urban imagery and its corresponding urban elements. The clustering method in the feature clustering identification model not only adopts the fuzzy logic method commonly used at present but also other methods can be used according to the actual situation or application effect. In addition, when the boundary of spatial environment classification is not very obvious, some corresponding preprocessing can be carried out before sorting. For example, some fusion of spatial environment feature information can be used to solve the problem of spatial environment information recognition under uncertain conditions with few paradigms and multiple feature spaces, especially for urban spatial environment imagery feature recognition, which has a good effect.

5. Analysis of Results

5.1. Neural Network Model Analysis of Color Element Data Analysis of the Urban Spatial Environment. Color harmony is essential in the overall control of the city such as the ratio between adjacent colors between the hue brightness vividness relationship and similar hues between environmental

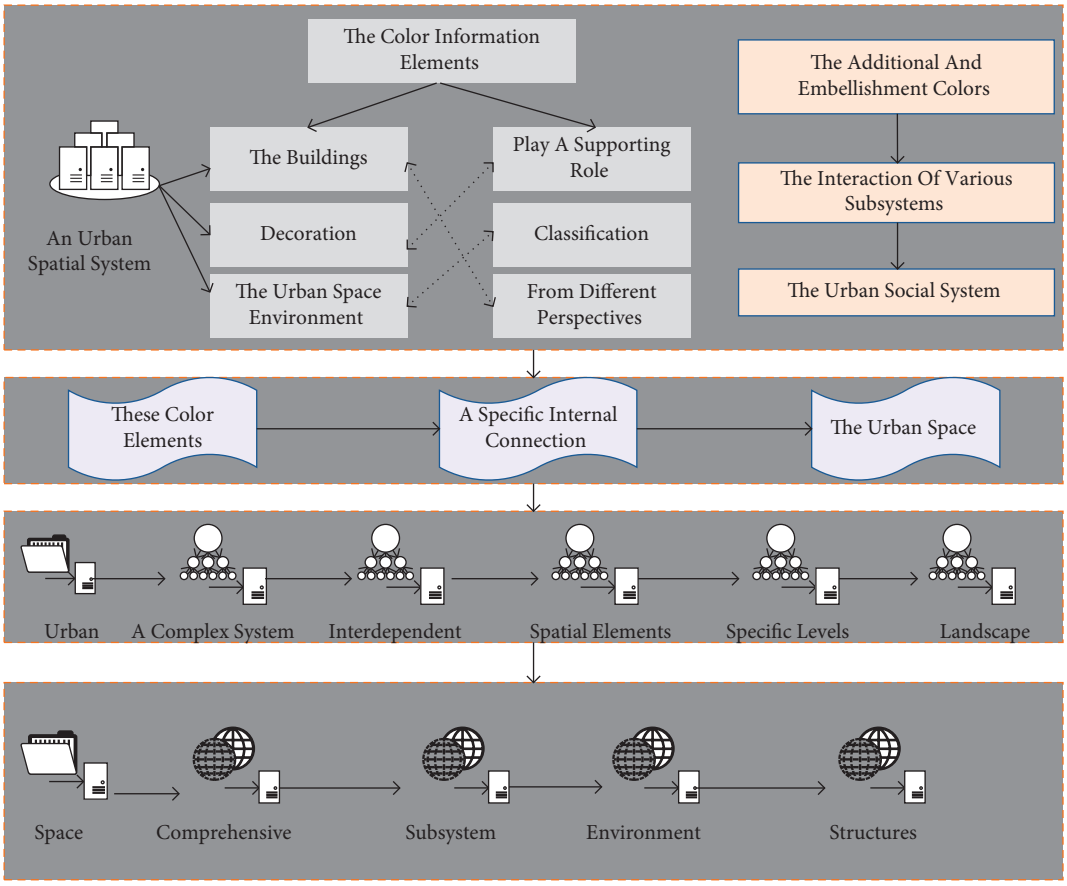


Figure 2: Urban spatial data analysis process.

colors, hue harmony, and hue harmony method to adjust the discordant urban color [23]. (1) The contrast relationship between adjacent colors: through the edge contrast of adjacent colors, people can identify subtle color differences, the hue also in this contrast, in the interaction of various visual sensory changes. (2) Environmental color harmony relationship: the color harmony with the form, building materials, and places will change accordingly.

This paper divides the primary data set into 90% of the data as the training set and the remaining 10% as the test set during the actual training. For the training parameters, the batch size is set to 32, the initial learning rate is set to 0.002, and the learning rate is multiplied by 0.2 every 30 training rounds for 500 games. The training data set is randomly disordered in each round. The weights of the Adam optimization algorithm were developed to decay from 4 to 10, and the two hyperparameters β_1 and β_2 were set to default values of 0.9 and 0.999, respectively. The variation of model accuracy over time is shown in Figure 4.

For convolutional neural networks, a well-designed network structure is a basis for the convergence of the model. At the same time, a suitable data set is needed to ensure whether the network can accomplish the set learning goals and whether the trained network model can meet the final task requirements. A good data set can speed up the convergence of the network so that the network can learn the mapping between the input and the target output faster and

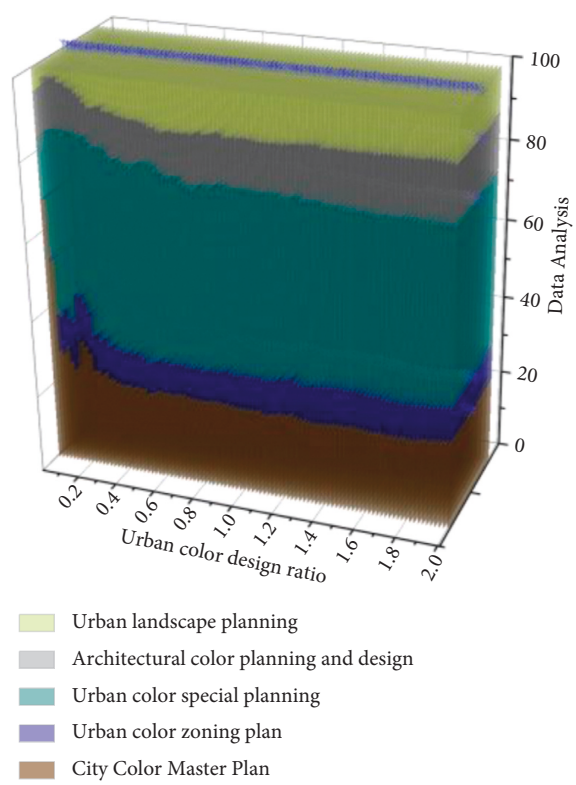


Figure 3: Data analysis of urban color design types.

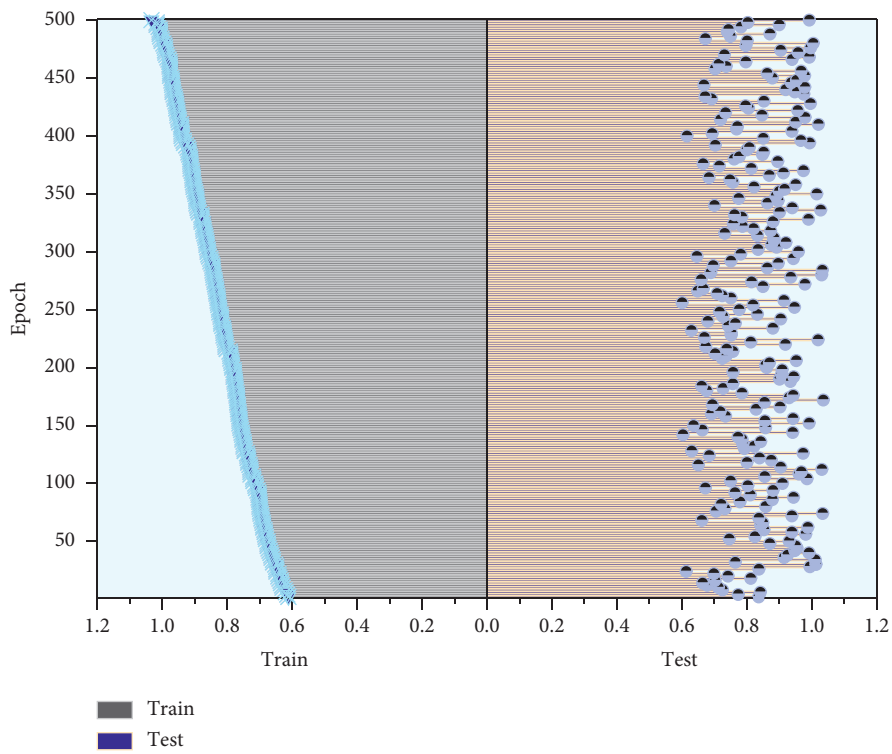


Figure 4: Variation of model accuracy with a period.

complete the task. On the other hand, it can significantly reduce the network's performance, making it unable to learn the correct parameters and thus underfitting or overfitting. This affects the generalization performance of the network and may even cause the network to fail to converge. Once the network meets the accuracy requirements, the web can be used to compute the example data. In the Keras framework, the h5py library was used to load the previously trained U-Net network for prediction, and all the image data of the streetscape were input in batches. After 9 hours and 469955 calculations, the pixel percentages of greenery, sky, buildings, pedestrian streets, and motorized streets in all the streetscape data were finally obtained. At this point, the task of the image segmentation network has been fully completed. The variation of model similarity with a period is shown in Figure 5.

A convolutional neural network was built to analyze and calculate the color elements of the urban spatial environment previously acquired, and the results of the classification of all color elements were obtained in a short time. The softmax function is used in the category, and its calculation formula is as follows:

$$\text{Softmax}\,(x-1) = \int_{x=1} \frac{e^{x_{i-1}} - 1}{\sqrt{e^{x_1} + e^{x_2}}}. \tag{7}$$

Equation (7) indicates here that it is one of two categories, in this case, 1 or 2, so the network outputs a matrix with two values, one for the prepredicted color value and the other for the postpredicted color value. It can be easily deduced that the sum of the two output values is 1. Since the two predicted values are reciprocal, when evaluating the color element scores, it is only necessary to focus on using the cheerful color expected value. Based on the above-mentioned principles, it is possible to quantify the color element data for all urban spatial environments.

5.2. Color Element Data Analysis Realization of the Urban Spatial Environment. The natural colors in the urban space environment are influenced by the city's geographic location, topography and terrain, natural vegetation, and other factors, presenting different visual perceptions. Natural colors are also divided into constant and nonconstant colors. Land, mountains, rivers, and other natural colors that can remain stable over a long period are standard colors with a particular law of change; nonconstant colors are represented by the sky, vegetation, and urban neon, which have the characteristics of changing with time. The natural colors are in people's daily lives and affect society's physical and mental feelings. Artificial colors in urban color, as the name suggests, are artificially planned and designed in urban advertisements, streets, public vehicles, equipment and facilities, and pedestrian clothing colors. City streets, high-rise buildings, unique buildings, and other relatively fixed colors; pedestrian clothing, cars, and other colors constitute the flow of color; city advertising, lighting, street decorations, signs, and other colors form a temporary color. Artificial colors accompany people's daily work and life, profoundly influence the surrounding environment, and harmonize with natural dyes to jointly determine the color appearance of the city. The results obtained by SVM classification based on the

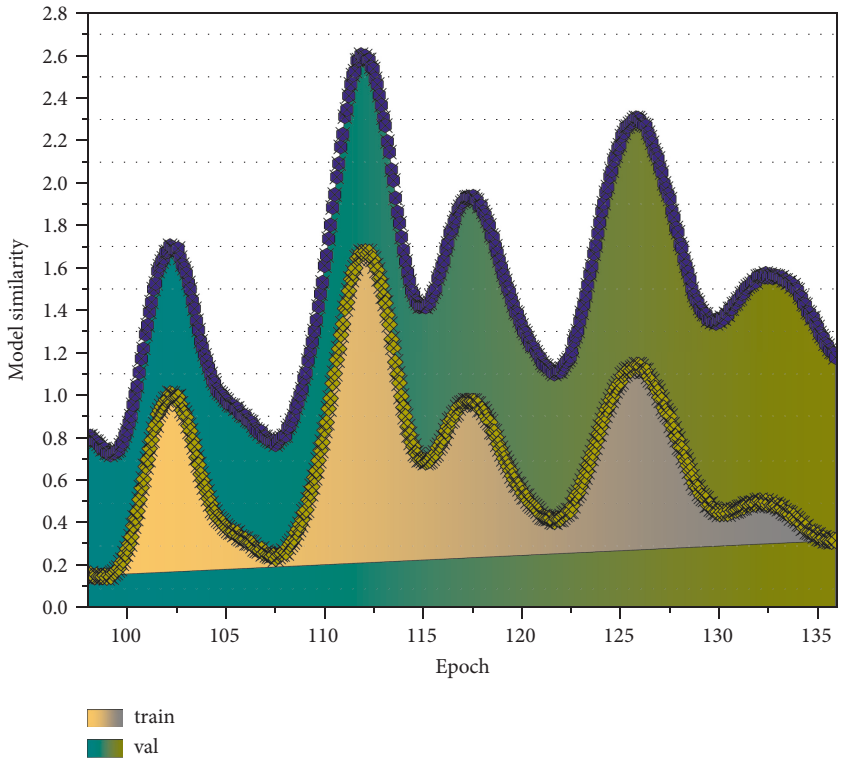


Figure 5: Variation of model similarity with a period.

existing CNN model are better; there are about 477 training data. To verify the effect of data size on the experimental results, fc8 is selected as the output layer, and the size of the experimental data will be adjusted to get the corresponding classification results. The number of all data sets is 716, of which 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, and 0.9 are selected as training data, and the others are used as test data for experiments. The visualized image of the training set is shown in Figure 6.

In each selection of training data, the entire data set will be randomly disrupted to ensure the fairness of the data. Still, because there is a specific error in each data selection, this thesis only uses the experimental data as a reference to analyze the overall change trend. For example, the cross-validation accuracy obtained when the number of training data was 358 is 94.8%. The result is relatively higher than the experimental results when the number of training samples was 430, 502, and 574, but this does not indicate that the experimental results corresponding to the number of pieces of training were 368 are better. The results must be analyzed based on the overall trend of change. As can be seen, the minimum number of training data is 72, the maximum number is 716, and the accuracy rate is guaranteed to be above 80%. Also when the amount of data reaches about 300, the overall trend is more stable. When the data come to 300 or more, the mean square error of the trained SVM model is not high, and it can be guaranteed to be around 0.1, which indicates that the experimental results are more reliable. The experimental results of different training sets are shown in

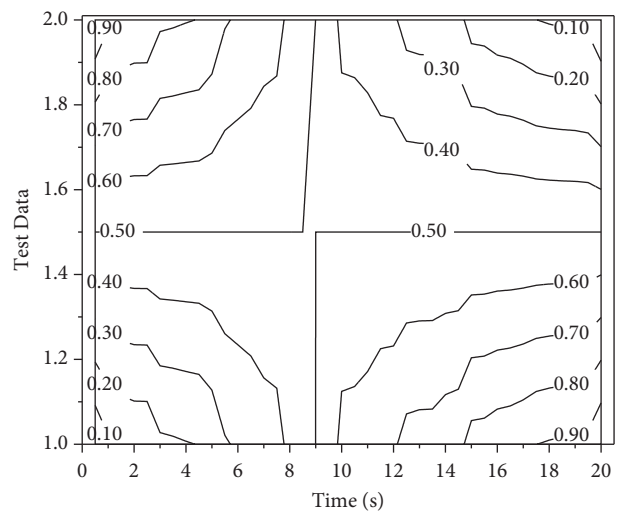


Figure 6: Visualization of the training set distribution.

Figure 7. The experimental data based on mass spectral images contain 149 images in the most categories and 53 shots in the least. The amount of data in all types of the network databases far exceeds the experimental data of mass spectral images, corresponding to the better classification results obtained from the experiments. Therefore, it can be concluded that the data analysis neural network model applied in this paper has a good recognition ability and is suitable for designing color elements in urban spatial environments.

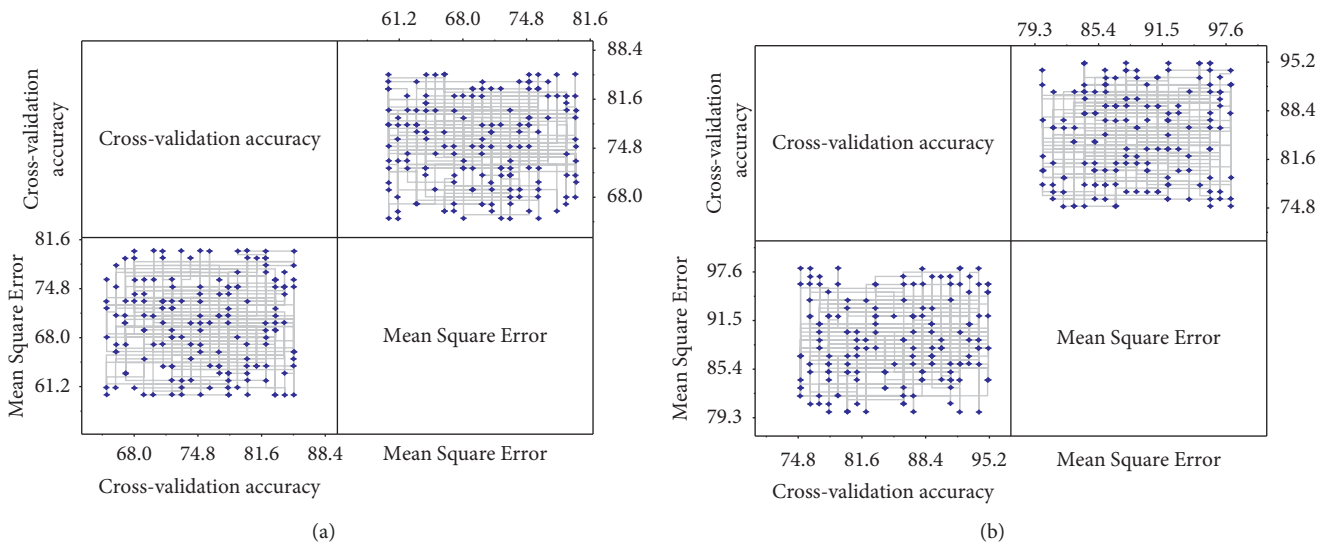


Figure 7: Experimental results of different training sets.

6. Conclusion

The generation and development of the urban spatial environment is a selfgrowing process with a long-time dimension. As an essential element affecting the urban landscape, the buildings in the urban spatial environment have a relatively long-term existence; that is to say, their influence on the urban landscape is relatively long from the date of their construction to the date of their demolition. Therefore, renovation and renewal, which only change the color of the building's appearance instead of demolishing the building itself, become an essential means of shaping and improving the urban landscape economically and operably. This paper proposes a convolutional neural network algorithm, Cen GCN, on a scale-free network. This paper presents a graph transformation method to capture the influence of the central node on its surrounding neighbors. The power of the central node on similar neighbors increases, and the impact on dissimilar neighbors decreases. This paper also proposes a significant awareness to aggregate the information of noncentral node neighbors. In experiments, this paper presents two variants, Cen GCN_D and Cen GCN_E. Both variants outperform the latest convolutional neural network algorithms in node classification, link prediction, node clustering, and network visualization. Even though the scale-free nature is based on degree features, significant results can still be achieved using eigenvalue centrality. This study shows that convolutional neural networks can be improved by considering central nodes and centrality. It is also demonstrated that the introduction of node centrality can deepen the convolutional neural network. When the number of neural network layers reaches ten, the algorithm still achieves long-lasting results. The neural network based on the AlexNet model has high feature learning performance, and the output fully connected layers can all represent the information contained in the mass spectral images relatively well; thus, obtaining high accuracy and low mean square error. In particular, the last fully connected layer, fc8, receives higher classification accuracy of 95.2% than other fully connected layers. The inherent neural network model based on millions of training images does not have high requirements in terms of training set size and has a relatively stable recognition capability when the data volume reaches 300 or more. In this paper, the trained classifier model is obtained to classify and predict the color element data analysis of the urban spatial environment based on CNN.

Data Availability

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

Conflicts of Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work was supported by the School of Urban Construction, Wuhan University of Science and Technology.

References

[1] S. Law, C. I. Seresinhe, Y. Shen, and M. Gutierrez-Roig, "Street-Frontage-Net: urban image classification using deep convolutional neural networks," *International Journal of Geographical Information Science*, vol. 34, no. 4, pp. 681–707, 2020.

[2] H. Moayedi, A. Moatamediyan, H. Nguyen, X. N. Bui, D. T. Bui, and A. SA. Rashid, "Prediction of ultimate bearing capacity through various novel evolutionary and neural network models," *Engineering with Computers*, vol. 36, no. 2, pp. 671–687, 2020.

[3] A. Jahani, S. Allahverdi, M. Saffariha, A. Alitavoli, and S. Ghiyasi, "Environmental modeling of landscape aesthetic value in natural urban parks using artificial neural network technique," *Modeling Earth Systems and Environment*, vol. 8, no. 1, pp. 163–172, 2022.

[4] Z. Cui, K. Henrickson, R. Ke, and Y. Wang, "Traffic graph convolutional recurrent neural network: a deep learning framework for network-scale traffic learning and forecasting," *IEEE Transactions on Intelligent Transportation Systems*, vol. 21, no. 11, pp. 4883–4894, 2020.

[5] X. Yang, X. Li, Y. Ye, R. Y. K. Lau, X. Zhang, and X. Huang, "Road detection and centerline extraction via deep recurrent convolutional neural network U-Net," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 57, no. 9, pp. 7209–7220, 2019.

[6] Z. Allam, "Achieving neuroplasticity in artificial neural networks through smart cities," *Smart Cities*, vol. 2, no. 2, pp. 118–134, 2019.

[7] R. D. Majd, M. Momeni, and P. Moallem, "Transferable object-based framework based on deep convolutional neural networks for building extraction," *Ieee Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 12, no. 8, pp. 2627–2635, 2019.

[8] R. Hang, Q. Liu, D. Hong, and P. Ghamisi, "Cascaded recurrent neural networks for hyperspectral image classification," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 57, no. 8, pp. 5384–5394, 2019.

[9] L. Zhou, S. Zhang, J. Yu, and X. Chen, "Spatial–temporal deep tensor neural networks for large-scale urban network speed prediction," *IEEE Transactions on Intelligent Transportation Systems*, vol. 21, no. 9, pp. 3718–3729, 2020.

[10] L. Zhao, Y. Song, C. Zhang et al., "T-GCN: a temporal graph convolutional network for traffic prediction," *IEEE Transactions on Intelligent Transportation Systems*, vol. 21, no. 9, pp. 3848–3858, 2020.

[11] W. Zeng, C. Lin, J. Lin et al., "Revisiting the modifiable areal unit problem in deep traffic prediction with visual analytics," *IEEE Transactions on Visualization and Computer Graphics*, vol. 27, no. 2, pp. 839–848, 2021.

[12] V. Nasiri, A. A. Darvishsefat, R. Rafiee, A. Shirvany, and M. A. Hemat, "Land use change modeling through an integrated Multi-Layer Perceptron Neural Network and Markov chain analysis (case study: arasbaran region, Iran)," *Journal of Forestry Research*, vol. 30, no. 3, pp. 943–957, 2019.

[13] E. Eslami, A. K. Salman, Y. Choi, A. Sayeed, and Y. Lops, "A data ensemble approach for real-time air quality forecasting using extremely randomized trees and deep neural networks," *Neural Computing & Applications*, vol. 32, no. 11, pp. 7563–7579, 2020.

[14] E. Saralioglu and O. Gungor, "Semantic segmentation of land cover from high resolution multispectral satellite images by spectral-spatial convolutional neural network," *Geocarto International*, vol. 37, no. 2, pp. 657–677, 2022.

[15] Y. Chen, K. Zhu, L. Zhu, X. He, P. Ghamisi, and J. A. Benediktsson, "Automatic design of convolutional neural network for hyperspectral image classification," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 57, no. 9, pp. 7048–7066, 2019.

[16] X. Hou, K. Wang, C. Zhong, and Z. Wei, "ST-trader: a spatial-temporal deep neural network for modeling stock market movement," *IEEE/CAA Journal of Automatica Sinica*, vol. 8, no. 5, pp. 1015–1024, 2021.

[17] H. Yang, B. Yu, J. Luo, and F. Chen, "Semantic segmentation of high spatial resolution images with deep neural networks," *GIScience and Remote Sensing*, vol. 56, no. 5, pp. 749–768, 2019.

[18] K. Gu, Y. Zhou, H. Sun, L. Zhao, and S. Liu, "Prediction of air quality in Shenzhen based on neural network algorithm," *Neural Computing & Applications*, vol. 32, no. 7, pp. 1879–1892, 2020.

[19] J. Xie, N. He, L. Fang, and A. Plaza, "Scale-free convolutional neural network for remote sensing scene classification," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 57, no. 9, pp. 6916–6928, 2019.

[20] P. Gogikar, B. Tyagi, and A. K. Gorai, "Seasonal prediction of particulate matter over the steel city of India using neural network models," *Modeling Earth Systems and Environment*, vol. 5, no. 1, pp. 227–243, 2019.

[21] R. Sun, Z. Tu, L. Fan et al., "The correlation analyses of bacterial community composition and spatial factors between freshwater and sediment in Poyang Lake wetland by using artificial neural network (ANN) modeling," *Brazilian Journal of Microbiology*, vol. 51, no. 3, pp. 1191–1207, 2020.

[22] S. Zorzi, E. Maset, A. Fusiello, and F. Crosilla, "Full-waveform airborne LiDAR data classification using convolutional neural networks," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 57, no. 10, pp. 8255–8261, 2019.

[23] V. C. Ebhota, J. Isabona, and V. M. Srivastava, "Environment-adaptation based hybrid neural network predictor for signal propagation loss prediction in cluttered and open urban microcells," *Wireless Personal Communications*, vol. 104, no. 3, pp. 935–948, 2019.

Hindawi
Computational Intelligence and Neuroscience
Volume 2022, Article ID 3278194, 8 pages
https://doi.org/10.1155/2022/3278194

Research Article

The New Economic Era Analysis of the Structure System of Chinese Household Consumption Expenditure Based on the ELES Model

Chaozhi Fan, Siong Hook Law, Saifuzzaman Ibrahim, and N. A. M. Naseem

School of Business and Economics, Universiti Putra Malaysia, Serdang, Selangor 43440, Malaysia

Correspondence should be addressed to Chaozhi Fan; gs57155@student.upm.edu.my

Received 18 June 2022; Revised 19 July 2022; Accepted 21 July 2022; Published 9 August 2022

Academic Editor: Luobing Dong

In recent years, the new economy has entered a phase of rapid development and upgrading China's service consumption is driving the continuous optimization of the population's consumption structure. To realize the rationalization of the Chinese household consumption structure, the ELES model is used to analyze the structure system of Chinese household consumption expenditure. This article constructs the ELES model, divides the types of Chinese household consumption expenditure structure systems, establishes consumption expenditure function, analyzes the influencing factors of the consumption expenditure structure system, and obtains the analysis results from static and dynamic aspects. Based on the statistics of Chinese household consumption expenditure data in recent years, this article obtains the analysis results of the consumption expenditure structure system: the basic consumption demand and marginal consumption tendency of food are in the first place, and the consumption expenditure structure system has gradually changed into the development-type and enjoyment-type consumption mode. Through increasing the income of rural residents, guiding reasonable consumption concept, optimizing consumption environment, and so on, we can promote the proposal and implementation of the optimization of China's household consumption expenditure structure system to improve the rationalization of China's household consumption structure system.

1. Introduction

Consumption is not only the end of social reproduction, but also the starting point of its logic. It is also one of the three main driving forces of economic growth and plays an important role in economic growth. Consumption structure is one of the key research contents in the field of current economics. It can not only reflect the living standard and quality of residents, but also reflect the development of economy. It is a very important index in economic research. The consumption structure is affected by many comprehensive factors, which in turn affect the development of these factors [1]. Through the study of consumption structure, we can deeply understand the specific content of residents' consumption, find the characteristics of residents' consumption, and play an important role in expanding domestic demand. To study the characteristics and types of Chinese family consumption support structure is helpful to recognize the situation of China's economic development, reasonably guide the effective allocation of resources and the path of economic growth, and optimize the path of economic development.

At this stage, most foreign scholars are more committed to the construction and application of demand function model in the study of residents' consumption structure, which brings a lot of innovation in research methods. At the same time, the entry point of foreign scholars' research on the consumption structure is novel and diverse. The research results of foreign scholars on consumption function and consumption system model have laid a solid theoretical and model foundation for the research on China's household consumption expenditure structure. However, most of the domestic scholars introduce new variables to innovate based on the existing models to inject new vitality into the research of consumption structure. Based on the research results at home and abroad, we can find that the AIDS model and Panel Data are mostly used in the research of the consumption structure, while VAR model, factor analysis, cluster analysis, and grey correlation analysis are mostly used in the research of influencing factors of the

consumption structure. However, these methods can only rank the influencing factors of the consumption structure, so the analysis results of the consumption expenditure system are one-sided, so the ELES model is introduced.

The ELES model is an extended linear expenditure system model, which assumes that people's demand for various goods in a certain period depends on people's income and the price of various goods. Moreover, people's demand for all kinds of goods is divided into two parts—basic demand and demand beyond basic demand, and they think that basic demand has nothing to do with income level. After the basic demand is met, residents arrange various nonbasic consumption expenditures according to certain marginal consumption tendency [2]. In the process of analyzing the Chinese household consumption expenditure structure system, the application of the ELES model is expected to improve the rationalization of the consumption expenditure structure.

2. Building ELES Model

According to the statistical data of the National Bureau of Statistics on the per capita annual income and expenditure of urban households, the ELES model takes the per capita disposable income of residents at time t as the education consumption expenditure and then adds the expenditure of cultural and entertainment supplies to other expenditures, to properly process the statistical data of the National Bureau of Statistics [3]. The basic expression of the extended linear expenditure system of China's household consumption expenditure is as follows:

$$V_{it} = p_{it}q_{it} + b_{it}\left(X_t - V_{0t}\right), \tag{1}$$

where t means time, Q_{it} is the demand of class i consumption items in year t, V_{it} is the consumption expenditure of category i consumption item in year t, and the consumption expenditure items represented by category i in V_{it} have been given above; p_{it} is the price of category i consumption item in year t, q_{it} is the basic demand of type i consumption item in year t; b_{it} is the marginal propensity to consume; and V_{0t} is the total basic demand expenditure in year t, and its expression is as follows:

$$V_{0t} = \sum_t p_{it}q_{it}. \tag{2}$$

In addition, X_t is the per capita disposable income of residents in year t. That is to say, the consumption expenditure of a certain kind of commodity in year t is equal to the sum of the basic consumption expenditure of this kind of commodity and the marginal consumption expenditure of this kind of commodity in year t [4]. Based on the basic model, the residual term is added to get the econometric model:

$$V_{it} = p_{it}q_{it} + b_{it}\left(X_t - V_{0t}\right) + \mu_i. \tag{3}$$

After deformation and finishing the above formula, the results are as follows:

$$V_{it} = a_{it} + b_{it}X_{it}. \tag{4}$$

The expression of parameter a_{it} is as follows:

$$a_{it} = p_{it}q_{it} - b_{it}V_{ot}. \tag{5}$$

Then, the total basic expenditure and various basic expenditures in the $t-th$ year are calculated as follows:

$$\left\{ V_{ot} = \sum_i \frac{a_{it}}{1 - \sum_i b_{it}} p_{it}q_{it} = a_{it} + b_{it}V_{ot} \right. . \tag{6}$$

Therefore, we can conclude that the cross-price elasticity of Chinese household consumption expenditure is as follows:

$$\varepsilon_{ijt} = \frac{\partial Q_{it}}{\partial p_{jt}} \cdot \frac{p_{jt}}{Q_{it}} = -\frac{b_{it}p_{jt}q_{jt}}{V_{it}}. \tag{7}$$

The self-price elasticity of Chinese household consumption expenditure demand is

$$\varepsilon_{iit} = \frac{\partial Q_{it}}{\partial p_{it}} \cdot \frac{p_{it}}{Q_{it}} = -\frac{b_{it}\left(X_t + p_{it}q_{it} - V_{ot}\right)}{V_{it}}. \tag{8}$$

In the same way, we can also get the basic demand expenditure of i commodity.

3. Analysis on the Structure System of Chinese Household Consumption Expenditure

3.1. Classification of Consumption Structure Types. The consumption demand of Chinese families is constantly showing the characteristics of diversification, the scope of consumption expenditure is gradually expanding, new consumption hot spots are constantly emerging, and the consumption structure of residents is also showing the characteristics of multilevel. According to the different levels of meeting consumption demand, consumption structure can be divided into survival data, development data, and enjoyment data [5]. Survival materials are the materials that meet people's most basic needs, that is, the necessary food, clothing, and housing for rest to meet people's basic needs of eating, wearing, and living. They are the materials that are necessary to maintain the simple reproduction of labor force and the physical and mental strength of workers. Development materials are the materials that meet the needs of human physical and intellectual development. To make people's physical and intellectual development in an all-round way, residents need to have the necessary living materials for education, science and technology, culture, health care, and social interaction. Means of enjoyment are the material and spiritual means of satisfying people's needs for enjoyment. They include high-end consumer goods and materials necessary for people's spiritual life, such as entertainment and tourism [6]. Survival data are the most basic consumption data, and its elasticity is the smallest; development data and enjoyment data are higher level consumption data. In the Statistical Yearbook published by the National Bureau of Statistics, consumption expenditure is divided into eight categories according to the specific forms

of consumption, including food, clothing, housing, family equipment, and transportation. From the perspective of the development level of consumption structure, the consumption structure can be divided into four types, as shown in Table 1.

3.2. Establishing Chinese Family Consumption Function. The consumption function is mainly used to measure the relationship between consumption and income. Assuming that consumption is mainly determined by permanent income, permanent income is the expected long-term income. Y_p is the permanent income. The expression is as follows:

$$Y_p = \theta Y + (1 - \theta) Y_{-1}, \tag{9}$$

where Y is the current income, Y_{-1} represents past revenue, θ is a constant, and the parameter value range is [0, 1]. Then, the form of consumption function is as follows:

$$C = kY_p. \tag{10}$$

The persistent income hypothesis suggests that rational consumers make consumption decisions to maximize their effects, not based on temporary income in the present period. Instead, consumption decisions are based on the level of income that can be maintained long-term, that is, the level of lasting income. The life cycle hypothesis relates consumption to lifetime income and property. It assumes that consumers are rational and able to use their income sensibly and consume. The only goal of consumer behavior is to maximize utility. The life cycle and persistent income hypotheses are highly similar, emphasizing that consumption is mainly determined by future income [7].

3.3. The Main Influencing Factors of Consumption Structure. There are many factors affecting the consumption structure of urban residents, which can be divided into macro- and microfactors, such as industrial structure, consumer prices, and family population; they can also be divided into economic and noneconomic factors, such as income, industrial structure, and social and cultural, scientific, and technological progress [8]. These factors do not exist alone but restrict and influence each other and work together on the consumption structure. Among them, the income of residents is the most important and direct economic factor. Because income constitutes the economic basis of residents' consumption demand, the amount of income directly affects the level of residents' consumption expenditure and the proportion of all kinds of goods or services consumption, namely, consumption structure [9]. Industrial structure and consumption structure promote each other. The industrial structure determines the product structure and then the consumer goods structure, which is further reflected in the consumption structure of residents. With the increase in income, people will put forward higher consumption demand. The industrial structure is constantly adjusted and optimized to meet the gradually upgraded consumption demand through the change of product structure, thus forming a virtuous cycle. In addition, consumption environment refers to the external and objective factors that consumers face in the process of survival and development and have a certain impact on consumers, including natural environment and social environment [10]. The natural environment includes sunlight, water, air, soil, and other natural conditions for human survival; the social environment includes infrastructure, consumption and circulation market, institutional environment, and other factors.

3.4. Analysis on the Structure System of Chinese Household Consumption Expenditure. Under the ELES model, through the analysis of the influencing factors and components of China's household consumption expenditure structure system, the final analysis results of China's household consumption expenditure structure system are obtained from the static and dynamic aspects, respectively, and the basic demand and marginal consumption tendency of China's household consumption expenditure structure system are shown in the results.

4. An Empirical Analysis of the Structure of Household Consumption Expenditure in China

4.1. Basic Needs Analysis. According to the definition of the consumption structure in the Statistical Yearbook, consumption expenditure is divided into food consumption, clothing consumption, residential consumption, household equipment and services consumption, medical and health care consumption, transportation and communication consumption, education, culture and entertainment services consumption, and other goods and services consumption, a total of eight consumption expenditures [11]. This classification method has been adopted since China Statistical Yearbook, and the specific data of consumption structure starting from 2015 are sorted out, as shown in Table 2.

The internal structure of household goods and services consumption expenditure is shown in Figure 1.

On this basis, the analysis results of the Chinese household consumption expenditure structure system are obtained from static and dynamic aspects.

4.2. Results of Structural Static Analysis

4.2.1. Consumption Level. Consumption level refers to the scale and level of life consumption and services used by individual consumers or the whole society in a certain period. The level of consumption is determined by social development and people's living standards [12]. According to the consumption situation of multiple regions, the trend of per capita disposable income of consumption level is obtained, as shown in Figure 2.

It can be seen from Figure 2 that great changes have taken place in China's household consumption level in the past two decades. The per capita disposable income has increased from 4091 yuan in 2001 to 25763 yuan in 2020, an increase of about 6.3 times. In the past two decades, the

Table 1: Types of household consumption expenditure structures in China.

Classification criteria	Types of consumption expenditure structures	Explain
The development level of consumption structure	Consumption structure of hunger and cold	Food consumption and clothing consumption are the focus, and the quality of food and clothing is poor; that is to say, the main part of consumption expenditure is used to meet the basic needs of life, and there are few development materials and enjoyment materials
	Food and clothing consumption structure	Food and clothing consumption structure is a medium-level consumption structure. Food consumption accounts for a large proportion of household consumption. When the income increases, the proportion of food consumption expenditure is basically unchanged, and the quality of food improves with the increase in income. The proportion of clothing consumption and housing consumption in household consumption is rising, and the quality of clothing and housing is improving. Consumption expenditure such as culture and services accounts for a small proportion
	Well-off consumption structure	The well-off consumption structure has reached a higher level. The proportion of food consumption in household consumption decreased significantly, the consumption expenditure of clothing and living decreased gradually, the consumption expenditure of medical and transportation increased continuously, and the consumption expenditure of culture and services increased significantly. The proportion of basic living consumption in household consumption is declining, mainly focusing on development materials and enjoyment materials
	Post well-off consumption structure	The proportion of food consumption expenditure in household consumption continued to decline, while the proportion of transportation and communication, education and culture, and service consumption expenditure increased significantly

Table 2: China's household consumption data from 2015 to 2020.

Project	Year					
	2015	2016	2017	2018	2019	2020
Disposable income	10129	11321	11759	13786	15781	19109
Consumption expenditure	7182	7943	8697	9997	11243	12265
Food	2710	2914	3112	3628	4260	4479
Clothing	687	801	902	1042	1166	1284
Live	733	809	904	983	1145	1229
Household appliances and services	407	446	498	602	692	787
Medical care	528	600	621	699	786	856
Traffic communication	844	997	1147	1358	1417	1683
Education, culture, and entertainment services	1033	1098	1203	1329	1358	1473
Other goods and services	240	278	309	358	418	474

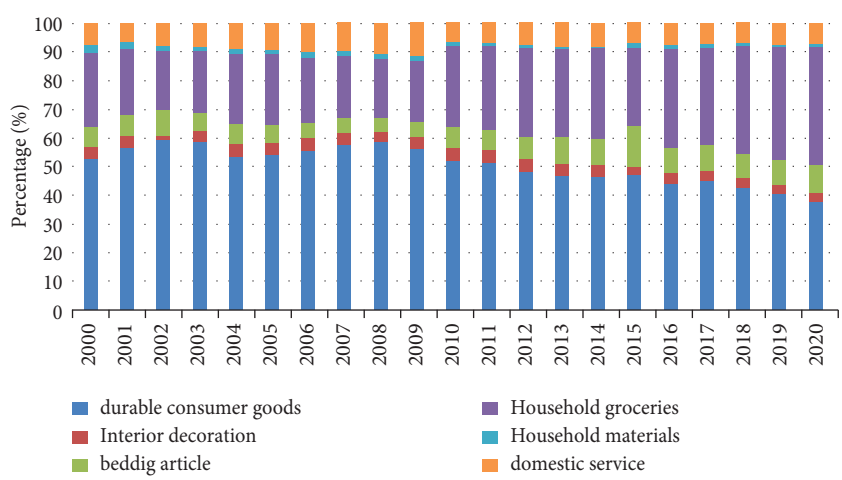


Figure 1: Internal structure of consumption expenditure on daily necessities and services.

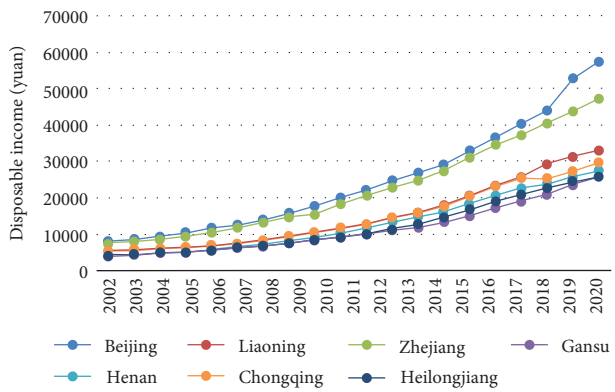


Figure 2: Trend of per capita disposable income in different regions.

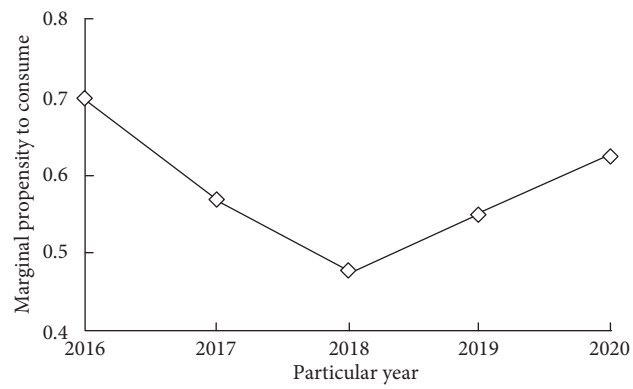


Figure 3: Marginal propensity to consume of Chinese household consumption structure.

growth rate of per capita disposable income has not fluctuated greatly. It shows a trend of increasing first, keeping stable in the medium term, and then declining. In the past two years, the economy has entered a new normal. Affected by the domestic economic slowdown and changes in the provincial economic structure, the income growth rate is slow [13]. The per capita consumption expenditure increased by 5.6 times from 3213.42 yuan in 2001 to 18145 yuan in 2020. In terms of the growth rate of consumption expenditure, the growth rate of consumption expenditure in 2004, 2007, 2011, and 2017 was rapid, while the growth rate in other years was small, and the growth rate was fluctuating. The per capita consumption expenditure and per capita disposable income showed a synchronous growth trend, but its growth rate was slightly lower than that of per capita disposable income.

4.2.2. Average Marginal Propensity to Consume. Marginal propensity to consume refers to the proportion of increased consumption in increased income. According to the annual per capita consumption expenditure, per capita disposable income, increased consumption, and income of Chinese households from 2016 to 2020, the value of marginal propensity to consume is calculated, as shown in Figure 3.

According to the law of diminishing marginal propensity to consume, although people's consumption increases with the increase in income, the part used to increase consumption in the increased income is less and less. From 2016 to 2020, China's household marginal propensity to consume is within the limit of marginal propensity to consume (between 0 and 1), basically fluctuates between 0.5 and 1, and does not show an obvious decreasing trend, which does not conform to the law of diminishing marginal propensity to consume, indicating that China's household consumption confidence is insufficient [14]. In 2018, the marginal propensity to consume dropped to the lowest value of 0.47394, indicating that the consumer confidence of Chinese households was once in a high state. However, due to the slowdown of macro-economic growth, the aggravation of the subprime mortgage crisis, the increasing pressure of inflation, and other adverse factors, the marginal propensity to consume showed an upward trend at the beginning of 2008, which indicated that the consumer confidence of Chinese households was unstable in recent years.

4.3. Results of Structural Dynamic Analysis

4.3.1. Structural Change of Consumption Expenditure. The change degree of consumption structure measures the change degree of residents' consumption structure every year. The larger the calculated value is, the greater the change range of residents' consumption structure in this period is. Combined with the consumption expenditure data of China in recent years, Chinese families are divided into three stages: 1997–1999, 2000–2013, and 2014–2020. These three stages correspond to the well-off, affluent, and richest life of Chinese families in turn. The specific statistics of structural change are shown in Table 3.

From the calculation results in Table 3, we can see that the average annual change of China's household consumption structure was 2.2 percentage points from 1997 to 1999, 0.73 percentage points from 2000 to 2013, and 3.73 percentage points from 2014 to 2020. According to the criteria for dividing social stages, the Engel coefficient was 45.9% in 1997 and 38.4% in 2000. At this stage, China has completed the transition from a well-off society to a relatively rich society [15]. Therefore, the change of consumption structure in this period is relatively large. From 2000 to 2013, China's household Engel coefficient was between 30% and 40%, showing a trend of first decline and then increase. Currently, the change of consumption structure is lower than that of the previous period. On the one hand, the change range of the previous period is large, which is difficult to maintain; on the other hand, it needs a long transition period from a well-off society to a more affluent society, and the residents' consumption structure is gradually improving, and the residents' rational consumption consciousness is also constantly improving. From 2014 to 2020, the Engel coefficient of Chinese families is less than 30%, and people's life has initially entered the richest stage. Therefore, the change of consumption structure in this period is relatively high.

Table 3: Changes of basic consumption structure of Chinese households.

	1997–1999 (%)	2000–2013 (%)	2014–2020 (%)
Household appliances and services	1.90	1.05	2.83
Medical care	1.39	0.83	0.93
Traffic communication	0.29	0.52	0.22
Education, culture, and entertainment	0.79	0.584	0.93
Live	0.68	0.59	0.41
Other goods and services	0.96	0.66	1.03
Changes in consumption structure	0.29	0.50	4.39
Average annual consumption structure change	0.30	0.41	0.45
Household appliances and services	6.60	5.10	11.18
Medical care	2.20	0.73	3.73

4.3.2. Elastic Analysis. According to the calculation principle of demand income elasticity, we can calculate the demand income elasticity of Chinese households' consumption expenditure. The demand income elasticity of food, clothing, housing, family equipment, medical treatment, transportation, culture and education, and other items are 0.763, 0.889, 0.958, 0.696, 0.721, 1.119, 0.648, and 0.842 respectively. Therefore, the demand income elasticity of residents' consumption is positive, which indicates that with the increase in disposable income, residents' consumption expenditure will increase accordingly.

Based on the analysis of basic consumption demand, marginal consumption tendency, and demand income elasticity of the ELES model, the basic consumption demand and marginal consumption tendency of food are in the first place; the marginal consumption propensity and demand income elasticity of transportation, communication, and housing are the largest, and the consumption of transportation, communication, and housing is developing into a hot area of urban residents' consumption in Hebei Province; people will pay more and more attention to medical treatment, culture, and education, and gradually change into the development type and enjoyment type of consumption [16].

5. Suggestions on Promoting the Optimization of China's Household Consumption Expenditure Structure

Consumption is of great significance in the development of national economy, and the current economic situation makes it particularly important to promote consumption. The government should optimize the structure of Chinese household consumption expenditure from many aspects.

5.1. Increasing the Income of Rural Residents through Various Ways. As far as the current economic situation is concerned, measures can be taken from the following aspects: first, accelerate the reform of state-owned enterprises, accelerate the improvement of the overall efficiency of state-owned enterprises, improve their economic efficiency and increase workers' wages; vigorously develop various forms of collective economy, encourage and support the healthy development of individual economy and private economy, create more employment opportunities, and increase residents' income. Support laid-off workers and other low-income people to ensure the minimum living standard of the city; encourage secondary entrepreneurship through technical training, policy support, financial assistance, etc., vigorously develop the individual and private economy, and gradually increase their income [17]. Encourage and support residents to broaden investment channels in various forms, such as investing in the stock industry; purchasing treasury bonds; and participating in insurance, real estate, education, tourism, and other industries, to realize the continuous appreciation of residents' assets. Finally, it is necessary to implement a moderate tax reduction policy for the individual economy and small- and medium-sized private economy. The administrative departments for industry and commerce and the health and epidemic prevention departments should increase their support, and establish corresponding support policies, to accelerate the prosperity of the individual and private economy and effectively improve the income of operators.

5.2. Guide Reasonable Consumption Concept. As the main body of consumption, consumers and consumption environment work together and promote each other. Therefore, it is necessary to guide consumers to consume reasonably and scientifically, change traditional consumption concepts, abandon the consumption habits of comparison and extravagance, and establish a correct and reasonable consumption concept. On the one hand, we need to improve the quality of consumers and update the concept of consumption through education and guidance, so that people can correctly understand the importance of scientific consumption mode to the harmonious development of human, society, and ecological environment. On the other hand, we should pay attention to the consumption of spiritual culture [18]. Complete material consumption is bound to fail to fully meet people's consumption needs. Therefore, the government needs to increase investment in spiritual culture, pay equal attention to spiritual consumption and material consumption, coordinate the development of the two, increase the proportion of Chinese education and entertainment in the consumption structure, improve the level of consumption, and further optimize the consumption structure, to effectively promote economic growth and achieve all-round social progress.

5.3. Optimize the Consumption Environment. As an external condition, consumption environment determines the utility that consumer goods can bring to consumers and then has an important impact on the consumption mode and behavior choice of residents. First, improve the consumer market laws and regulations. We should strengthen legal publicity, strengthen market supervision, ensure product quality, crack down on fake and shoddy commercial fraud, establish a good and orderly market order, and purify the consumption environment. To protect the legitimate consumption rights and interests of residents. Secondly, strengthen infrastructure construction [19]. Ensure the basic service conditions of power supply network, water supply network, information network, etc., according to the characteristics of different cities or regions, classify and make reasonable planning, and establish shopping places that meet the economic development level of Hebei Province [20, 21].

5.4. Expanding the Field of Consumption and Cultivating New Hot Spots of Consumption. In the process of adjusting the consumption structure of residents, in addition to the diversification and personalized supply of consumer goods in the traditional consumption field, we should continue to expand new consumption fields. According to the existing traditional market, we should find new potential consumption markets, such as information consumption and green consumption, to form new consumption hot spots and realize the sustainability of consumption structure optimization.

6. Conclusion

From the national level, according to the marginal propensity to consume, income elasticity of demand, and price elasticity of all kinds of consumer goods, we can know that the marginal propensity to consume food in China's household consumption expenditure structure is gradually decreasing, the marginal propensity to consume health care and housing is gradually increasing, and the household consumption expenditure structure is gradually changing to the development-type and enjoyment-type structure. From the regional level, due to different income and prices, there are differences between different provinces. Therefore, analyzing the results of residents' consumption expenditure can optimize the structure of residents' consumption expenditure and help rationalize the system of household consumption expenditure in all regions of China.

Data Availability

The labeled dataset used to support the findings of this study is available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This study was sponsored by Universiti Putra Malaysia.

References

[1] A. Galkin, K. Mykola, I. Balandina et al., "Assessing the impact of population mobility on consumer expenditures while shopping," *Transportation Research Procedia*, vol. 48, pp. 2187–2196, 2020.

[2] K. C. Konduri, S. Astroza, B. Sana, R. M. Pendyala, and S. R. Jara-Díaz, "Joint analysis of time use and consumer expenditure data: examination of two approaches to deriving values of time," *Transportation Research Record*, vol. 2231, no. 1, pp. 53–60, 2011.

[3] M. Siahpush, P. A. Farazi, S. I. Maloney, D. Dinkel, M. N. Nguyen, and G. K. Singh, "Socioeconomic status and cigarette expenditure among US households: results from 2010 to 2015 Consumer Expenditure Survey," *BMJ Open*, vol. 8, no. 6, Article ID e020571, 2018.

[4] N. B. Bulamu, B. Kaambwa, L. Gill, I. D. Cameron, and J. Ratcliffe, "An early investigation of individual budget expenditures in the era of consumer-directed care," *Australasian Journal on Ageing*, vol. 39, no. 1, pp. e145–e152, 2020.

[5] F. Aschauer, I. Rösel, R. Hössinger, H. B. Kreis, and R. Gerike, "Time use, mobility and expenditure: an innovative survey design for understanding individuals' trade-off processes," *Transportation*, vol. 46, no. 2, pp. 307–339, 2019.

[6] J. Sen and D. Das, "Commodity-specific consumer expenditure in India: pro-poor or pro-rich," *Economics Bulletin*, vol. 38, no. 1, pp. 501–508, 2018.

[7] P. Alexander and L. Poirier, "Did US consumers respond to the 2014-2015 oil price shock? Evidence from the Consumer Expenditure Survey," *Energy Journal*, vol. 41, no. 1, 2020.

[8] A. Sood, A. Mishra, and A. Ambekar, "Trends in alcohol consumption and expenditure: analysis of household data from nationally representative sample from India," *Asian Journal of Psychiatry*, vol. 40, pp. 116–123, 2019.

[9] K. M. Cutright, "The beauty of boundaries: when and why we seek structure in consumption," *Journal of Consumer Research*, vol. 38, no. 5, pp. 775–790, 2012.

[10] L. Cherchye, T. Demuynck, and B. D. Rock, "Bounding counterfactual demand with unobserved heterogeneity and endogenous expenditures," *Journal of Econometrics*, vol. 211, no. 2, pp. 483–506, 2019.

[11] Q. Brummet, D. Flanagan-Doyle, J. Mitchell, J. Voorheis, L. Erhard, and B. McBride, "What can administrative tax information tell us about income measurement in household surveys? Evidence from the Consumer Expenditure Surveys," *Statistical Journal of the IAOS*, vol. 34, no. 4, pp. 513–520, 2018.

[12] K. Forrester and J. Klein, "An analysis of female labor supply, home production, and household consumption expenditures," *Journal of Demographic Economics*, vol. 84, no. 3, pp. 257–307, 2018.

[13] A. Khan, N. Javaid, and M. I. Khan, "Time and device based priority induced comfort management in smart home within the consumer budget limitation," *Sustainable Cities and Society*, vol. 41, pp. 538–555, 2018.

[14] O. Desyatnyuk and V. Sidliar, "Fiscal regulation of the Ukrainian consumer market," *World of finance*, vol. 1, no. 1(62), pp. 78–94, 2020.

[15] Z. Yang, Y. Fan, and L. Zhao, "A reexamination of housing price and household consumption in China: the dual role of

housing consumption and housing investment," *The Journal of Real Estate Finance and Economics*, vol. 56, no. 3, pp. 472–499, 2018.

[16] O. Shittu, "Emerging sustainability concerns and policy implications of urban household consumption: a systematic literature review," *Journal of Cleaner Production*, vol. 246, Article ID 119034, 2020.

[17] Y. Ling, G. Shu, L. Wenchang, and H. U. Lin, "Rural household energy consumption investigation and structural pattern analysis of xuanwei study area in central Yunnan IOP conference series: earth and environmental science," *IOP Conference Series: Earth and Environmental Science*, vol. 565, no. 1, Article ID 012017, 2020.

[18] B. S. Lee, "Effects of household income and rate of consumption-oriented spending on children on mother's parenting stress: a short-term longitudinal study using a latent growth model," *Korean Journal of Child Studies*, vol. 39, no. 1, pp. 61–74, 2018.

[19] D. Borozan, "Regional-level household energy consumption determinants: the European perspective," *Renewable and Sustainable Energy Reviews*, vol. 90, pp. 347–355, 2018.

[20] J. Li, J. Zhang, D. Zhang, and Q. Ji, "Does gender inequality affect household green consumption behaviour in China?" *Energy Policy*, vol. 135, Article ID 111071, 2019.

[21] K. Mori, H. Muta, and Y. Ohtori, "Development of interaction model on the risk assessment method for nuclear facilities using a system model with a multi-layer structure," *Journal of Nuclear Science and Technology*, vol. 58, no. 5, pp. 542–566, 2021.

Hindawi
Computational Intelligence and Neuroscience
Volume 2022, Article ID 6453499, 10 pages
https://doi.org/10.1155/2022/6453499

Research Article

Research on Classroom Emotion Recognition Algorithm Based on Visual Emotion Classification

Qinying Yuan

Xi'an Jiaotong University, Shaanxi, Xi'an 710049, China

Correspondence should be addressed to Qinying Yuan; 150776@peihua.edu.cn

Received 10 June 2022; Revised 15 July 2022; Accepted 19 July 2022; Published 8 August 2022

Academic Editor: Ning Cao

In this paper, we construct a classroom emotion recognition algorithm by classifying visual emotions for improving the quality of classroom teaching. We assign weights to the training images through an attention mechanism network and then add a designed loss function so that it can focus on the feature parts of face images that are not obscured and can characterize the target emotion, thus improving the accuracy of facial emotion recognition under obscuration. Analyze the salient expression features of classroom students and establish a classification criteria and criteria library. The videos of classroom students' facial expressions are collected, a multi-task convolutional neural network (MTCNN) is used for face detection and image segmentation, and the ones with better feature morphology are selected to build a standard database. A visual motion analysis method with the fusion of overall and local features of the image is proposed. To validate the effectiveness of the designed MTCNN model, two mainstream classification networks, VGG16 and ResNet18, were tested and compared with MTCNN by training on RAF-DB, masked dataset, and the classroom dataset constructed in this paper, and the final accuracy after training was 78.26% and 75.03% for ResNet18 and VGG16, respectively. The results show that the MTCNN proposed in this paper has a better recognition effect. The test results of the loss function also show that it can effectively improve the recognition accuracy, and the MTCNN model has an accuracy of 93.53% for recognizing students' facial emotions. Finally, the dataset is extended with the training method of expression features, and the experimental study shows that the method performs well and can carry out recognition effectively.

1. Introduction

Human emotions are closely related to behavior. Emotions originate from objective facts and are the key to determining the quality of human daily life, and they are related to human needs. When human needs are met, humans tend to generate positive emotions and show positive behaviors [1]. The classroom teaching environment is different from the one-on-one online tutoring teaching environment and offline tutoring teaching environment because the number of students is large, and the teacher cannot pay attention to each student's emotional condition and always give feedback while considering the course schedule [2]. The emotional state of students in the learning process may have a positive or negative impact on the overall learning results, which leads to students in a depressed mood because they cannot adjust their state in time to reduce enthusiasm for learning, affecting the efficiency of the entire course, which is undoubtedly half the effort. The pursuit of teaching efficiency is the essential characteristic of teaching, which is also an important goal of the current curriculum reform and an inevitable requirement for the internal development of education [3]. In the process of classroom teaching, teachers hope to understand students' emotional changes and learning status in real time, adjust the course progress in time, and ensure that students can grasp the content taught in a timely and comprehensive manner. According to Mehrabian, 55% of emotional information is expressed through facial expressions [4]. Therefore, facial expression recognition can determine students' emotional state and provide timely feedback to the teacher, so that the teacher can judge the current classroom students' emotional change in real time, make a timely course adjustment, activate the classroom

atmosphere, mobilize students' learning enthusiasm, and improve the overall classroom learning efficiency.

Sentiment analysis, also called opinion mining, is the process of processing and analyzing user-generated content to study the emotions, opinions, and attitudes expressed by users and to judge or evaluate them. Sentiment, as an important attribute of humans in social behavior, represents an individual's viewpoint and attitude toward a specific goal and has a great impact on all aspects of human life such as cognition, socialization, decision-making, and reasoning [5]. The current work on sentiment analysis in social media is mainly focused on text analysis, and a lot of research results have been achieved, but the research on sentiment analysis based on visual content such as images and videos is still very limited. As the saying goes, "a picture is worth a thousand words," compared with textual content, images and video content are simpler and more intuitive; especially, with the development of multimedia technology and the popularity of photographic devices, more users tend to use visual content such as images to express their views and emotions, and the amount of visual content in social media is therefore growing. The amount of visual content in social media is growing as a result [3]. The identification and retrieval of these visual contents have been widely used in various fields; however, the emotions and opinions of users conveyed by these visual contents largely reflect their behavioral preferences and needs, so it is also important to deeply explore the attitudes and emotions in them for public opinion monitoring and enterprise decision-making. Although the emotions embodied in visual contents are complex and subjective, there are commonalities in human cognition of emotions, which is the basis of visual sentiment analysis [6]. Traditional research on visual sentiment analysis mainly uses image low-level visual features for sentiment analysis, such as color features, texture features, and contour features. However, these algorithms are not sufficient to bridge the huge semantic gap between low-level visual features and high-level sentiment semantics of images, so the effect of sentiment classification is not satisfactory.

In recent years, artificial intelligence plays an increasingly important role in education, and AI and deep learning-related technologies (such as image recognition, semantic recognition, and speech recognition) are gradually implemented into specific fields, providing technical support for the development of adaptive education. The conference pointed out that one of the applications of AI in education is real-time monitoring of students' learning expressions, learning postures, and other valuable information [7]. In traditional classrooms, the number of students is large and the teacher needs to control the teaching pace, so it is difficult to pay attention to each student's emotional state and make corresponding adjustments to feedback during the classroom teaching process, which makes students unable to adjust their state in a time when they are depressed, thus reducing their enthusiasm for learning, and affecting teaching efficiency, which is undoubtedly half the effort [8]. The traditional classroom evaluation methods, such as postclass conversations and questionnaires, are too delayed and cannot help teachers make timely adjustments. Therefore, how to provide teachers with timely and effective feedback on students' emotional states in class has become an urgent problem in classroom teaching. This paper constructs an algorithmic model for classroom emotion recognition with the help of artificial intelligence, which has positive significance for technology to focus on students' classroom emotions and improve the quality of classroom teaching.

2. Related Works

Traditional visual sentiment analysis methods mainly use knowledge from fields such as visual cognition and psychology to extract underlying visual features such as color, texture, and shape of images and perform sentiment analysis through statistical methods [9]. The University of Chile has developed a texture retrieval system based on soft computing techniques. By studying the human subjective psychological perception of texture, the system obtains qualitative descriptions of texture consisting of 12 adjectives, which can be further used for texture querying, and evaluates 100 textures by selected adjectives to associate texture features with these subjective qualitative descriptions through neural networks. Yin et al. achieved high accuracy by detecting the shape of straight lines in images, further extracting the straight-line orientation histogram features of the images, and fusing other features for sentiment recognition [10]. However, these methods extract features based on simple statistical methods and do not consider the human emotion element. Hossain combines human emotion and art painting theory based on psychological principles, thus achieving emotion category labeling of art images [7]. Since traditional visual sentiment analysis methods are not sufficient to overcome the semantic gap problem, some researchers have started to construct intermediate semantic representations around the semantic content of images to better express the sentiment of images. The most representative of these works is the approach proposed by Kim H H et al. which is by designing an adjective noun pair (ANP) and using the ANPs corresponding to different images as intermediate representations of visual sentiment, e.g., beautiful flower, lovely dog, etc. [11]. However, there are still some problems in the practical application of the visual emotion ontology library VSO. First, the ANPs in VSO can only describe the visual concepts in the images and cannot identify which ANP is highly related to the main emotion in the images; second, the images in social networks have different themes, which can be divided by themes, and a theme usually contains many related images. However, the VSO model cannot fuse the sentiment information of multiple related images under the same topic and thus cannot select more representative ANPs to describe the sentiment of images. To solve the above problems, Li et al. proposed a visual emotion-based topic model, which tries to enhance the understanding of visual semantic objects in images and their emotions through topics in social networks [12].

A good teacher can accurately grasp classroom emotions to improve students' attention, make them feel happy and satisfied, and eventually make them enjoy their lessons,

while a teacher who is not good at grasping classroom emotions may cause the classroom to fall into a boring narrative, students' attention is not focused, and learning is inefficient. Scholars have done enough research on classroom emotions and have deeply analyzed teachers' and students' behaviors and emotional expressions in the classroom and what kind of teaching effect the classroom will have after these expressions. Professor Luo H argues that the affective dimension can be seen as a continuum of hierarchical order and describes emotions not only as simple or complex, figurative, or abstract interpretations but also as characteristics of control from the unconscious to the internal and external, which can be called internalization [13]. Rani et al. proposed a classification theory of educational goals, for which the goal classification system needs to be divided according to the emotions in it, and the dimensions of its division are mainly reflected in five levels, namely, acceptance, reaction, value judgment, organization, and characterization of value and value complexes, where each of the levels has its emotional meaning, and which also has a sublevel corresponding to that level [14]. For emotion recognition in a classroom environment, Zeng et al. obtained facial features through the AAM model combined with constrained local model (CLM) and classified classroom emotion categories into five types: listening, doubt, understanding, resistance, and disdain, and based on Izard's maximum discriminative facial action coding system, judged by facial offset angle, eye and eyebrow angle, and mouth angle [15]. The five expression states were judged by the changes in the characteristics of the three parts of the face, the angle between the eyes and the eyebrows, and the corner of the mouth.

In classroom teaching, teachers should pay attention to students' learning emotions and promptly adjust their teaching strategies based on the feedback from expressions to mobilize students' positive learning emotions. Scholars have conducted many studies on classroom teaching and learning, which have provided comprehensive and in-depth analyses of classroom teaching and learning processes from many different perspectives, and have thus contributed to further development of teaching and learning. Since 2000, the integration of computer sentiment analysis and educational teaching has been greatly promoted. Researchers have used video analysis to understand learners' emotional states, providing teachers with timely feedback so that they can revise course content, adjust course difficulty, select teaching methods, and control the teaching schedule, greatly promoting changes in the teaching model and improving teaching quality [16]. The biggest problem of emotion recognition research in a classroom teaching environment is the need to detect the unique emotion category for the unique environment, there is a serious lack of publicly available emotion datasets for the classroom environment, and due to the large change in students' posture and orientation in the classroom environment, the existence of such actions as lowering the head and turning sideways will lead to incomplete face pictures and coupled with the influence of background, illumination, and occlusion, the available face images are less. How to recognize the emotion and achieve a high accuracy rate despite the presence of facial occlusion is also a focus of the research.

3. Visual Emotion Classification Classroom Emotion Dataset Construction

In classroom teaching, students' expressions and behaviors can greatly express their mental and emotional states. For courses they are interested in or like, they tend to show happy, focused, and positive emotions, with facial expressions such as smiles and concentration. For courses that they do not like or even dislike, they tend to show negative emotions such as resistance and disdain, and their faces are accompanied by negative expressions such as frowns. These negative emotions are partly caused by the difficulty of the course and the difficulty of adapting to the teacher's teaching style, while another part is caused by the students' psychological problems. Therefore, the identification and analysis of students' emotions in the classroom environment can not only help teachers to make better adjustments to the course teaching but also help to detect whether students have psychological problems in time and provide timely psychological counseling and other treatments for students with psychological problems, to promote the overall development of students' body and mind. This paper realizes the prediction of dimensional emotional data based on related networks such as recurrent neural network and finally uses the quantitative algorithm of emotional eigenvalues to calculate emotional intensity, which provides a database for the system to perform adaptive adjustment.

Get the students' dynamics in class in real time through the camera, then extract features through image processing, machine learning, deep learning, and other technologies, classify emotions, and feed back the real-time emotional status to the teacher in time. The teacher makes corresponding adjustments according to the students' current emotional status. When students are depressed, adjust the teaching progress, activate the classroom atmosphere, and improve students' learning motivation. When students are generally in a low mood, the teacher will adjust the teaching schedule to improve students' learning enthusiasm and efficiency; when students are generally in a high mood, the teaching schedule can be accelerated to further improve the teaching efficiency of the course.

The convolutional neural network is a special kind of feed-forward neural network with numerous neurons in its structure, and these neurons come with weights and biases. It uses local connections to avoid the problem of redundant data caused by full connections [17]. In the convolutional neural network, its local connectivity can effectively reduce the number of parameters in the network, which can reduce the dependence of the model on a large amount of training data. In simple terms, for a two-dimensional image data, the color image is input to the neural network by default, and then the size of the color image is set to the standard data format: depths $\times$ height $\times$ width. Multi-task convolutional neural network (MTCNN) provides a multi-task cascade framework for face detection, which is characterized by two steps of face detection and key point localization being

performed simultaneously. When an arbitrary image is given, MTCNN scales it to different scales to form an image pyramid for scale invariance.

$$p(v, h|\theta) = \ell^{Ev,z(\theta)}, \tag{1}$$

m is the number of neurons in the hidden layer, n is the number of neurons in the visible layer, and W_{mn} is the weight matrix between them, $v = [v_1, v_2, \ldots, v_n]$ is the state of the visible layer, and $h = [h_1, h_2, \ldots, h_m]$ is the state of the hidden layer. For example, if we assume that there are n visible units in the visible layer and m hidden units in the hidden layer, the energy function between the nodes in the visible layer and the nodes in the hidden layer can be expressed as follows when the values of both visible and hidden units are 0 or 1.

$$E(v, h|\theta) = \sum_{i=1} a_i v_i + \sum_{j=1}^{n} v_i b_j w_{ij}. \tag{2}$$

In the formula, v_i represents the state of the ith unit of the visible layer, h_j represents the state of the jth unit of the hidden layer, w_{ij} represents the connection weight between v_i and h_j, and a_i and b_j represent v_i and the bias value of h_j.

A deep neural network (DNN) is also a neural network, but it has multiple layers of perceptron and multiple hidden layers and consists of a series of Boltzmann machines (RBMs) stacked and trained layer by layer. The lower layers represent the original information of data features, while the higher layers represent the attribute categories of data, so its learning process proceeds from the lower to the higher layers. In the learning process, the higher layers will continuously acquire deep abstract features, which is the essential feature of deep learning data. According to the above research status, although researchers have proposed many feature extraction methods and recognition classification methods for pictures, there are still some problems to be solved if we want to realize a system of expression recognition with high accuracy and in real time: one is the problem of how to perform multi-person face detection in real time; the other is the problem that some external factors and similarity between faces under natural conditions can affect the accuracy rate.

$$p(h_i = 1|v) = \delta\left(\sum_{i=1} a_i w_{ij} - b_j\right) \tag{3}$$

In 1976, Paul Ekman created the facial action coding system (FACS), which divided the face into independent and related AU—action unit—based on the anatomical characteristics of different muscles of the face. The FACS provides a detailed classification and elaboration of motor units and expression composition, which has become the main theoretical basis for later research, and many classical algorithms or expression databases are based on the FACS. The classroom expression status categories should indicate whether the learners are interested in the content, whether they agree with the teacher's teaching style, and whether they are comfortable with the pace of knowledge taught. Teachers can use this as a basis for adjusting the content and teaching process and providing timely emotional compensation. Teaching observers can use this to evaluate the quality of teaching, reflect on teaching activities, and make changes. Since the focus and application scenarios of classroom expressions are different, they are both the same and different from the common expressions.

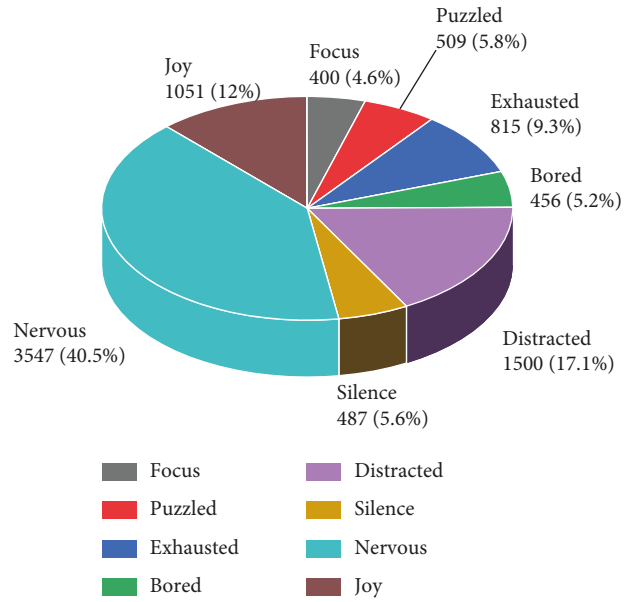


Figure 1: Number of samples collected for each type of emotion.

The discrete emotion description model describes emotions into discrete categories, such as happy, angry, and sad. Ekman, an American psychologist, proposed 6 basic human emotions: happy, angry, surprised, sad, disgusted, and afraid. The Chinese emotion corpus of the Chinese Academy of Sciences Institute of Automation (CASIA) also basically follows these six basic emotion categories, except that disgust is replaced with neutrality according to the corpus data [18]. The present dataset refers to a wide range of human emotions as well as the actual emotional performance in traditional secondary school classrooms and finally identifies eight categories of emotions: attentive, confused, tired, bored, distracted, silent, nervous, and pleasant. The emotions of each speech sample in the dataset were jointly labeled by three people, and the final emotion labels of the samples were determined by the principle of minority rule. The number of samples in each category of emotions in the dataset is shown in Figure 1. The database contains three kinds of labels: emotion learning behavior label, discrete emotion label, and dimensional emotion label (two-dimensional (arousal-valence) emotion), subjects labeled using specific emotion words, labeled learning behavior, and discrete learning expressions. The recorded videos were labeled using discrete learning emotion labels (common emotions in online learning: confused, focused, distracted, tired, happy, and bored), and the experimenter helped the subjects to examine the videos and extract the emotion clips related to academic emotions. Ensure the completeness of image sequences, the consistency of similar emotion annotations, and re-tagging for images that do not meet the conditions. By combining previous articles and comparing with the latest research, the

article finally identified 8 classroom recognition emotions of concentration, confusion, exhaustion, boredom, distraction, silence, tension, and pleasure.

Because of the flexible nature of human facial expression, the change of expression is nonrigid, and considering the requirement of real-time detection of multiple targets, this study will use the multi-task convolutional neural network (MTCNN) proposed in 2016 in the part of implementing face detection. It has very good experimental and application effects, MTCNN uses three different depths of network types to achieve face detection, and its cascade structure + CNN + bounding box method can achieve real-time face detection, which can better meet the task of multiple face detection in classroom scenes. The captured video is fed into a multi-task convolutional neural network (MTCNN), which detects the faces obtained from the video in real time and locates the key parts of the face: eyes, nose, and mouth. This model mainly uses three cascade networks, P-Net, R-Net, and O-Net, to detect and locate faces more accurately by stepwise fine-tuning.

The original face images have the problem of pose and scale differences due to different shooting conditions. Geometric normalization includes face correction and scale scaling, which can correct the angularly shifted faces and scale the faces of different scales to a uniform scale for subsequent dataset construction and use. Face correction is based on the angular difference between the original face pose and the standard face pose, and the original image is corrected to the standard face pose by coordinate rotation to ensure the consistency of face orientation. The center coordinates of the left and right eyes in the intercepted face image are obtained by feature point detection, the angle between the line connecting the center coordinates of the two eyes and the horizontal direction is calculated, and the original image is rotated by coordinates with the center of the line connecting the center coordinates of the two eyes as the rotation center to obtain the orthographic image.

$$\theta = \arccos \frac{y_{\text{right}} - y_{\text{left}}}{x_{\text{left}} + x_{\text{right}}}. \tag{4}$$

The original face images often have different face sizes and scales due to differences in shooting locations and individual differences. Through scale normalization, the faces in the original images are scaled to a uniform standard to reduce or even eliminate noise interference, making the accuracy of the subsequent expression recognition algorithm evaluation more reliable. Scale normalization includes cropping and scaling of the corrected face images. The feature points of the corrected face image are calculated according to the previous (4) as the distance d between the two eyes, the center of the line connecting the two eyes is used as the origin, the area with distance d is cropped to the left and right, and the area with distance d and $1.5\,d$ is cropped upward and downward, respectively, to obtain the standard face rectangular area. The intercepted standard face area is scale transformed to a uniform size by bilinear interpolation to achieve normalization in scale. This paper realizes the prediction of dimensional emotional data based on related networks such as recurrent neural network and finally uses the quantitative algorithm of emotional eigenvalues to calculate emotional intensity, which provides a database for the system to perform adaptive adjustment.

4. Classroom Emotion Recognition Model Design

To achieve a deeper network to extract deeper features while reducing the network parameters and improving the computational efficiency, a bottleneck architecture is used for the residual unit. A 1×1 convolutional kernel is used at the beginning and the end, and only a 3×3 convolutional kernel is used in the middle. The information transfer in the residual cell consists of two types of mapping, namely, residual mapping and constant mapping. The mapping generated by the convolutional and activation operations is called residual mapping; at the same time, it allows the direct transfer of information between residual units in the form of jump connections, and the mapping generated by this direct connection is called constant mapping, which is directly used as the input of the next residual unit and propagated from the top to the bottom layer. In traditional neural networks, some information is lost when the information is passed in the convolutional and fully connected layers. The information transfer between residual units can be expressed as the following equation.

$$x_{l+1} = x_l + f(x_l) \times f(w_l), \tag{5}$$

x_l denotes the constant mapping, f denotes the residual mapping, and w_l denotes the convolution kernel parameter. From this equation, the output of any layer can be represented by the output of any layer shallower than it and the sum of the residuals between the two layers. The input image is passed through the deep residual network, and the final output is obtained as the feature representation $F \in R^{W \times H \times C}$ of the overall image feature.

For the training set $(x_i, y_i) = 1$, where x_i denotes the sentiment image samples, $y_i \in \{1, 2, \ldots, n\}$ like the sentiment labels corresponding to the samples, N denotes the number of image samples in the training set, and k denotes the number of sentiment categories. For each sentiment image sample, the overall feature, i.e., the output of the last convolutional layer of the deep convolutional neural network, is $F \in R^{W \times H \times C}$. For the feature map of each channel, the richer the sentiment information contained in a location, i.e., the more strongly the location expresses the sentiment, the larger the value of that location in the feature map. The sentiment activation map is then generated by CAM using the image-level sentiment labels.

$$\begin{aligned} V_i &= (WH) \sum_{n=1}^{m} f(m, n), \\ \text{frature} &= \sum_{i=1} \varpi_i \cdot \text{sum}(r_i) + N. \end{aligned} \tag{6}$$

The process of emotion recognition algorithm is mainly divided into three steps: face detection, feature extraction, and emotion recognition. The original image to be detected

first needs to be preprocessed, and the preprocessing includes some digital image processing operations on the original image and face detection operations. Digital image processing operations such as histogram equalization are used on the original image to make it easier to detect and extract faces in the face detection process [19]. Since the existing images for emotion recognition are of high quality and face detection is the most important part of the preprocessing process, most of the literature related to emotion recognition uses face detection as preprocessing. Next is the feature extraction process, where the acquired faces are extracted by feature extraction methods to characterize the current facial expressions. Finally, the emotion classification is performed, and the emotion features extracted by the feature extraction algorithm are classified by existing classification methods to give the corresponding emotion list labels of the input face images. Compared with the traditional emotion recognition process, for a specific classroom environment, it is necessary to first obtain the real-time video of the classroom teaching process through the camera, use tools such as OpenCV to obtain each frame of the acquired real-time video as the original image, and then carry out the traditional emotion recognition process to get the corresponding emotion category labels of the input face images. It is necessary to analyze the emotions of the classroom students identified in a period. Finally, we need to analyze the classroom students' emotions identified in a period and provide timely feedback to the teacher, to help the teacher adjust the teaching strategy. The classroom student emotion recognition algorithm designed in this paper is shown in Figure 2.

In recent years, an increasing number of researchers have started to combine deep learning techniques with visual sentiment analysis. The performance of a deep learning model is largely determined by the structure of the deep network, and therefore, deep learning models with different structures are often designed and employed to suit different tasks. In this section, classical deep learning models are introduced, with a particular focus on convolutional neural networks, which are widely used in visual sentiment analysis tasks. The partial connection method it uses can effectively avoid the redundant data problem caused by the full connection. And in the convolutional neural network, its local connection method can effectively reduce the amount of parameters of the network, which can reduce the dependence of the model on the huge amount of training data.

Visual feature extraction mainly uses computer vision-related technologies and digital image processing-related technologies to extract visual features closely related to human emotions and moods from image and video data according to the psychological and physiological characteristics of human beings. According to the level of the extracted features, the features can be classified into three categories: bottom-level features, middle-level features, and top-level features. Traditional machine learning methods rely more on manual design and extraction of features when processing raw data, which makes traditional machine learning algorithms have great limitations. Compared with traditional machine learning algorithms, deep learning enables automatic learning of robust features from large-scale data, abandoning complex feature engineering, and at the same time, deep learning techniques have strong domain adaptation and model generalization capabilities. Due to the massive data available on the Internet and the development of computer hardware, deep learning techniques are widely used in various fields.

This section focuses on a qualitative approach to demonstrate the reliability of the labeled reporter labeling results, and the arousal-valence emotion model can be used to better identify students' emotions during the learning process compared to the discrete emotion model. As shown in Figure 3, the distribution of the eight emotion categories in the arousal-valence space shows that (1) a single emotion (e.g., pleasant) can lead to multiple arousal-valence values [20]. This suggests that each category of emotion may have a different arousal-valence distribution, which means that traditional discrete emotion categories may not accurately describe a person's internal emotions. (2) There is overlap between emotions, suggesting that different emotion categories may have similar arousal-valence distributions. For example, the arousal, valence values of some "focused" and "pleasant" images are very close to each other. This suggests that each person has a different understanding of linguistic features. In terms of description, the consistency of human categorization markers for emotions is quite poor. As can be seen, it is not easy to choose an emotion from many clear words to describe a person's emotion, because there are nuances between some emotion labels or relationships between emotions.

To provide a benchmark for automatic emotion classification on the database (SHZ-LSD), this paper validates the recognition based on convolutional neural network and recurrent neural network database contents, while the best model for emotion recognition is used in the adaptive learning system to build the emotion recognition module of the student model. The experiments are divided into emotion recognition for discrete emotions and emotion recognition for two-dimensional arousal-valence. Discrete emotion recognition uses a convolutional neural network, the experiments are performed on the original image data for face cropping, to increase the generalization ability of the model, data augmentation is used, and the leaky ELU activation function is used for the experiments to get the best model; compared with discrete emotion, dimensional emotion data can not only reflect the spatial information of the data but also describe the temporal information of the data. To this end, this paper implements the prediction of dimensional sentiment data based on recurrent neural networks and other related networks, and finally uses a quantitative algorithm of sentiment eigenvalues to calculate the sentiment intensity, which provides the database for the system to perform the adaptive adjustment.

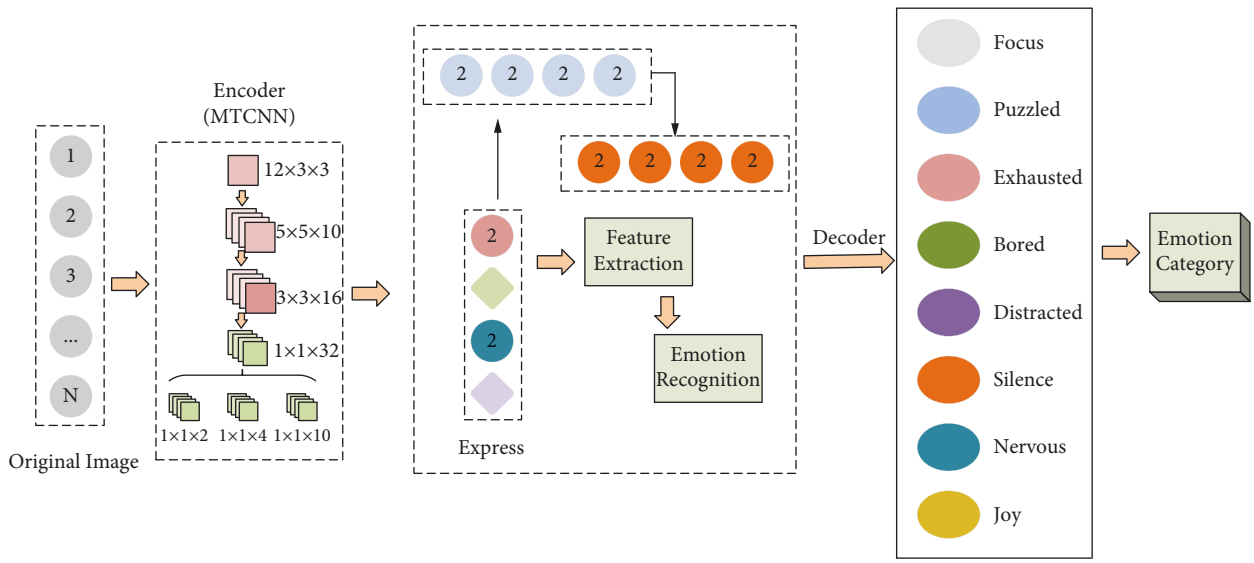


Figure 2: Classroom student emotion recognition algorithm.

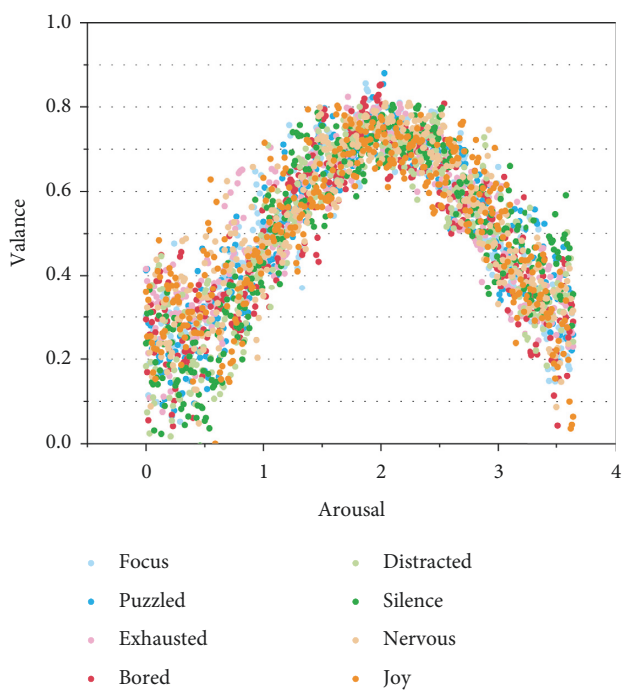


Figure 3: Spatial distribution of the 8 emotions in the arousal-valence dimension.

5. Classroom Emotion Recognition Model Performance Test Results

In this chapter, the global features extracted by IS10 and IS13 are screened using FC for better feature fusion, and the experimental results show that the combined features (IS10 + IS13) of global features extracted by IS10 and IS13 work best compared with the global features extracted by each of IS10 and IS13 [21]. The temporal features were also extracted with IS10 and IS13 and screened using 1D-CNN, and the experimental results showed that the combined features (IS10IId + IS13lld) with temporal features extracted by IS10 and IS13 worked better compared with MFCC. Therefore, the last dense layer output of (IS10 + IS13) in FC is used as the global optimal feature, and the last dense layer output of (IS10Ild + IS131ld) in 1D-CNN is used as the temporal optimal feature. To reduce the correlation between global and temporal features, the CoreNet network is used to fuse the filtered global optimal features and temporal optimal features to obtain fused features, which are trained on FC, and the results show that the results of fused features based on CoreNet are optimal, thus proving the feasibility of the method. One of the applications of artificial intelligence in education is to monitor valuable information such as students' learning expressions and learning postures in real time, pay attention to each student's emotional state, and make corresponding adjustments for feedback, and teachers adjust the rhythm of the classroom according to the specific situation. Classroom efficiency will be significantly improved.

In this study, a single ID photo of 100 people is selected as the original sample and a small sample dataset is constructed. Because the ID photos have the same background, it can make the faces in the images at the same angle, which can eliminate the influence of different image backgrounds and avoid the influence of face skin color, etc. Based on this, data enhancement techniques such as mirror transformation, multi-region cropping, Gaussian noise, symmetric expansion, and bit-plane method are carried out to expand the sample database [22]. To achieve expression recognition, we use adversarial generative networks to expand the samples, extract expressions from expression feature samples, and use convolutional neural network models to "attach" them to a single face image and train them to generate face images with different expressions. The loss test is shown in Figure 4, which shows that the loss value decreases and smooths out as the number of iterations increases.

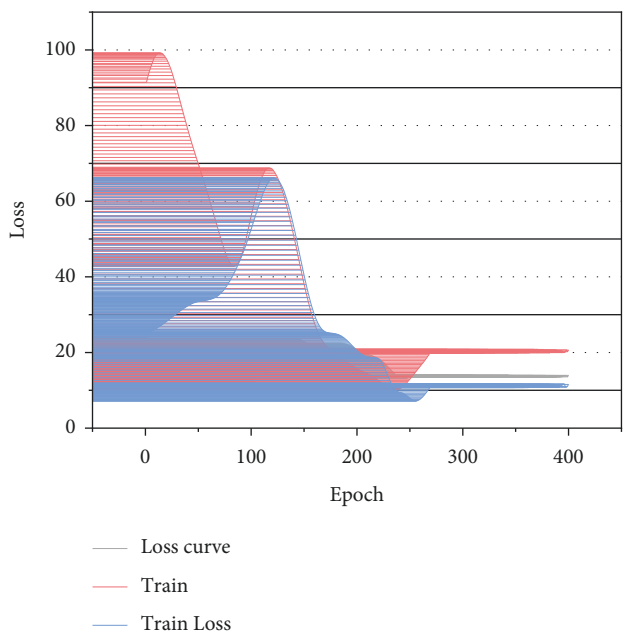


Figure 4: Change of loss value during training.

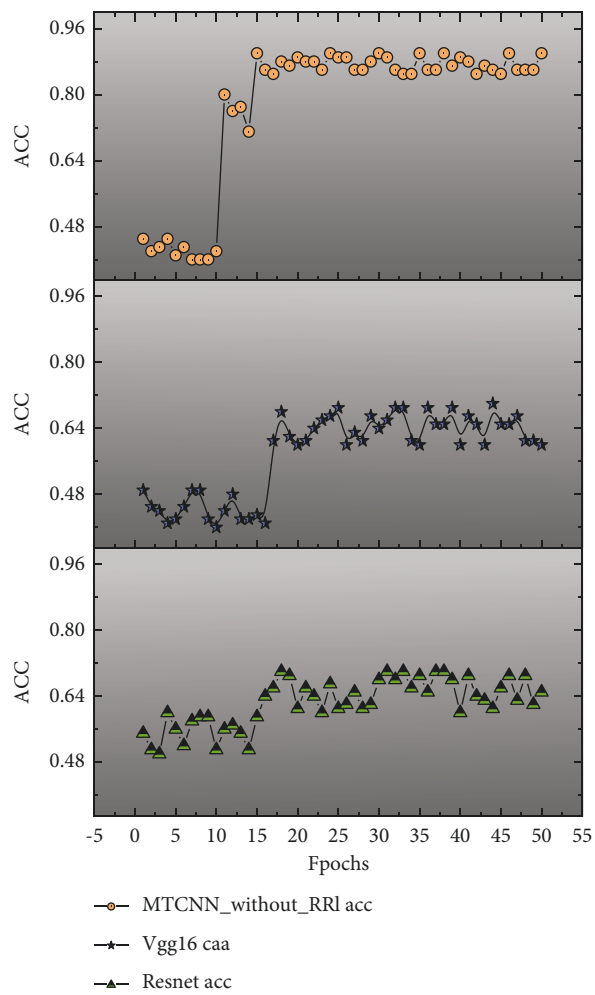


Figure 5: MTCNN accuracy comparison chart.

To further verify the effect of the representative region loss function designed in this paper on the model recognition, the MTCNN without RRL was tested in RAF-DB with added feature images. The accuracy of the MTCNN without RRL on the RAF-DB test set is 76.5%, which compares with that of the pretrained model in Figure 5. The blue curve represents the test accuracy of the MTCNN model without RRL, which nearly overlaps with the accuracy curves of ResNet18 and VGG16 in the fitting part, indicating that the MTCNN model without RRL has the same recognition effect as the two pretrained models without RRL. The difference between the recognition results of the MTCNN model without RRL and the two pretrained models is small.

The MTCNN model with RRL has a faster fitting speed with the same training parameters and achieves the fit in the 15th round, while the model without RRL has oscillations in the test set, a slow fitting speed, and relatively low accuracy [23]. Therefore, RRL can effectively improve the performance of the MTCNN model and has an improvement effect on the recognition of the MTCNN model.

The proposed method in this paper was evaluated on three datasets Twitter I, Twitter II, and Emotion ROI to demonstrate its effectiveness. The three datasets were divided into 80% training set and 20% test set using random partitioning. The proposed method achieves 79.83% and 78.25% classification accuracy on Twitter I and Twitter II datasets, respectively, which are higher than the traditional visual sentiment analysis method GCH and Sent bank based on intermediate semantic representation [7]. The deep learning-based visual sentiment analysis methods outperform the traditional visual sentiment analysis methods in terms of performance. The classification results of the proposed methods in this paper are improved compared with the comparison methods on both datasets, and the COIS model with the best comparison results has improved the classification accuracy by 0.93% and 1.42% on the two datasets, respectively. The classification results of the test method and the five comparison methods on the meta classifier sentiment image dataset Emotion ROI are evaluated in terms of accuracy %, and the accuracy comparison is shown in Figure 6.

The proposed method achieves 49.34% classification accuracy on the multi-category sentiment image dataset Emotion ROI, which is more accurate than the traditional visual sentiment analysis method GCH and the intermediate semantic representation-based visual sentiment analysis method Sent bank. The classification accuracy of DA-MLCNN is 6.81% and 3.88% higher than that of Deep-SentiBank and VGGNet-16, and 1.78% and 1.21% higher than that of PCNN and COIS models, respectively. By comparing the classification results of the various methods on the multiclassification dataset, it can be shown that the DA-MLCNN method proposed in this paper can also be adapted to the multiclassification task of visual emotion [6]. The combined classification performance on both dichotomous and meta classifier sentiment image datasets shows that the proposed method can learn more discriminative

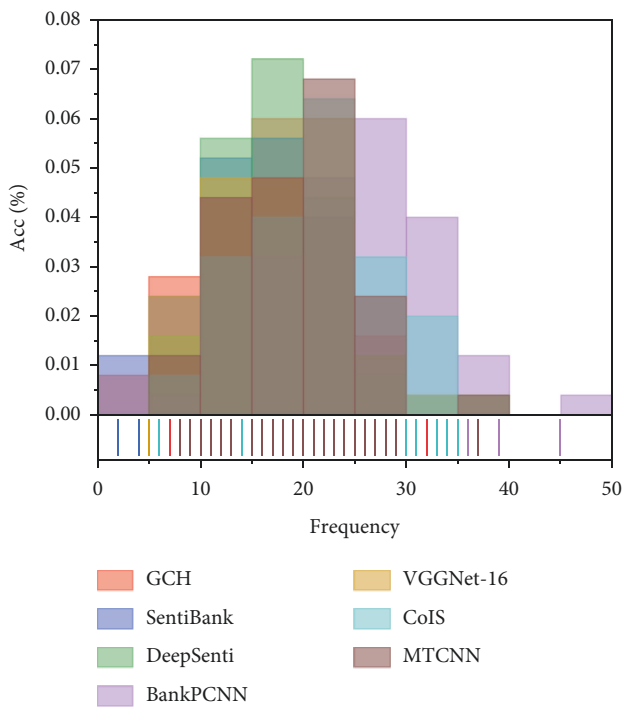


Figure 6: Classification accuracy of different methods on the Emotion ROI dataset.

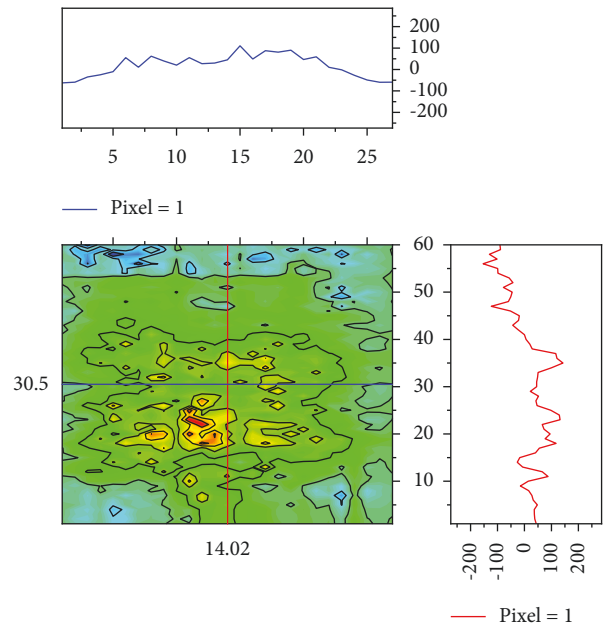


Figure 7: Joint probability distribution of parameters.

visual features and thus improve the visual sentiment analysis.

Through scale normalization, the face in the original image is scaled to a unified standard to reduce or even eliminate noise interference, which makes the evaluation accuracy of the subsequent expression recognition algorithm more reliable. First, a class is sampled and analyzed as Classroom A. Classroom A is analyzed by running the program. The X-axis represents the values of the parameters, and the corresponding Y-axis represents the corresponding posterior probabilities; if the posterior probabilities are larger, then the probability that the parameters μ andσ are the true values will be larger [3]. In such a plot, it is intuitive to obtain more reasonable values from the posterior. It can also be seen that the curves of the parameters in the left plot are smoother, and the right plot looks like white noise, which means that there is a good degree of mixing. And the maximum posterior estimate of each variable, which is the peak in the left-hand distribution, is very close to the true parameters. As shown in Figure 7, the parameters μ andσ do not have any correlation, meaning that the two parameters are independent of each other and will not be linearly correlated. Finally, a maximum posterior density (HPD) interval is run. An HPD interval is the smallest interval that contains a certain proportion of probability density and is often used to describe the dispersion of the posterior distribution of the parameters.

When making statistical inferences, the size of the sample that can be relied upon is often limited by various conditions. Direct use of means is intuitive and simple, but the meaningfulness of statistical inference cannot be determined by the size and complexity of the computation alone. The mean value calculated directly is easily affected by the sample characteristics, the way the sample is drawn, and other factors, especially when the sample size is limited, making it more difficult to make reasonable inferential statistics. Using the Bayesian probability model for statistical analysis, the MCMC under the framework of the PyMC3 probability model estimates the posterior distribution of parameters by random sampling, and then the highest probability value in the posterior distribution is used as the estimator of the parameters, the more the samples, the more the posterior distribution converges, and the resulting estimates will be closer to the real situation.

6. Conclusion

In this paper, we perform face detection and image segmentation with the help of a multi-task convolutional neural network (MTCNN), select the ones with better feature morphology to build a standard database, and further improve and build a classroom-based facial expression classification standard. A visual motion analysis method with the fusion of overall and local features of images is proposed. The method designs a multi-scale full convolutional neural network for detecting the salient regions of images and extracting the features of the salient regions of images, while generating class activation mappings of emotional images using only the image-level emotional labels, and finally generating emotional activation maps of images and extracting the features of emotional regions by superimposing class activation mappings of multiple emotional categories. Suitable parameters are designed according to the characteristics of the experimental data, and the constructed dataset is trained. After testing and analysis, the results show that the model of constructing a task convolutional neural network performs well on the student classroom facial

expression dataset, and the recognition rate using the test set is at 91%, which indicates that the method is scientifically feasible and can reduce the reliance on the training sample size, reduce the collection workload, and alleviate the pressure of data storage. However, due to the small sample size of the classroom emotion dataset constructed in this paper, the emotion labels are representative but incomplete and cannot fully characterize the recognition effect of the model designed in this paper. For the construction of the emotion dataset in a classroom environment, future work, other emotion categories, or even compound emotion categories can be added to further expand the sample size of the dataset. Increase the number of labelers to further enhance the accuracy of the dataset picture labels.

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 or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] S. Li and W. Deng, "Reliable crowdsourcing and deep locality-preserving learning for unconstrained facial expression recognition," *IEEE Transactions on Image Processing*, vol. 28, no. 1, pp. 356–370, 2018.

[2] E. Avots, T. Sapiński, and M. Bachmann, "Audiovisual emotion recognition in wild[J]," *Machine Vision and Applications*, vol. 30, no. 5, pp. 975–985, 2019.

[3] P. Song and W. Zheng, "Feature selection based transfer subspace learning for speech emotion recognition," *IEEE Transactions on Affective Computing*, vol. 11, no. 3, pp. 373–382, 2020.

[4] A. Mehrabian and J. A. Russell, *An Approach to Environmental Psychology*, p. 266, MIT Press, Cambridge, MA, 1980.

[5] A. K. Hassan and S. N. Mohammed, "A novel facial emotion recognition scheme based on graph mining," *Defence Technology*, vol. 16, no. 5, pp. 1062–1072, 2020.

[6] A. Khattak, M. Z. Asghar, M. Ali, and U. Batool, "An efficient deep learning technique for facial emotion recognition," *Multimedia Tools and Applications*, vol. 81, no. 2, pp. 1649–1683, 2022.

[7] M. S. Hossain and G. Muhammad, "An audio-visual emotion recognition system using deep learning fusion for a cognitive wireless framework," *IEEE Wireless Communications*, vol. 26, no. 3, pp. 62–68, 2019.

[8] F. Noroozi, M. Marjanovic, A. Njegus, S. Escalera, and G Anbarjafari, "Audio-visual emotion recognition in video clips," *IEEE Transactions on Affective Computing*, vol. 10, no. 1, pp. 60–75, 2019.

[9] R. Szczepanowski, J. Traczyk, M. Wierzchoń, and A Cleeremans, "The perception of visual emotion: comparing different measures of awareness," *Consciousness and Cognition*, vol. 22, no. 1, pp. 212–220, 2013.

[10] X. Yin and X. Liu, "Multi-task convolutional neural network for pose-invariant face recognition," *IEEE Transactions on Image Processing*, vol. 27, no. 2, pp. 964–975, 2018.

[11] H. H. Kim, J. K. Park, J. H. Oh, and D. J Kang, "Multi-task convolutional neural network system for license plate recognition," *International Journal of Control, Automation and Systems*, vol. 15, no. 6, pp. 2942–2949, 2017.

[12] X. Li, L. Zhao, L. Wei et al., "DeepSaliency: multi-task deep neural network model for salient object detection," *IEEE Transactions on Image Processing*, vol. 25, no. 8, pp. 3919–3930, 2016.

[13] H. Luo, Y. Yang, B. Tong, F. Wu, and B Fan, "Traffic sign recognition using a multi-task convolutional neural network," *IEEE Transactions on Intelligent Transportation Systems*, vol. 19, no. 4, pp. 1100–1111, 2018.

[14] K. S. K. Rani, D. Ramya, and J. Manikandan, "Monitoring emotions in the classroom using machine learning[J]," *Int J Sci Technol Res*, vol. 9, no. 01, pp. 3723–3726, 2020.

[15] H. Zeng, X. Shu, Y. Wang et al., "*EmotionCues*: emotion-oriented visual summarization of classroom videos," *IEEE Transactions on Visualization and Computer Graphics*, vol. 27, no. 7, pp. 3168–3181, 2021.

[16] Y. Cui, S. Wang, and R. Zhao, "Machine learning-based student emotion recognition for business English class," *International Journal Of Emerging Technologies In Learning (IJET)*, vol. 16, no. 12, pp. 94–107, 2021.

[17] M. N. Divya, "Smart teaching using human facial emotion recognition (fer) model[J]," *Turkish Journal of Computer and Mathematics Education (TURCOMAT)*, vol. 12, no. 11, pp. 6925–6932, 2021.

[18] Y. Wang, M. Zhang, A. Lin et al., "Mining structure–property relationships in polymer nanocomposites using data driven finite element analysis and multi-task convolutional neural networks," *Molecular Systems Design & Engineering*, vol. 5, no. 5, pp. 962–975, 2020.

[19] K. Wu, Z. Zhao, R. Wang, and G. W Wei, "Top*P-S*: p," *Journal of Computational Chemistry*, vol. 39, no. 20, pp. 1444–1454, 2018.

[20] Y. Cheng, S. Fu, M. Tang, and D Liu, "Multi-task deep neural network (MT-DNN) enabled optical performance monitoring from directly detected PDM-QAM signals," *Optics Express*, vol. 27, no. 13, pp. 19062–19074, 2019.

[21] Y. Cheng, W. Zhang, S. Fu, M. Tang, and D Liu, "Transfer learning simplified multi-task deep neural network for PDM-64QAM optical performance monitoring," *Optics Express*, vol. 28, no. 5, pp. 7607–7617, 2020.

[22] S. Zhai, S. Liu, X. Wang, and J Tang, "FMT: fusing multi-task convolutional neural network for person search," *Multimedia Tools and Applications*, vol. 78, no. 22, pp. 31605–31616, 2019.

[23] O. Gervasi, V. Franzoni, M. Riganelli, and S Tasso, "Automating facial emotion recognition," *Web Intelligence*, vol. 17, no. 1, pp. 17–27, 2019.