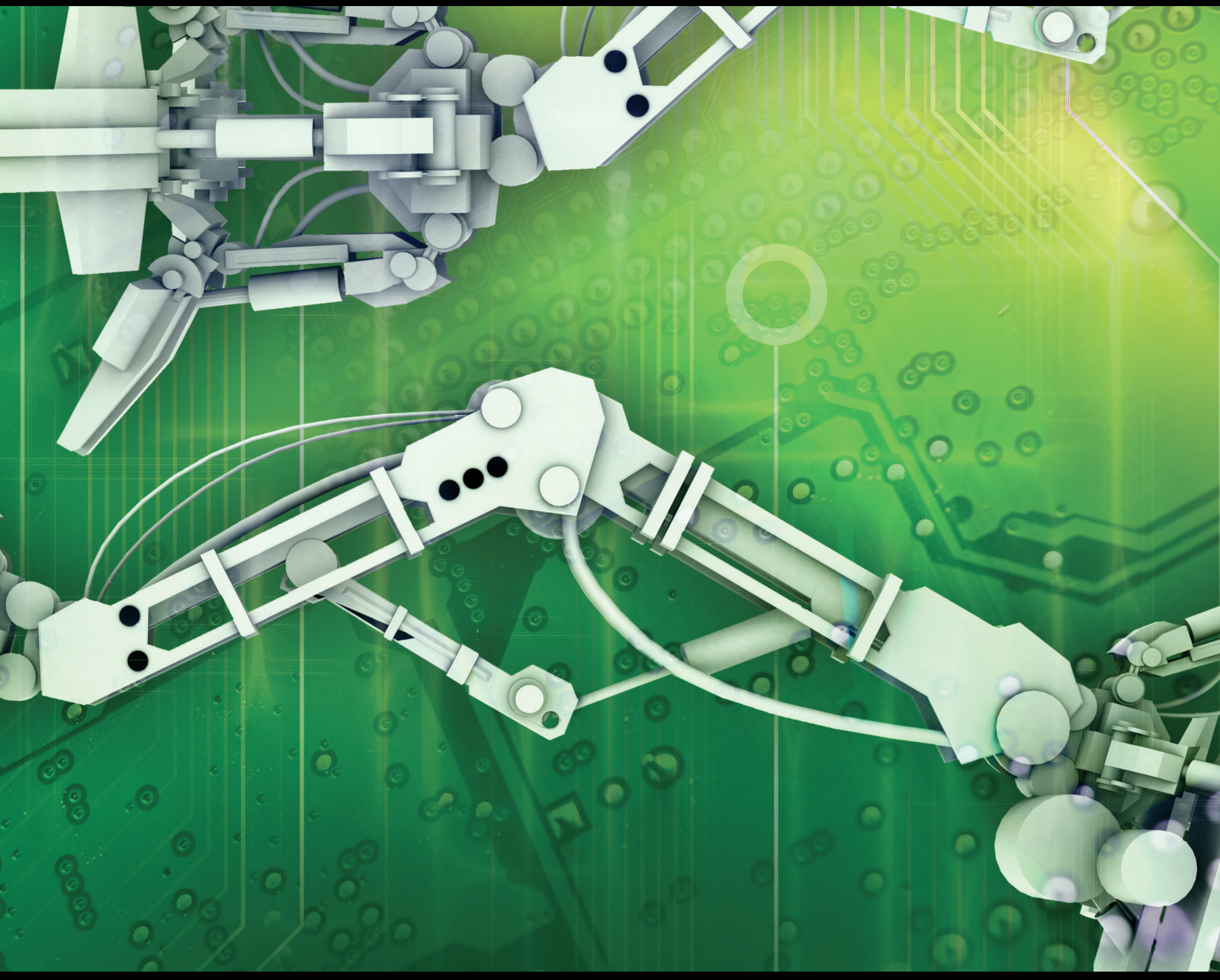
Journal of Robotics

Internet of Robotic Things-Enabled Edge Intelligence Cognition for Humanoid Robots

Lead Guest Editor: Kaijian Xia
Guest Editors: Joon Huang Chuah and Shan Zhong



Internet of Robotic Things-Enabled Edge Intelligence Cognition for Humanoid Robots

Journal of Robotics

Internet of Robotic Things-Enabled Edge Intelligence Cognition for Humanoid Robots

Lead Guest Editor: Kaijian Xia
Guest Editors: Joon Huang Chuah and Shan Zhong

Contents

Contents

Hindawi
Journal of Robotics
Volume 2023, Article ID 9785409, 1 page
https://doi.org/10.1155/2023/9785409

Retraction

Retracted: Research on Clue Mining in Criminal Cases of Smart Phone Trojan Horse under the Background of Information Security

Journal of Robotics

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This article has been retracted by Hindawi following an investigation undertaken by the publisher [1].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

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

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

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

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

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

References

[1] L. Gang and Y. Wen, "Research on Clue Mining in Criminal Cases of Smart Phone Trojan Horse under the Background of Information Security," *Journal of Robotics*, vol. 2022, Article ID 7568110, 11 pages, 2022.

Hindawi
Journal of Robotics
Volume 2023, Article ID 9860830, 1 page
https://doi.org/10.1155/2023/9860830

Retraction

Retracted: Research on the Detection Countermeasures of Telecommunication Network Fraud Based on Big Data for Killing Pigs and Plates

Journal of Robotics

Received 19 September 2023; Accepted 19 September 2023; Published 20 September 2023

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

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

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

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

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

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

References

[1] G. Li and Y. Wen, "Research on the Detection Countermeasures of Telecommunication Network Fraud Based on Big Data for Killing Pigs and Plates," *Journal of Robotics*, vol. 2022, Article ID 4761230, 11 pages, 2022.

Hindawi
Journal of Robotics
Volume 2022, Article ID 8951671, 6 pages
https://doi.org/10.1155/2022/8951671

Research Article

Emotional State Analysis Model of Humanoid Robot in Human-Computer Interaction Process

Boxin Peng

School of Computer Science and Technology, Southeast University, Dhaka, Bangladesh

Correspondence should be addressed to Boxin Peng; 213193474@seu.edu.cn

Received 22 December 2021; Revised 27 March 2022; Accepted 30 March 2022; Published 6 May 2022

Academic Editor: Shan Zhong

The traditional humanoid robot dialogue system is generally based on template construction, which can make a good response in the set dialogue domain but cannot generate a good response to the content outside the domain. The rules of the dialogue system rely on manual design and lack of emotion detection of the interactive objects. In view of the shortcomings of traditional methods, this study designed an emotion analysis model based on deep neural network to detect the emotion of interactive objects and built an open-domain dialogue system of humanoid robot. In affective state analysis language processing, language coding, feature analysis, and Word2vec research are carried out. Then, an emotional state analysis model is constructed to train the emotional state of a humanoid robot, and the training results are summarized.

1. Introduction

With the progress of science and technology, robots have gradually entered every aspect of people's lives. From industrial use to military applications, home service, education, and laboratories, robots are playing a significant role [1]. According to the three laws of robotics [2], the ultimate goal of robot development is to make robots imitate human intelligent behavior, to help humans better complete tasks, and to achieve goals [3]. In the process of human and robot cooperation to complete tasks, human beings cannot avoid the need to better communicate with the robot in order to complete tasks more efficiently [4, 5]. Traditional interaction between human and computer is human mainly through keyboard, mouse, and other manual input equipment to tell the computer information and computer through the display and other peripherals to feedback information to human. This interaction is inconvenient and requires a large number of peripherals. And in ordinary life, it is not guaranteed that everyone can use a computer [6]. Different from the traditional interaction between human and computer, the interaction between human and machine is carried out through some well-known natural channels, such as speech, vision, touch, hearing, proximity, and other human interactions to complete tasks [7]. This kind of interaction is familiar to most people and is more concise and efficient. So it can help people and robots interact more effectively to complete tasks [8]. The emotion analysis model of humanoid robot can analyze and identify the emotional information of the interactive object in the process of interaction, which is an important part of the dialogue system [9]. In the process of interaction, the language of the object of interaction contains rich emotional information and the text content information is a high-level expression of human thinking.

2. Literature Review

The main implementation methods of traditional text sentiment analysis are generally divided into two kinds: sentiment dictionary and machine learning algorithm. Text emotion analysis is usually based on emotion dictionary. At present, relatively well-known emotion dictionaries include HowNET, Chinese Polar Emotion Word (NTUSD) from Taiwan University, and English emotion dictionary Word Net from Preston University [10]. The emotion analysis process based on the emotion dictionary is shown in Figure 1.

Machine learning is used for sentiment analysis, which is regarded as a text classification task. The commonly used

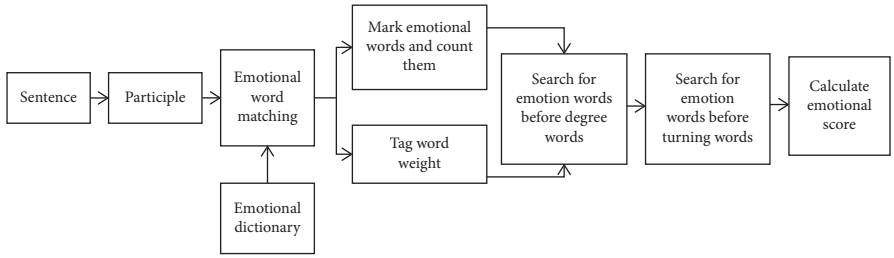


Figure 1: The emotion analysis process based on the emotion dictionary.

methods include Naive Bayes, SVM, and CRF. Li [11] compared Naive Bayes, maximum entropy model, and support vector machine algorithm in the emotion classification task of film reviews and found that SVM achieved the best classification effect. Huang [12] (2021) used the multistrategy method of SVM hierarchical structure to classify the emotional polarity of Chinese microblogs. The experiment shows that the SVM-based multistrategy method has the best effect, and the introduction of theme-related features can improve the accuracy to some extent. Lu [13] (2018) experimented with SVM, Bayes, and other classification algorithms and information gain and other feature selection algorithms in Chinese microblog sentiment analysis and took TF-IDF as the feature weight. Experimental results show that TF-IDF can be used as a feature weight, SVM can be used as a classification algorithm, and information gain can be used as a feature selection algorithm to achieve the best classification effect.

With the development of deep learning, deep learning models have also been applied to text classification. Law [14] (2013) proposed the reinforcement learning framework DISA based on CNN and LSTM by taking Chinese audio information and pinyin as emotion analysis features and achieved good results. Cela [15] (2013) applied the D-S evidence theory to integrate emotional information from visual, sound, and other aspects and analyzed the transfer rule of emotional state caused by the simultaneous action of the two factors. Finally, the emotion model is applied to the emotion robot system, so that the robot can generate emotions according to the external stimulus and make the corresponding expression. The experimental results show that the affective model is effective. Tidoni [16] (2014) combined the idea of a recurrent neural network with that of a convolutional neural network to improve the limitations of CNN in long-distance context capture and proposed RCNN for text classification. BaTula [17] (2017) proposed a game-based cognitive and emotional interaction model for robots based on PAD (pleasure-arousal-dominance) emotion space, aiming at the problem of lack of emotion and low participation of members in existing open-domain human-computer interaction systems. Experimental results show that compared with other cognitive interaction models, the proposed model can reduce the dependence of robot on external emotional stimuli and guide members to participate in human-computer interaction effectively.

There are various forms of information transmission. Due to the limitation of technology, it is impossible for robots to completely obtain the information of interactive objects, so intention prediction becomes important and essential [18]. Human interaction usually requires continuous prediction of intention. For example, in a conversation, people are constantly trying to predict the direction of a future conversation or the reactions of others through intention prediction [19]. Therefore, in order to make human-robot interaction more like human-human interaction, intention prediction in human-computer interaction is essential. According to the classification of human-computer interaction, intention prediction also has different processing situations and forms [20]. In cooperative human-computer interaction, intention prediction is mainly caused by incomplete interactive information. Humans and robots should cooperate to complete tasks. With the completeness of information, it is necessary to add the prediction of human intentions to complete tasks better and more efficiently.

Due to the linguistic phenomenon of polysemy and irony in Chinese, the method based on emotion dictionary cannot achieve high accuracy and is not suitable for cross-field research. With the geometric increase in information content, the establishment of a data-driven machine learning model for emotion analysis of irregular documents has a good application prospect.

3. Emotion Recognition Process and Data Acquisition Preprocessing

3.1. Emotion Recognition Process. In the process of interaction, to obtain text information, the voice content of the interactive object needs to be recorded through the microphone and converted into an audio and then the voice recognition module obtains the text information through voice recognition. The text information is preprocessed and fed into the emotion analysis model, from which the emotional state of the interactive object is output. As for the construction of emotion analysis models, this paper adopts the idea of machine learning to build a data-driven emotion analysis model [21]. The algorithm is selected to conduct offline training through data sets, and the model after training is reserved. The saved model is loaded and used for prediction. The text emotion analysis process based on machine learning is shown in Figure 2.

3.2. Data Acquisition and Preprocessing. The data includes data acquisition and data preprocessing. The specific contents are as follows.

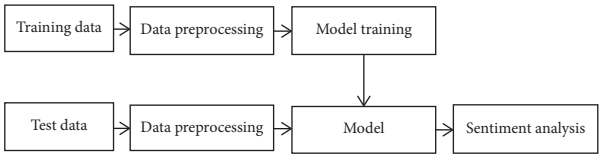


Figure 2: Text emotion analysis process based on machine learning.

3.2.1. Data Acquisition. In building the emotion analysis model of humanoid robot, we used the data set from the "Microblog Cross-Language Emotion Recognition Dataset" published by the International Conference on Natural Language Processing and Chinese Computing (NLPCC) in 2019 and 2020. The corpus is divided into positive and negative categories. Among them, there are 12,153 positive corpus categories and 12,178 negative corpus categories. The corpus content is from microblogs, and the sentences are colloquial, which is suitable for training the emotion analysis model.

3.2.2. Data Preprocessing. In data preprocessing, in order to ensure the uniformity of positive and negative categories in the corpus, 25 items were deleted from the negative label corpus through downsampling and the positive and negative samples were unified into 12,153 items. Since most of the corpus is taken from Weibo, it contains many emojis and repetitive punctuation marks. In addition, in practical application, the sentence after speech recognition will not have multiple repeated punctuation marks [22]. Based on the above points, redundant punctuation marks and emoticons were deleted in the preprocessing, which were not regarded as features. During word segmentation, the Jieba word segmentation tool kit was used in this study. A comparison table of punctuation and facial expressions in textual processing is shown in Table 1.

In the text vector space model, the commonly used feature selection methods include the chi-square test, information gain, mutual information method, and TF-IDF. TF-IDF combines word distribution information among documents in the word bag model and highlights key information by calculating absolute word frequency and inverted document frequency.

In the subsequent construction of emotion analysis models, we experimented with traditional machine learning models such as Bernoulli Bayes, Polynomial Bayes, and support vector machine (SVM) and neural network models such as Bi-LSTM, Bi-LSTM combined with attention mechanism, and Text-CNN.

4. Affective State Analysis Language Processing

4.1. Language Code. In natural language processing (NLP), sentence segmentation is generally carried out, with characters, words, and phrases as the minimum unit of estimation of Chinese language. Unique thermal coding is the simplest representation of this kind of feature. In unique coding, each different feature has its special state bit.

For example, "It's a good book!"; after word segmentation, the result reads "This is a good book with good content!."

Table 1: Comparison table of punctuation and facial expressions in textual processing.

Punctuation and facial expression	Textual processing
? ? ?	Doubt
! ! !	Amazing
...	End, helpless
😫😤😠	Plaint
😂😅😓	Awkward
😃😊😄😄	Happy
😡😓😱	Unhappy

["This," "book," "Content," "Good," "!"] There are altogether 5 independent features in the above examples, which can be represented by independent thermal coding according to the order of word occurrence. The independent thermal coding of some features can be expressed as follows:

"This": [1, 0, 0, 0, 0, 0, 0, 0, 0]
"book": [0, 1, 0, 0, 0, 0, 0, 0, 0]
"Content": [0, 0, 1, 0, 0, 0, 0, 0, 0]
"Good": [0, 0, 0, 1, 0, 0, 0, 0, 0]
"!": [0, 0, 0, 0, 0, 0, 0, 0, 1]

In unique thermal coding, each feature has a single dimension and the dimensions of unique thermal coding are the same as the number of different features. To a certain extent, the unique thermal coding plays a role in expanding the features, but when the database dictionary content is large, this representation method takes up a lot of space and the calculation dimension is large.

"This is a good book with good content!": [1, 2, 2, 1, 1, 1, 1, 1, 1].

The word bag model is a vector space model in which the number of words is represented in the corresponding position according to the word index to achieve the representation of the whole sentence.

4.2. Characteristics Analysis. In the text vector space model, the commonly used feature selection methods include the chi-square test, information gain, mutual information method, and TF-IDF. TF-IDF combines word distribution information between documents in the word bag model, and keywords are highlighted by calculating absolute term frequency (TF) and inverse document frequency (IDF).

Absolute word frequency (TF) represents the spelling of the feature item in the training text D_j. Important words in a text are often emphasized multiple times, and absolute word frequency can easily highlight these words.

Calculation of IDF of inverted document frequency is shown in the following formula:

$$\mathrm{IDF} = \log \frac{N}{n_i}. \tag{1}$$

N represents the total number of documents in the training set, and n_i represents the number of documents in which feature item appears in the training set. IDF highlights some words that appear less frequently but have strong

classification ability. In the actual calculation process, IDF will carry out smooth processing in order to avoid the missing of rare words in the corpus.

$$\text{TF} - \text{IDF} = \text{TF} \times \text{IDF} = tf_{ij} \times \log \frac{N+1}{n_i+1}. \tag{2}$$

4.3. Word2vec. Word2vec was proposed by Google in 2013. It is a way to represent words through dense feature representation, also known as distributed representation. There are two models for Word2vec training, namely, the Continuous Bag-of-Words (CBOW) model and Skip-Gram model. The improved methods of Word2vec are divided into two types, one is based on Hierarchical Softmax and the other is based on Negative Sampling, both of which are for simplifying computation and accelerating training. In Hierarchical Softmax, the output of projection layer under the CBOW model is the mean value of input word vector sum and the output of projection layer under the Skip-Gram model is the same as the input. In order to avoid calculating the probability of all words, the Hierarchical Softmax approach uses Huffman tree instead of Softmax mapping of the projection layer to output layer. Negative Sampling Word2vec is widely used in various natural language models, and word vector is also a pretraining method that can bring a neural network to a better training starting point and make the network easier to optimize [23].

Compared with the independent thermal coding, the dense feature is easy to calculate and does not have the problem of dimension explosion, which has a strong generalization ability. Dense feature representation can provide similarity information between features. This distributed representation of the word vector is widely used in natural language processing, such as Chinese word segmentation, sentiment analysis, and reading comprehension.

5. Construction of the Affective State Analysis Model

5.1. Introduction to the Model. Support vector machine (SVM) is a binary model algorithm that can also be used for text classification. The basic idea of a support vector machine is to find the hyperplane with the largest interval in the feature space. Its advantages are that it is effective in high-dimensional space and still has good effect when the number of dimensions is larger than the number of samples. Different kernel functions can be specified during the design of support vector machines. However, when the number of features is much larger than the number of samples, the performance of SVM is poor.

For the training data set T, T is $\{(x_1, y_1), (x_2, y_2), \ldots, (x_N, y_N)\}$ and $i = 1, 2, \ldots, N$, $y_i \in \{-1, +1\}$ stands for negative and positive labels, x_i is a sample or sentence, and N is the number of samples. Optimization problems solved by support vector machines are shown in the following formula:

$$\begin{aligned} &\min_{w,b,\zeta} \frac{1}{2} w^T w + C \sum_{i=1}^{n} \zeta i \\ &\text{s.t.}\, y_i\left(w^T \phi(x_i) + b\right) \geq 1 - \sum_{i=1}^{n} \zeta i, \quad \zeta i \geq 0,\ i = 1, 2, \ldots, N. \end{aligned} \tag{3}$$

In formula (3), w is the normal vector separating the hyperplane, ζi is the relaxation variable, and $\phi(x_i)$ is the mapping function. The dual form of the problem can be expressed as

$$\begin{aligned} &\min_{\alpha} \frac{1}{2} \alpha^T Q \alpha - e^T \alpha \\ &\text{s.t.}\, y^T \alpha = 0, \quad 0 \leq \alpha_i \leq C,\ i = 1, 2, \ldots, N. \end{aligned} \tag{4}$$

In formula (4), e stands for all vectors 1, C is the upper bound of Lagrange everyday number, and Q is a semidefinite matrix of shape (N, N).

$$Q_{ij} = y_i y_j K\left(x_i, x_j\right). \tag{5}$$

In formulas (3–5), $K(x_i, x_j) = \phi(x_i)^T \phi(x_j)$ is the kernel. The decision function Gram of the support vector machine is expressed as $\text{sign}\left(\sum_i^N y_i \alpha_i K(x_i, x) + b\right)$.

5.2. Training Process. In order to make the support vector machine output category probability, Platt Scaling was used in this paper. This is a parameterized method that uses the logistic regression model to fit the output values; that is, the Sigmoid function is used to map the values to between $[0, 1]$ and finally the output values of the original model are mapped to probability values, as shown in the following formula:

$$P(y = 1|x) \approx P_{A,B}(f) \equiv \frac{1}{1 + \exp(Af + B)}. \tag{6}$$

In formula (3), $f(x)$ is the decision function of the support vector machine, which can output corresponding labels to any input X, and parameter A and parameter B are trainable parameters.

The objective function of training is the loss of cross entropy, as shown in the following formulas:

$$\min_{z=(A,B)} F(z) = -\sum_{i}^{l} \left(t_i \log(P) + (1 - t_i) \log(1 - P_i)\right), \tag{7}$$

$$t_i = \begin{cases} \dfrac{N_+ + 1}{N_+ + 2} & \text{if } y_i = +1 \\ \dfrac{1}{N_- + 2} & \text{if } y_i = -1 \end{cases} \qquad i = 1, 2, \ldots, l. \tag{8}$$

By Platt Scaling, support vector machines can output probability values of categories. The basic idea is that the closer the points are to the interface, the less likely they are to

match and the farther the points are from the interface, the more likely they are to match.

5.3. Training Results. Support vector machines generally adopt a linear kernel in text classification. The test results of the SVM model with a linear kernel as the kernel function and a word bag model as input are as follows: F1 value is 0.763, the accuracy is 76.81%, and the AUC value is 0.821. The test results of the SVM model with a linear kernel as the kernel function and TF-IDF as the input are as follows: F1 value is 0.795, the accuracy is 78.94%, and the AUC value is 0.863.

Compared with the word bag model, SVM with TF-IDF as feature input achieves better results under the current data set.

6. Conclusion

The main steps of building an emotion analysis model include emotion recognition process and data acquisition and preprocessing, emotional state analysis language processing, emotional state analysis model construction, and integration [24]. The effects of TF-IDF characteristics and the word bag model on the training results are analyzed and explored in the experiment. The results show that in a single model, the TF-IDF feature combined with the support vector machine achieves the optimal result. The stacking strategy and soft voting strategy were compared in the model integration, and the optimal performance was obtained by stacking with a support vector machine learner. This is the main innovation of this paper. The introduction of the methods used in each component is too superficial, which is the main shortcoming of this paper. In this paper, the design principle of a sentiment analysis model based on the support vector machine is studied. Based on the experimental data, the influence of attention mechanism on the neural network model is explored and the performance of traditional machine learning model is compared with different inputs. In the single model experiment, the support vector machine combined with TF-IDF achieved the best classification effect, with a F1 value of 0.795, an accuracy rate of 78.94%, and an AUC value of 0.863.

Data Availability

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

Conflicts of Interest

The author declares no conflicts of interest.

References

[1] L. Gabriella, G. Márta, K. Veronika et al., "Emotion attribution to a non-humanoid robot in different social situations," *PLoS One*, vol. 9, no. 12, Article ID e114207, 2014.

[2] A. Rozanska and M. Podpora, "Multimodal sentiment analysis applied to interaction between patients and a humanoid robot pepper," *IFAC-Papers Online*, vol. 52, no. 27, pp. 411–414, 2019.

[3] M. Viríkova and S. Peter, "Teach your robot how you want it to express emotions," *Advances in Intelligent Systems and Computing*, vol. 316, pp. 81–92, 2015.

[4] X. Ke, Y. Shang, and K. Lu, "Based on hyper works humanoid robot facial expression simulation," *Manufacturing Automation*, vol. 137, no. 1, pp. 118–121, 2015.

[5] F. Azni Jafar, N. Abdullah, N. Blar, M. N. Muhammad, and A. M. Kassim, "Analysis of human emotion state in collaboration with robot," *Applied Mechanics and Materials*, vol. 465-466, pp. 682–687, 2013.

[6] Z. Shao, R. Chandramouli, K. P. Subbalakshmi, and C. T. Boyadjiev, "An analytical system for user emotion extraction, mental state modeling, and rating," *Expert Systems with Applications*, vol. 124, no. 7, pp. 82–96, 2019.

[7] J. Hernandez-Vicen, S. Martinez, J. Garcia-Haro, and C. Balaguer, "Correction of visual perception based on neuro-fuzzy learning for the humanoid robot TEO," *Sensors*, vol. 18, no. 4, pp. 972-973, 2018.

[8] A. Zaraki, D. Mazzei, M. Giuliani, and D. De Rossi, "Designing and evaluating a social gaze-control system for a humanoid robot," *IEEE Transactions on Human-Machine Systems*, vol. 44, no. 2, pp. 157–168, 2014.

[9] J. Wainer, B. Robins, F. Amirabdollahian, and K. Dautenhahn, "Using the humanoid robot KASPAR to autonomously play triadic games and facilitate collaborative play among children with autism," *IEEE Transactions on Autonomous Mental Development*, vol. 6, no. 3, pp. 183–199, 2014.

[10] L. Tang, Z. Li, X. Yuan, W. Li, and A. Liu, "Analysis of operation behavior of inspection robot in human-machine interaction," *Modern Manufacturing Engineering*, vol. 3, no. 3, pp. 7-8, 2021.

[11] Z. Li and H. Wang, "Design and implementation of mobile robot remote human-computer interaction software platform," *Computer Measurement & Control*, vol. 25, no. 4, pp. 5-6, 2017.

[12] H. Huang, N. Liu, M. Hu, Y. Tao, and L. Kou, "Robot cognitive and affective interaction model based on game," *Journal of Electronics and Information Technology*, vol. 43, no. 6, pp. 8-9, 2021.

[13] Lufei, Y. Jiang, and G. Tian, "Autonomous cognition and personalized selection of robot service based on emotion-spatiotemporal information," *Robot*, vol. 40, no. 4, pp. 9-10, 2018.

[14] J. Law, P. Shaw, and M. Lee, "A biologically constrained architecture for developmental learning of eye-head gaze control on a humanoid robot," *Autonomous Robots*, vol. 35, no. 1, pp. 77–92, 2013.

[15] A. Cela, J. Yebes, R. Arroyo, L. R. Bergasa, and E. López, "Complete low-cost implementation of a teleoperated control system for a humanoid robot," *Sensors*, vol. 13, no. 2, pp. 1385–1401, 2013.

[16] E. Tidoni, P. Gergondet, A. Kheddar, and S. M. Aglioto, "Audio-visual feedback improves the BCI performance in the navigational control of a humanoid robot," *Frontiers in Neurorobotics*, vol. 8, 2014.

[17] A. M. BaTula, Y. E. Kim, and H. Ayaz, "Virtual and actual humanoid robot control with four-class motor-imagery-based optical brain-computer interface," *BioMed Research International*, vol. 2017, Article ID 1463512, 13 pages, 2017.

[18] T. Sato, Y. Nishida, J. Ichikawa, and Y. Hatamura, "Active understanding of human intention by a robot through monitoring of human behavior," in *Proceedings of the IEEE/RSJ/GI International Conference on Intelligent Robots and*

Systems, IROS'94, pp. 405–414, IEEE, Munich, Germany, September 1994.

[19] H. Clint, *Modelling Intention Recognition for Intelligent Agent Systems*, DTIC, Mexico City, Mexico, 2004.

[20] K. A. Tahboub, "Intelligent human-machine interaction based on dynamic bayesian networks probabilistic intention recognition," *Journal of Intelligent and Robotic Systems*, vol. 45, no. 1, pp. 31–52, 2006.

[21] T. Koolen, S. Bertrand, G. Thomas et al., "Design of a momentum-based control framework and application to the humanoid robot atlas," *International Journal of Humanoid Robotics*, vol. 13, no. 1, 2016.

[22] W. Wu and H. Li, "Artificial emotion modeling and human-computer interaction experiment in PAD emotion space," *Journal of Harbin Institute of Technology*, vol. 51, no. 1, pp. 9-10, 2019.

[23] M. S. Erden, "Emotional postures for the humanoid-robot nao," *International Journal of Social Robotics*, vol. 5, no. 4, pp. 441–456, 2013.

[24] B. Browatzki, V. Tikhanoff, G. Metta, and H. H. C. Bulthoff, "Active in-hand object recognition on a humanoid robot," *IEEE Transactions on Robotics*, vol. 30, no. 5, pp. 1260–1269, 2014.

Hindawi
Journal of Robotics
Volume 2022, Article ID 1188617, 9 pages
https://doi.org/10.1155/2022/1188617

Research Article

A MultiModal Detection Method for UHV Substation Faults Based on Robot Inspection and Deep Learning

Rong Meng, Zhao-lei Wang, Zhi-long Zhao, Jian-peng Li, and Wei-ping Fu

State Grid Hebei Extra High Voltage Company, Shijiazhuang, Heibei, 050070, China

Correspondence should be addressed to Zhao-lei Wang; wzlabc@163.com

Received 8 February 2022; Revised 22 March 2022; Accepted 30 March 2022; Published 23 April 2022

Academic Editor: Shan Zhong

Aiming at the problem of multi-modal fault detection of different equipment in ultrahigh voltage (UHV) substations, a method for based on robot inspection and deep learning is proposed. First, the inspection robot is used to collect the image data of different devices in the station and the source data is preprocessed by standard image augmentation and image aliasing augmentation. Then, the HSV color space model based on saliency area detection is used to extract equipment defect areas, which improves the accuracy of defect image classification. Finally, the traditional YOLOv3 network is improved by combining the residual network and the K-means clustering algorithm, and the detailed flow of the corresponding detection method is proposed. The proposed detection method and the other three methods were compared and analyzed under the same conditions through simulation experiments. The results show that the detection accuracy and recall rate of the method proposed in this study are the largest, which are 95.9% and 91.3%, respectively. The average detection accuracy under multiple intersection ratio thresholds is also the highest, and the performance is better than the other three comparison algorithms.

1. Introduction

At present, with the increasing number of UHV substations, their importance in the power system is gradually getting promoted [1]. The operation state of the UHV substation has a great impact on the power system. The effective safety detection of the UHV substation is an effective means to ensure its safe and stable operation, which is of great significance [2–4]. The safety detection of the UHV substation is mainly to check whether the equipment fails. At present, the main detection means include fixed monitoring, manual inspection, and robot inspection [5, 6]. Using robots to perform inspection tasks instead of manual workers has broad application scenarios and huge development space, and it is a research hotspot for intelligent development of UHV substations at home and abroad [7, 8].

2. Related Work

In view of the mentioned problems related to fault detection of the UHV substations, the most prospective and research-based detection method is robot inspection that can automatically obtain infrared images of power equipment by installing inspection robots in UHV substations and judge the status of power equipment on this basis [9]. Finally, the online analysis and decision on the operation status of the UHV substation are realized. Aiming at the fault detection problem of GIS (GAS Insulated SWITCHGEAR) combined appliances, [10] designs a GIS equipment maintenance robot with a flexible structure and convenient controls. The whole GIS equipment can be checked without blind spots. However, this method is only used to detect the internal equipment of GIS integrated appliances and has some limitations. With the goal of practical inspection robot, [11] developed a set of intelligent inspection technologies that can realize automatic charging of robots and task planning. It improves the inspection efficiency of UHV substations based on inspection robots to a certain extent. However, this method has some disadvantages, such as low accuracy and a long time. Aiming at the problem that substation indoor inspection robot could not accurately identify the inspection range under some complicated circumstances, [12] proposes

a flexible cable-driven inspection robot system scheme based on the design of hardware equipment and system software of the robot. Ref. [13] designs a trackless robot with a robotic arm that can navigate autonomously and inspect equipment for the narrow and complex indoor environment of the UHV substations. Autonomous movement and detection of the robot based on simultaneous positioning and mapping are realized. The above two methods only study the indoor inspection technology based on the inspection robot but do not analyze the outdoor situation, so their application scope is limited. Ref. [14] studies the key technologies of the UHV substation inspection robots in image acquisition, image recognition, intelligent inspection, and comprehensive diagnosis. By integrating the multisource data of the whole station and combining the corresponding sensors and acquisition equipment, it puts forward a kind of joint inspection system based on artificial intelligence analysis technology combined with automatic inspection. However, this method does not analyze the control complexity and motion precision of the robot itself. The goal of [15] is to improve the detection accuracy of the equipment and reduce the detection time in UHV substations. Aiming at the infrared image library collected by conventional UHV substation inspection robot equipped with infrared thermal imager, an infrared image detection method of UHV substation equipment based on improved Gaussian convolution kernel is proposed by combining deep learning. However, the image acquisition effect of the inspection robot in this method is poor in resolution, and the information is not intuitive and requires secondary processing. By designing the hardware structure and software system and integrating the existing technologies, [16] proposes a fully autonomous robotic system LongSword. Corresponding technical solutions are proposed for the process of equipment inspection using optical zoom cameras, thermal imaging cameras, and partial discharge detectors. However, the method does not consider the acquisition of sound information, and the detection accuracy needs to be improved.

Aiming at the core problems of small application range and low accuracy of current fault detection methods in UHV substations, this study proposes a multimodal fault detection method for UHV substations based on robot inspection and deep learning based on the above analysis. The basic idea is as follows: (1) The equipment image is obtained by inspecting the robot, and the image is augmented by standard image and image aliasing. (2) By introducing a HSV color space model to extract equipment defect areas, the accuracy of image classification can be improved. (3) The feature extraction capability and detection accuracy are improved by improving the traditional YOLOv3 network. Compared with traditional detection methods, the innovations of the proposed method are listed:

(1) Both standard image augmentation and image aliasing augmentation increase the amount of data and improve the generalization ability of the model.

(2) A defect region segmentation algorithm based on HSV is introduced to separate the defect region from the background, which improves the accuracy of the subsequent image classification algorithm.

(3) The traditional YOLOv3 network is improved by using the residual network and K-means clustering algorithm, which improves the detection accuracy.

3. Fault Multimodal Detection Method

3.1. The Architecture of Multi-Modal Detection of Faults in UHV Substations. Due to the complexity and diversity of the UHV substation inspection environment, the design of the robot inspection system must ensure real-time performance and robustness, as well as consider the needs of system optimization and equipment replacement. It is necessary to design extensible functional modules for easy maintenance and upgrade. The overall scheme of the UHV substation robot inspection system is shown in Figure 1. The distributed architecture and modular design strategy are adopted to design different functional modules to realize the final fault detection.

As shown in Figure 1, the work content of the UHV substation robot inspection system mainly includes three aspects:

(1) The inspection robot is used to collect image data for the operation status of various equipment in the UHV substation, and the collected image data is preprocessed. Then the processed equipment image data is sent to the remote monitoring system.

(2) Based on the image data sent by the inspection robot, the remote monitoring system classifies the data and carries out network training and fault detection in combination with its own deep learning algorithm.

(3) It outputs fault detection results.

3.2. Data Acquisition and Preprocessing. In the process of inspection in the UHV substation, the inspection robot mainly detects multisource data, such as meter reading, split and close state, temperature distribution, and foreign body recognition of different equipments in the station. Therefore, it can be seen that the inspection content includes the external conditions of the equipment and the temperature distribution of the equipment, and the comprehensive inspection in the UHV substation cannot be realized only by relying on single-source visual inspection. In addition, the structure of different devices is different, so it is difficult to achieve universal detection of infrared and visible images for all devices. The specific inspection contents and approaches of the inspection robot are as follows:

(1) Meter reading detection: It is mainly for transformer and other oil-filled equipment oil level meter detection. Visible light is used for detection.

(2) Status detection: It mainly detects the status of circuit breakers, disconnecting switches and signal lights. Visible light is used for detection.

(3) Temperature distribution detection: The temperature changes in the body and the joint of the primary equipment in the UHV substation are mainly detected. Infrared light is used for detection.

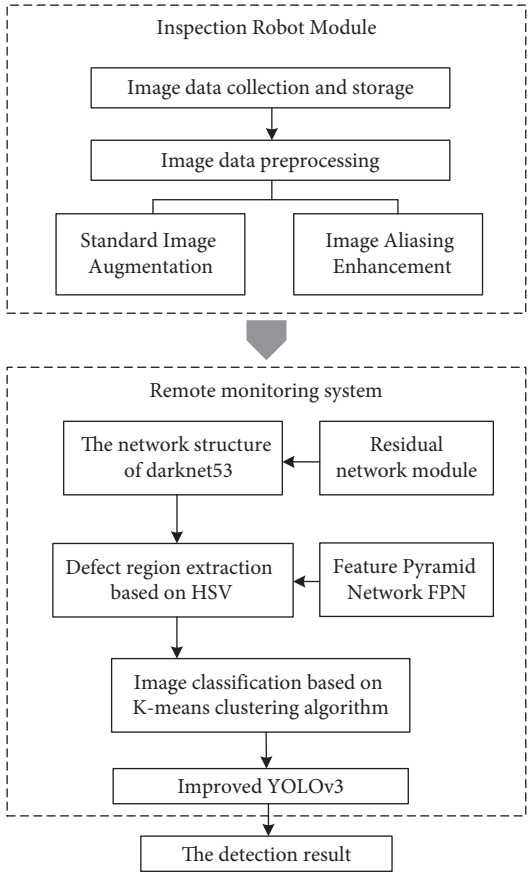


Figure 1: Overall scheme of robot inspection system for UHV substations.

(4) Foreign body recognition and detection: It mainly detects the appearance of different equipment in the UHV substation and whether there are foreign bodies. Visible light is used for detection.

In the automatic detection of power equipment, it is necessary to identify the classification information of power equipment while locating the target area in the image so as to facilitate the later data processing and abnormal output. Therefore, it is necessary to construct a multisource heterogeneous data set for visible and infrared images of all the equipments in the UHV substation and clearly mark the equipment area and specific categories of power equipment before conducting a safety inspection for the equipment in the whole station.

The LabelImage tool is adopted here which can support multicategory simultaneous labeling and can select VOC, YOLO, and other data sets to save labels. Based on visible light and infrared detection data, equipment area and specific categories of power equipment in the obtained multisource heterogeneous images of the UHV substation are marked. The VOC2007 standard dataset is used for annotation. After annotation, corresponding label files containing image file name, size, position of annotation box, and annotation target category are generated in the Annotations folder. And labels, images, and related settings files of the dataset are generated in the VOC2007 folder to complete the establishment of the dataset. The VOC2007 storage format is shown in Figure 2.

In the process of actual data acquisition, it is difficult to obtain enough image data sets of different equipment in the UHV substation due to environmental factors and actual conditions. In this case, data augmentation is generally considered to solve the problem of insufficient sample data. The principle of data augmentation is to generate new data based on existing sample data through transformation, so as to increase the number of data and improve the ability of model generalization. Standard image augmentation and image aliasing augmentation are used for data augmentation.Standard image augmentation is realized through geometric transformation or color space transformation such as translation and rotation, flip and zoom, color channel transformation, and artificial addition of voice. The specific process is shown in .

$$\begin{bmatrix} a & b & 1 \end{bmatrix}^T = B\begin{bmatrix} a_0 & b_0 & 1 \end{bmatrix}^T, \tag{1}$$

where in (1), (a, b) represents the original image coordinates and (a_0, b_0) represents the image coordinates after transformation. B represents the geometric transformation relation matrix.

Image translation transformation is to move the original image pixels in different directions, and the movement

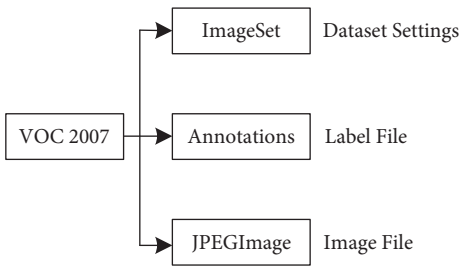


Figure 2: VOC2007 storage format.

process can be represented by matrix B_1. The specific process is shown in .

$$B_1 = \begin{bmatrix} 1 & 0 & \Delta a \\ 0 & 1 & \Delta b \\ 0 & 0 & 1 \end{bmatrix}. \tag{2}$$

In (2), Δa is the offset in the horizontal direction andΔb is the offset in the vertical direction.

Image rotation needs to consider the content and rotation angle of the original image to ensure the effectiveness of the transformed image. Generally, the central point of the image is the rotation center for rotation. The transformation matrix is B_2. The specific process is shown in (3).

$$B_2 = \begin{bmatrix} \cos\alpha & -\sin\alpha & 0 \\ \sin\alpha & \cos\alpha & 0 \\ 0 & 0 & 1 \end{bmatrix}. \tag{3}$$

In (3), α is the angle of rotation.

Image flipping transformation includes horizontal flipping and vertical flipping. Their transformation matrices are B_3 andB_4, respectively. The specific process is shown in (4) and (5).

$$B_3 = \begin{bmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \tag{4}$$

$$B_4 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix}. \tag{5}$$

Image scaling transformation needs to consider the scaling multiple to meet the detection requirements and ensure its effectiveness while scaling the image at any scale. The scaling transformation matrix is as shown in (6).

$$B_5 = \begin{bmatrix} \beta_a & 0 & 0 \\ 0 & \beta_b & 0 \\ 0 & 0 & 1 \end{bmatrix}. \tag{6}$$

In (6), β_a and β_b represent the scaling coefficients on the corresponding horizontal and vertical axes, respectively.

The image before and after transformation is shown in Figure 3.

The label corresponding to the image can be calculated, and the usual format is (a_l, b_l, a_r, b_r, c), where a_l andb_l are the coordinates of the upper left corner of the original image label box, a_r and b_r are the coordinates of the lower right corner of the original image label box, and c indicates the category of marked objects in the label box. The new label box is derived from the geometric transformation expression (1).

Image aliasing data augmentation is achieved by stitching and aliasing different images. Mosaic data enhancement method is adopted here for stitching aliasing. It can realize the stitching of four images and each image has a target box. The image processed by Mosaic contains the images and labels of the four images, which can greatly enrich the background of the detected object and improve the performance of the model. The implementation process of the Mosaic data enhancement method is as follows:

(1) Four photos are randomly read each time from the "JPEGImage" image file of the VOC2007 dataset as the original images of Mosaic.

(2) For the above four original images, geometric transformation, color change, and other operations are performed. After the operation is completed, the offset coordinates will be randomly generated and the original pictures will be placed in the order of top left, bottom left, bottom right, top right, and bottom right.

(3) The combination of images and boxes. The fixed area of four pictures is intercepted and spliced into a new picture by means of a matrix. The new picture contains a series of contents such as label box.

(4) Finally, the label box is processed. If the clipping sample contains part of the label frame, it will be discarded. The label frame that is still intact after clipping will be retained.

3.3. Multimode Fault Detection of UHV Substation Based on Improved Deep Learning. The data collected and preprocessed by robot inspection is transmitted to the remote monitoring system, and the improved deep learning module is used for feature training and fault detection.

3.3.1. Equipment Defect Extraction. Before the defect classification of the infrared image of power equipment, it is necessary to preprocess the image to judge whether the power equipment has defects according to the characteristics of the infrared image itself. If the power equipment is judged to have defects, the defect area of the power equipment is marked and the defect area is separated from other background areas. The part of the image that only contains the defect area is extracted, which can improve the accuracy of subsequent defect image classification.

For human vision, saliency region detection helps to extract important regions in a specific visual scene or an image. The color space models commonly used for significance region detection are HSV and RGB color space. RGB color space is also known as the three primary color mode, which is composed of red, green, and blue. All colors in RGB color space are generated through different degrees of fusion in the color space. In HSV

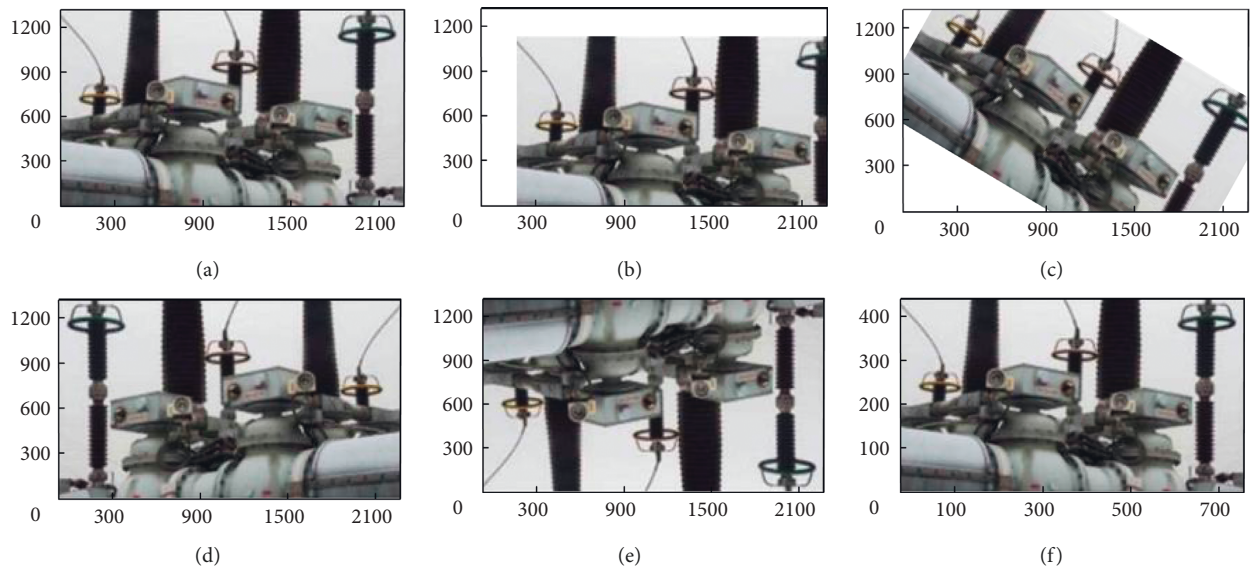


Figure 3: Comparison before and after image transformation. (a) The original image. (b) Translation transformation. (c) Image rotation. (d) Image horizontal flip. (e) Image flip vertically. (f) Image scaling.

color space, H represents chroma, S represents saturation, and V represents brightness. Infrared images mainly describe temperature, and the part with higher temperature is usually brightly reflected in the image. Compared with RGB color space, HSV color space can better perceive the connection between colors, and different H values can represent different colors.

For the defective areas of power equipment in the infrared image, we are more concerned with the areas that had yellow and white colors. To improve the accuracy of the subsequent image classification algorithm, the defect region segmentation algorithm based on HSV is adopted to separate the defect region from the background in the infrared image of faulty power equipment. Since it is impossible to determine the defect type by simply analyzing the fault area, it is necessary to segment it based on mathematical morphology according to the location of the defect area. It can reduce the interference caused by the background area in the infrared image to the subsequent classification of the defect types.

The steps of the defect region extraction algorithm based on HSV are as follows:

(1) The center point of image clustering is initialized so that each cluster center point is evenly distributed in the image with the distance d, and the pixel point with the smallest gradient is selected as the center point of the cluster in each 3×3 neighborhood.

(2) The processed image was normalized and the value of H was calculated.

(3) The threshold value was set, and the rough equipment defect area in the infrared image was separated by the threshold value. OTSU threshold segmentation algorithm is used to segment the original image to obtain the binary image I_2, and each isolated defect point is connected according to the mathematical morphology closed operation, and finally, the defect area Q_d is determined.

(4) Each connected region Q_0 in I_2 is searched for If (7) is true in the defective region, this region is the defective power equipment region to be extracted.

$$Q_0 \cap Q_d > 0.8Q_d. \tag{7}$$

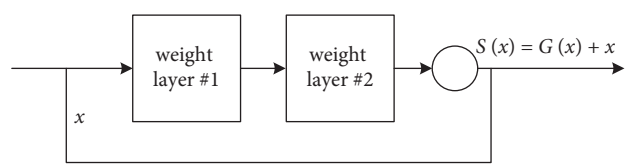


Figure 4: Residual network module.

3.3.2. Equipment Defect Training and Fault Detection

(1) Improved Deep Learning Algorithm. In this study, the target detection algorithm YOLOv3 with fast operation speed is mainly used as the main frame of detection. The main network of YOLOv3 is improved by combining residual network and deformable convolution and the convolutional neural network is redesigned so that the multiscale detection of YOLOv3 has a stronger ability of small target detection.

In the process of deep learning, problems such as gradient dispersion and gradient explosion make the model training difficult to converge. The residual network is shown in Figure 4, and the function $S(x)$ represents the output of the residual block$S(x) = G(x) + x$. The problem of gradient disappearance can be solved by transferring the main input x and using residual $S(x) = G(x) - x$ to represent the part that needs to be optimized. The residual block retains the

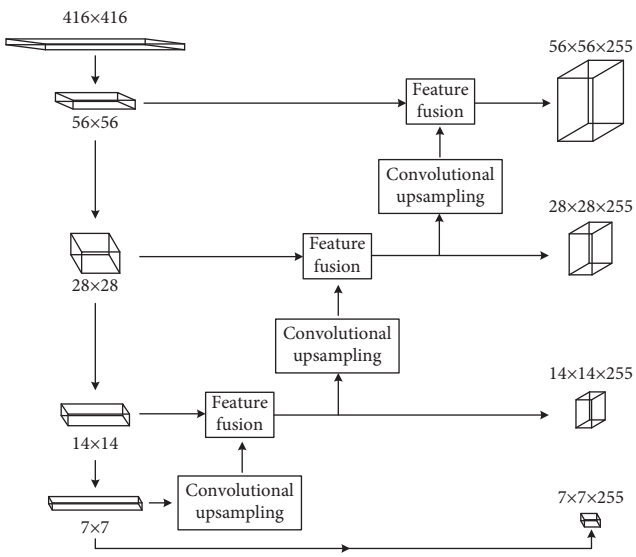


Figure 5: The basic structure of the improved YOLOv3.

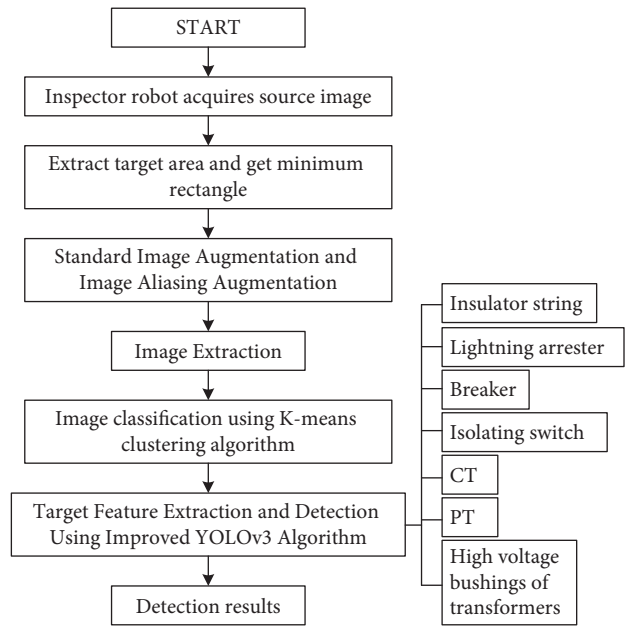


Figure 6: The process of the proposed detection method.

main information well while amplifying the sensitivity to small changes.

Based on the algorithm YOLOv2, YOLOv3 uses multiscale fusion to achieve target positioning and detection by adding an anchor mechanism to the model. The convolutional network related to YOLOv3 is the Darknet535 network. Darknet53 network does not contain the full convolutional network of the layer. YOLOv3 extracts different feature graphs of convolution for feature fusion, and respectively predicts three outputs. Finally, the detection results are obtained by the feature fusion of the three scales. Due to the different sizes of feature images, the unfused feature images cannot be directly calculated, and an interpolation algorithm is needed to fill pixels.

To improve the poor detection effect of the traditional convolution algorithm, the offset is added to the original convolution for improvement so as to extract features more effectively. The mathematical expression of the deformable convolution is shown in (8).

$$y(g_0) = \sum_{g_m \subset R} [w(g_m) x(g_0 + g_m + \Delta g_m)]. \quad (8)$$

In (8), g_0 represents the specified point in the image and g_m represents each element of the convolution kernel, and Δg_m represents the translation distance of the convolution kernel.

Since the substation equipment environment may present very small targets, there will be missed detection in the detection process. Therefore, the YOLOv3 algorithm can be optimized and improved by increasing the number of feature fusion 4 times to further improve the model's ability to detect small targets. The basic structure of the improved YOLOv3 is shown in Figure 5.

The K-means clustering algorithm aims to complete the division which has the advantages of fast classification speed and high accuracy. Therefore, it is considered to generate an anchor frame based on K-means clustering algorithm, and its objective function can be expressed as a data matrix. According to the objective function of K-means clustering, (9) is established.

$$\sum_{i=1}^{p} \sum_{j=1}^{q} \left\| x_j - \delta_i \right\|^2 = \| X - uV \|^2 = \left\| X - V^{\mathrm{T}} \left(V V^{\mathrm{T}} \right)^{-1} V \right\|^2. \quad (9)$$

In (9), $X \in i^{u \times q}$ represents the vector and $x_j \in R^u$ represents the matrix, and $U \in R^{u \times p}$ represents the matrix of class center point $\delta_i \in R^u$, and $V \in R^{p \times p}$ represents the matrix of binary indicator variable. It is shown in (10).

$$V_{ij} = \begin{cases} 1, & x_j \in L_i, \\ 0, & x_j \notin L_i. \end{cases} \quad (10)$$

In (10), $L_i\,(i = 1, 2, \ldots, p)$ represents that the data set x is divided into p classes.

To enable the improved YOLOv3 algorithm to converge quickly and improve the accuracy of detection during training, the size of the most appropriate anchor point prediction box is found by clustering and calculating the real coordinate box in the test. Then the classification is completed by calculating the distance between the samples and iterating continuously. K-means calculation is carried out for the labeled Ground True Box. The improved YOLOv3 has four detection scales, and 12 prediction boxes are need to be generated in total, so the clustering number is 12. In the target detection task, the intersection ratio between the anchor point prediction frame and the real sample frame is used to define the calculation method of clustering distance. It is shown in (11).

$$d_C = 1 - \sigma, \quad (11)$$

where in (11), σ is the intersection ratio of anchor point prediction box and the real sample box.

(2) Network Training and Fault Detection. The training samples come from the infrared images taken in the UHV substation. According to the source images, the target region is firstly extracted and the minimum rectangle of the target in the region is obtained, and then the boundary is appropriately expanded. Then the size of the image is adjusted to 32×32 to obtain the network input image. To enlarge the training samples and improve the recognition ability and generalization ability of the network, the sample images obtained by the above methods are augmented by standard image augmentation and image aliasing augmentation.

Based on the overall idea of image classification, image processing, and fault detection, a new infrared detection method is proposed combined with inspection robot and an improved deep learning algorithm. It is the multimode detection method of UHV substation fault based on robot inspection and deep learning. The specific process of this method is shown in Figure 6.

As shown in Figure 6, the specific process of multimode detection method for UHV substation faults based on robot inspection and deep learning includes the following steps:

(1) The inspection robot is used to collect image data of various equipment in the UHV substation.

(2) The target region is extracted from the source data and the minimum rectangle is obtained.

(3) Standard image augmentation and image aliasing augmentation are performed on the image data.

(4) Extract the defect area of the image and use the K-means clustering algorithm to generate the anchor box, and classify the image on this basis.

(5) The improved YOLOv3 algorithm is used for target feature extraction and fault detection.

(6) It outputs fault detection results and faulty equipment information.

4. Experiment and Analysis

4.1. Experimental Environment. The UHV substation equipment defect data set used in the experiment is mainly from a UHV substation of a provincial power company, and some images contain multiple detection objects. The data set involves a total of 14,000 images of insulators, lightning arresters, circuit breakers, isolation switches, current transformers, voltage transformers, and transformer high voltage bushing, with 2,000 images of each object category. The number of object types, labels, images, and instances detected in the dataset is shown in Table 1.

In the whole data set, 20% (2,800 pieces) are randomly selected and divided into test sets. Then 60% (8,400 pieces) are randomly selected and divided into training sets. The remaining 2,800 pieces are verification sets.

4.2. Comparison of Detection Accuracy. According to the detection results, the detection accuracy, recall rate, and average accuracy (AP) at multiple intersection ratio (IoU) thresholds are used as evaluation indexes to analyze the performance.

Table 1: Data statistics of the dataset.

Target	Number of images	Label	Number of instances
Insulator string	2000	I	2357
Lightning arrester	2000	LA	2163
Breaker	2000	B	2462
Isolating switch	2000	IS	2711
Current transformer	2000	CT	2593
Voltage transformer	2000	PT	2479
High voltage bushings	2000	HVB	2831

Table 2: The recall rate and accuracy of detection results of different methods.

Method	Accuracy (%)	The recall rate (%)	Target detection score
Proposed method	95.9	91.3	5825
Ref. [13]	83.3	89.1	3576
Ref. [14]	85.6	87.6	3695
Ref. [16]	84.8	88.7	3622

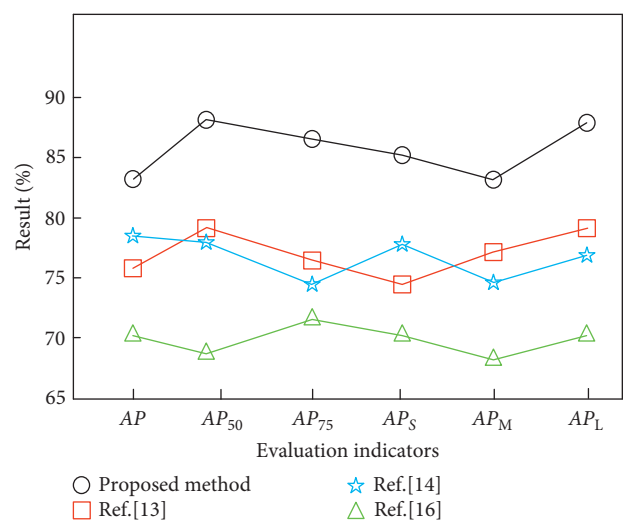


Figure 7: AP values for different methods.

AP is generally calculated by interpolation, and the calculation process is shown in (12).

$$\mathrm{AP} = \sum_{i=0} (R_{i+1} - R_i) P(R_{i+1}). \tag{12}$$

In (12), P and R respectively represent the accuracy rate and recall rate of a certain point on the original P-R curve, and i represents the interpolation point.

The multimodal detection method of UHV substation fault based on robot inspection and deep learning proposed in this study is compared with the algorithms in [13, 14, 16].

Firstly, the recall rate and detection accuracy of detection results of different algorithms are calculated, and the results are shown in Table 2.

It can be seen that the recall rate and accuracy of the proposed method are the highest according to the calculation results of the proposed method and other comparison methods in Table 2. In terms of recall rate, the improvement of the proposed method is relatively small compared with other methods. The maximum is 3.7% and the minimum is 2.2%. The recall rate of the four algorithms is relatively close. However, the accuracy of the proposed method reaches 95.9%, which is greatly improved compared with the other three methods. The maximum is 12.6% and the minimum is10.3%. The significant improvement in detection accuracy also makes the target detection score of the proposed method reach 5,825 points, which is greatly improved compared with the other methods. The maximum increase is 2,249 points and the minimum increase is 2,130 points. In conclusion, the proposed detection method has a better target detection performance compared with other methods.

The average accuracy of detection by different algorithms at multiple intersection ratio thresholds is compared, and the results are shown in Figure 7.

In Figure 7, AP_{50} and AP_{75} represent AP when the value of the official evaluation index of Microsoft Common Objects in Context (MS COCO) data set is 50 and 75, respectively. AP_S, AP_M, and AP_L represent the *Aps* of small size target (area <322), medium-size target (area <962), and large-size target (area >962), respectively.

It can be seen from Figure 7 that the proposed method has effectively enhanced the target detection effect of UHV substation equipment defects compared with other comparison algorithms, especially in multiscale detection. The *APs* of small-sized, medium-sized, and large-sized targets are all significantly improved. The number of feature fusions of the improved YOLOv3 network structure increases, and the low-dimensional feature map keeps the same scale size while changing the number of feature map channels. So the high-dimensional sampling feature map and low-dimensional feature map can be better superimposed and merged to improve the detection effect.

5. Conclusion

Aiming at the problem of automatic fault detection of equipment in UHV substation, a multimode fault detection method of UHV substation based on robot inspection and deep learning is proposed. The comparison results show that source augmentation can improve the generalization ability of the detection method to a certain extent. Using the HSV color space model to extract defect areas in images can effectively improve the accuracy of subsequent image classification. By introducing residual network and K-means clustering algorithm into traditional YOLOv3, the accuracy of the detection method improves effectively. Future work will conduct further research on how to judge the cause of the fault based on the image analysis technology on the basis of using the infrared image of the equipment to detect the fault of the power equipment.

Data Availability

The authors declare that all data sources are original.

Conflicts of Interest

The authors declare no conflicts of interest.

Acknowledgments

The authors are thankful to the science and technology project funding from State Grid Corporation of China (Project number: kj2021-059).

References

[1] B. Jalil, G. R. Leone, M. Martinelli, D. Moroni, M. A. Pascali, and A. Berton, "Fault detection in power equipment via an unmanned aerial system using multi modal data," *Sensors*, vol. 19, no. 13, pp. 108–114, 2019.

[2] S. Dong, C. Niu, and K. Dai, "Study on automatic control method of substation inspection robot based on deep reinforcement learning," *High Voltage Apparatus*, vol. 57, no. 2, pp. 172–177, 2021.

[3] R. Ding, J. Yang, M. Wu, and H. Hu, "Design and implementation of high precision positioning augmentation system for inspection robot in substations," *Intstrument Technique and Sensor*, vol. 7, no. 10, pp. 43–46, 2018.

[4] X. Zhang, S. G. Liu, and Z. Xiang, "Optimal inspection path planning of substation robot in the complex substation environment," in *Proceedings of the Chinese Automation Congress 2019*, pp. 5064–5068, Piscataway, NY, USA, 2019.

[5] X. Wang, P. Zhou, J. Hou, S. Wang, and X. Lin, "Research on algorithm of inspection path planning for substation robot," *Computer Engineering and Application*, vol. 57, no. 14, pp. 245–250, 2021.

[6] F. Kun, L. Ma, and Y. Sun, "Navigation of substation inspection robot based on laser sensor," *Transducer and Microsystem Technology*, vol. 38, no. 2, pp. 118–120, 2019.

[7] W. Xu, S. W. Yang, S. Y. Mu, R. Huang, J. Li, and H. Wang, "Modularization design of inspection robot based on substation," in *Proceedings of the 3rd IEEE Advanced Information Technology, Electronic and Automation Control Conference (IAEAC)*, pp. 2330–2333, Piscataway, NY, USA, 2018.

[8] Bo Niu, F. Ma, P. Ding et al., "Intelligent inspection and location technology of GIS partial discharge and its application," *High Voltage Apparatus*, vol. 56, no. 1, pp. 188–196, 2020.

[9] X. Zhang and Y. L. Qian, "An automatic defect detection method for gas insulated switchgear," in *Proceedings of the 4th IEEE Information Technology, Networking, Electronic and Automation Control Conference (ITNEC)*, pp. 1217–1220, Piscataway, NY, USA, 2020.

[10] Y. Yan, W. Jiang, Z. Luo, J. Zhang, and W. Liu, "System optimization and robustness stability control for GIS inspection robot in complex microgrid networks," *Complexity*, vol. 2021, Article ID 6691905, 12 pages, 2021.

[11] X. Peng, J. Liang, R. Wang, L. Yi, and G. Chen, "Substation robot intelligent inspection technology and its application," *High Voltage Apparatus*, vol. 55, no. 4, pp. 223–232, 2019.

[12] R. Wu, D. Li, J. Qin, and Y. Lan, "System design of an indoor inspection robot driven by a flexible cable in a substation," *Power System Protection and Control*, vol. 49, no. 10, pp. 89–97, 2021.

[13] C. Wang, L. Yin, Q. Zhao, W. C. Wang, and B. Luo, "An intelligent robot for indoor substation inspection," *Industrial Robot: The International Journal of Robotics Research and Application*, vol. 47, no. 5, pp. 705–712, 2020.

[14] C. Zhang, Z. Lu, and X. Liu, "Joint inspection technology and its application in a smart substation," *Power System Protection and Control*, vol. 49, no. 9, pp. 158–164, 2021.
[15] T. Wu, J. Guo, X. Gou, Q. Huang, and W. Zhou, "Method of detecting substation equipment in infrared images based on improved Gaussian convolution kernel," *Infrared Technology*, vol. 43, no. 3, pp. 230–236, 2021.
[16] M. Cheng and D. Xiang, "The design and application of a track-type Autonomous inspection robot for electrical distribution room," *Robotica*, vol. 38, no. 2, pp. 185–206, 2020.

Hindawi
Journal of Robotics
Volume 2022, Article ID 3851506, 8 pages
https://doi.org/10.1155/2022/3851506

Research Article

A Personalized Travel Route Recommendation Model Using Deep Learning in Scenic Spots Intelligent Service Robots

Qili Tang

School of Economics and Management, Aba Teachers University, Aba, Sichuan 623002, China

Correspondence should be addressed to Qili Tang; 20109637@abtu.edu.cn

Received 8 February 2022; Accepted 19 March 2022; Published 21 April 2022

Academic Editor: Shan Zhong

This paper proposes a personalized tourist interest demand recommendation model based on deep neural network. Firstly, the basic information data and comment text data of tourism service items are obtained by crawling the relevant website data. Furthermore, word segmentation and word vector transformation are carried out through Jieba word segmentation tool and Skip-gram model, the semantic information between different data is deeply characterized, and the problem of very high vector sparsity is solved. Then, the corresponding features are obtained by using the feature extraction ability of DNN's in-depth learning. On this basis, the user's score on tourism service items is predicted through the model until a personalized recommendation list is generated. Finally, through simulation experiments, the recommendation accuracy and average reciprocal ranking of the proposed algorithm model and the other two algorithms in three different databases are compared and analyzed. The results show that the overall performance of the proposed algorithm is better than the other two comparison algorithms.

1. Introduction

Service robot refers to an autonomous or semiautonomous robot [1] that completes useful service activities instead of human beings, but does not engage in production work. Its role is to replace service personnel and provide services required by human beings. Service robot contains many scientific knowledge including mechanical engineering, automation, computer science, and control engineering [2–4]. With the continuous development of artificial intelligence, service robots are gradually moving towards intelligence [5, 6]. Due to the increasingly obvious trend of personalized tourism service selection by users, the research on intelligent tourism service robot has become a hot spot in intelligent service robot [7, 8], which is innovative and forward-looking in the field of robot application in tourism industry.

At present, the traditional method of providing services by service personnel cannot meet people's personalized needs [9, 10]. Reference [11] developed a robot partner for information support and proposed a new method that can flexibly recommend all kinds of information according to human intention. However, this method does not divide the user's access sequence into different interest segment sequences according to time, so the recommendation accuracy is low. For the dynamic traveling repairman problem and dynamic vehicle routing problem, Reference [12] combined with time-varying requirements, sent the available robots to the nearest service request location, sent multiple robots for each service request arriving in the system, and proposed a new model-free operation strategy independent of load factor. However, this method does not have model algorithm and cannot be applied to more complex situations. In order to realize long-term autonomous operation, Reference [13] proposed a modular general software framework for intelligent mobile robot, which can use complex human voice commands to interact with humans. However, this method only considers the home service robot. Reference [14] introduced local attention and nonlinear attention to capture local and global project information at the same time. On this basis, a nonlinear attention similarity model (NASM) was proposed for project-based collaborative filtering through local attention embedding. However, the algorithm cannot accurately capture human high-order sequence

behavior, and it is difficult to realize complex recommendation. Reference [15] proposed a personalized robot service system centered on robot thinking, which can migrate with the user's geospatial movement at any time, so that it can continue to grow with the user. However, this method does not consider the environmental factors of users, and the personalized growth cycle is long. Reference [16] constructs the intelligent robot control system based on the principle of human-computer interaction and designs the corresponding model-based control algorithm to identify the dynamic model of the robot. However, this method is difficult to obtain the prior distribution, and it is difficult to characterize the high-dimensional semantics of users. In Reference [17], aiming at the path-planning problem of hospital service robot in drug delivery, medical insurance order and other services, based on the automatic control robot with visual recognition ability, combined with the three-dimensional reconstructed image and the route area shunting method using edge calculation, an image edge detection algorithm based on three-dimensional features is proposed. However, this method is only suitable for special people in specific areas and does not have the characteristics of in-depth personalized service.

Based on the above analysis, aiming at the problem of personalized travel route recommendation of intelligent service robot in the scenic spot, a personalized tourist interest demand recommendation model based on deep learning using word embedding technology is proposed. The basic idea is to (1) reduce the sparsity of data vector and improve the recommendation accuracy of the algorithm by preprocessing the original data and (2) build a depth prediction model to deeply mine the relationship between users and scenic spots. Compared with the traditional service robot travel route recommendation method, the contributions of the proposed method are as follows:

(1) The Skip-gram model in the word2vec word embedding method is used to transform the word vector of data, and the effective extraction of topic feature vector, geographic factor feature vector, and user access feature vector is realized.

(2) The proposed model uses the deep neural network to transform the recommendation of tourists' interests and needs into the task of binary classification, which improves the ability of extracting the features of the original data and effectively enhances the performance of predicting users' ratings.

The rest of this paper is organized as follows: the second part introduces the personalized tourist interest demand recommendation model based on deep neural network; the third part compares with the existing recommendation model to realize the feasibility and optimization of the method proposed in this paper; the fourth part is the conclusion of this paper.

2. Proposed Model

2.1. Overall Framework. Collaborative filtering technology is the most used in the recommendation system, which can be applied in various occasions and achieve good results, as well as tourism service recommendation [18]. Although collaborative filtering technology has many advantages, such as good processing of unstructured data and high degree of personalization and automation of recommendation, collaborative filtering technology also has the problem of data sparsity [19]. Generally, in the application of recommendation system, the data scored by users are insignificant for the overall data, which will lead to some problems. In this case, the use of collaborative filtering technology often cannot achieve good results. In other words, for collaborative filtering technology, if an item has less scores, it is difficult to recommend it to other users. In addition, if a user scores very little, it will be difficult to get some recommendations. It is difficult to achieve excellent results by using traditional collaborative filtering technology.

Therefore, this paper proposes a tourism service recommendation model based on deep learning to solve the related problems. The proposed model is mainly divided into four modules: data preprocessing, construction of depth prediction model, network training, and final recommendation list generation. The principle and function of each module are shown in Figure 1:

(1) Data acquisition and preprocessing—this module mainly obtains the basic information data and comment text data of tourism service items by crawling the relevant website data and then preprocesses these data.

(2) Construction of prediction model—this module uses deep learning technology to predict users' scores on tourism service items.

(3) Training network—the module uses sample data to train the model network, mine the potential relationship between users and tourism service items, and learn the interaction between users and tourism service items, so as to obtain a predictable model.

(4) Generate personalized recommendation list—the main function of this module is to test the experimental data. The module inputs the experimental data into the trained model. The model predicts users' scores on tourism service items, sorts them according to the score, and finally generates a personalized recommendation list for each user to complete the recommendation.

2.2. Data Acquisition and Preprocessing. The goal of data acquisition mainly includes three aspects: self-built database, data from Foursquare, and data from Tokyo [20, 21]. The self-built database adopts the data set established by collecting a large amount of travel information from the MaFengWo, with about 390000 user travel access records. Foursquare's data includes the long-term (about 10 months) check-in data collected in New York from May 2013 to March 2014, filtering out users with less than 8 access records and locations with less than 8 visits. There are 1158 users and 5092 locations for the experiment, with a total of 257221 check-in records in the Foursquare data set. The data

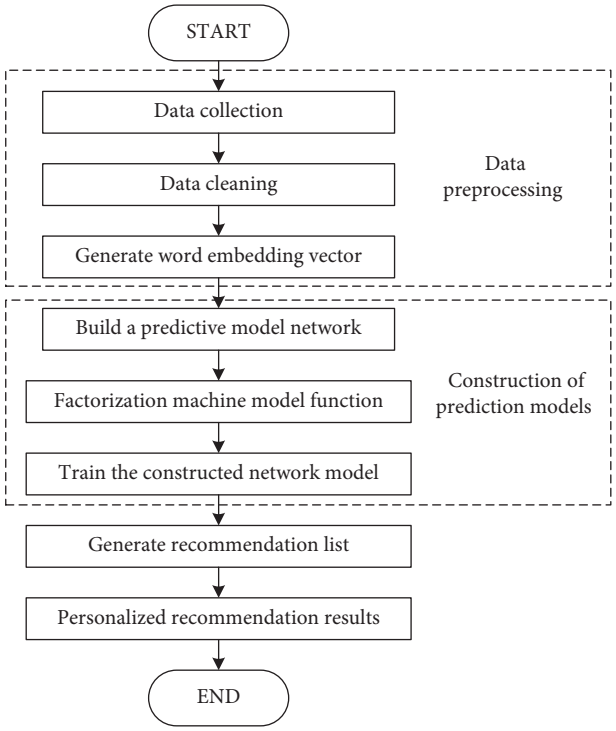


Figure 1: Implementation flow of the proposed recommendation model.

from Tokyo are similar to the data from Foursquare, and the same filtering operation is adopted for the data. Finally, there were 2095 users and 8246 locations for the experiment, with a total of 605893 check-in records.

The specific statistics of the three data sets are shown in Table 1.

Table 1: Recommended data set of tourism series.

Database	Self-built	Foursquare	Tokyo
Number of users	9328	1158	2095
Number of locations	1643	5092	8246
Number of check-in records	392031	257221	605893

2.3. Vector Representation of Text. Preprocessing the data in the above three databases mainly includes the following steps:

(1) Chinese word segmentation technology is used for word segmentation. In this paper, Jieba word segmentation tool is used to segment data [22]. Jieba word segmentation tool adopts the Chinese word segmentation algorithm of NShort, which is the python implementation of the algorithm. It has the characteristics of simple principle, easy to understand, low model resource occupation, and easy training. In addition, NShort Chinese word segmentation algorithm has excellent efficiency in large-scale word segmentation application scenarios and is widely used in various commercial fields. Moreover, the model supports incremental expansion. It is the mainstream algorithm of Chinese word segmentation and is used by most search engine companies.

(2) Convert the divided data into word vectors. As a method of text feature representation, vocabulary is expressed as a feature vector, which is called word vector [23]. It is a common expression of word vectors to represent text information with feature vectors on a word list. The data can be expressed as feature vectors using the frequency and TF-IDF value of words. A typical representation is One-Hot Encoding. However, this method essentially uses a vector containing only one 1 and the others are 0 to uniquely represent words, and its dimension is the number of words in the whole vocabulary list. However, the vocabulary list used in this representation method is very large, so the dimension of feature vectors represented by this method is very large, which eventually leads to the problem of very high vector sparsity. Secondly, such methods cannot express the relationship between words and cannot well reflect the deep semantic information between words and text. Here, the word embedding method is used for word vector conversion [24]. Different from the traditional lexical feature expression methods, the word embedding method can represent words by dense real number vectors in low dimensional space. This cannot only represent words in vector form but also calculate the effective distance between two words and describe the semantic information between words. It is a very effective method to process text information. At present, word2vec method is the most widely used word embedding technology. Word2vec model is divided into Skip-gram model and CBOW model [25]. Skip-gram model predicts the generation probability of context vocabulary through target vocabulary, while CBOW model uses context vocabulary to predict the generation probability of target vocabulary. This paper mainly uses Skip-gram model for word vector representation, and its basic structure is shown in Figure 2.

As can be seen from Figure 2, Skip-Gram model is a neural network model, including input layer, hidden layer, and output layer. Firstly, the vocabulary is transformed into One-Hot encoding form and input into the input layer and then calculated in the hidden layer. The output layer outputs the probability of the target context vocabulary. When the model training is completed, the weight from the input layer to the hidden layer can be used to represent the target word vocabulary. This is because in the weights at this time, only the weights at the position of "1" in the one-hot encoder are activated, and the number of these weights is the same as the number of hidden layer nodes so that the vector composed of these weights can represent the target vocabulary. The position of 1 of one-hot encoder of different words is different, so the target word is uniquely mapped into a low-dimensional dense vector.

The kip-Gram model predicts the probability of target vocabulary context vocabulary using the following formula (1):

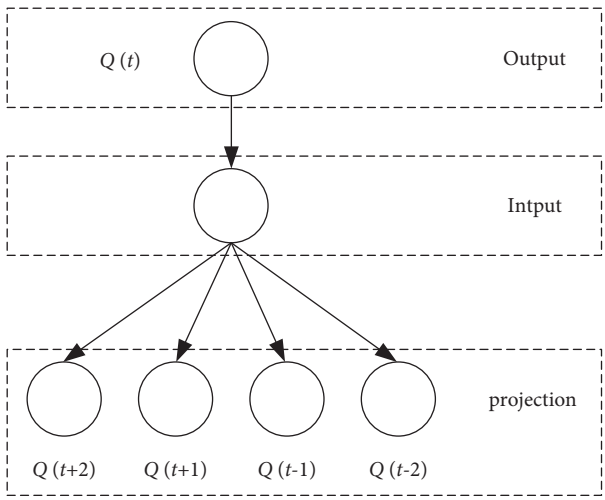


Figure 2: Basic architecture of Skip-gram model.

$$P[S(\chi)\,|\,\chi] = \prod_{a\in S(\chi)} P(v\,|\,\chi), \tag{1}$$

where, $S(\chi)$ represents the context vocabulary of the target vocabulary χ. When using softmax to calculate $P(v\,|\,\chi)$, the calculation efficiency is often very low. Therefore, the existing models often use hierarchical softmax for efficient calculation. Combined with hierarchical softmax, formula (1) is expanded into formula (2):

$$P[S(\chi)|\chi] = \prod_{a\in S(\chi)} \prod_{i=2}^{k^v} \left[\delta\left(v_\chi^T x_{i-1}^v\right)\right]^{\left(1-z_i^v\right)} \cdot \left[1-\delta\left(v_\chi^T x_{i-1}^v\right)\right]^{z_i^v}, \tag{2}$$

where, k^v represents the path length of the context vocabulary in the output hierarchy tree. v represents the context-sensitive vocabulary of the target vocabulary χ. v_χ represents the output vector of the target vocabulary. x_{i-1}^v represents the output word vector at the corresponding level under a context word.

z_i^v represents the logistic output indicator variable, when $z_i^v = 0$, it is expressed as $P(z_i^v|v_\chi, x_{i-1}^v) = \delta(v_\chi^T x_{i-1}^v)$, and when $z_i^v = 1$, it is expressed as $P(z_i^v|v_\chi, x_{i-1}^v) = 1 - \delta(v_\chi^T x_{i-1}^v)$.

In the process of word embedding model training, some negative samples are often added to improve the training speed and improve the quality of the obtained word vector. At this time, the objective function of Skip-gram model training can be expressed as follows:

$$G = \sum_{\chi\in C} \sum_{\tilde{\chi}\in S(x)} \sum_{v\in\{\chi\}\cup M_{\tilde{\chi}}} \left\{ y_v^\chi \cdot \log\left[\delta\left(v_{\tilde{\chi}}^T x_v\right)\right] + \left(1-y_v^\chi\right) \cdot \log\left[1-\delta\left(v_{\tilde{\chi}}^T\right)\right]\right\}. \tag{3}$$

x_v, $v_{\tilde{\chi}}$, and y_v^χ represent parameter vector, context word embedding vector, and logistic indicator variable, respectively, in which the sampled vocabulary set of vocabulary χ is represented by $M_{\tilde{\chi}}$. In deep learning tasks based on neural networks, the word embedding vector generated by Skip-gram model can be used as good input data. This method is used to map word vectors.

For the comment information that has been divided into words, the ultimate goal is to express the comment information as a word vector matrix and input it into the neural network. In order to achieve this goal, all comments written by user v, that is, user comments, are integrated into a single document and recorded as $z_{1:m}^v$. Document $z_{1:m}^v$ consists of m words in total, as shown in equation (4).

$$K_{1:m}^v = \Psi\left(z_1^v\right)\oplus\Psi\left(z_2^v\right)\oplus\Psi\left(z_3^v\right)\oplus\cdots\oplus\Psi\left(z_m^v\right), \tag{4}$$

where, z_1^v represents the first word in comment document, $z_{1:m}^v$, z_2^v represents the second word in comment document $z_{1:m}^v$, and so on. Function $\Psi(z_1^v)$ is used to map the first word into a low-dimensional dense real number word vector through word embedding technology. Symbol $\oplus$ is an association operation, which combines the word vectors of each word by line to form a word vector matrix. Compared with the traditional word bag technology, using this method to process the formed word vector matrix maintains the order of words in the sentence to the greatest extent, so that the order information of words can be well-preserved in the generated word vector matrix $K_{1:m}^v$, which is of great help and a good advantage for further processing.

This method is also used to process the comment information of tourism service items. All comment texts obtained from tourism service item v are also integrated into a single document $z_{1:m}^v$ and then the document is transformed into word vector matrix $K_{1:m}^v$ by the above method.

Information other than the comments of users and tourism service items is called other information, which includes the basic information of users and tourism service items. The age in the user's basic information is normalized by $x = x/120$, the gender is directly normalized to the real value of $[0, 1]$, and the occupation and city are directly converted to the word embedding vector. In addition to the comment text, the user's historical evaluation information and the evaluated item name also need to be considered. Because in terms of tourism service projects, the name of tourism service projects can also reflect the characteristics of this project. For example, for scenic spots, if a scenic spot is called "XX mountain," we can know that the scenic spot may be natural scenery. If a scenic spot is called "XX Museum," we know that the scenic spot may belong to buildings or museums. Therefore, the user's historical evaluation item name is transformed into a word embedding vector. For tourism service items, the names and locations in its basic information can be directly transformed into word embedding vectors. For tags, tags are mainly used to describe a

few sentences summarizing tourism service items, so its processing is also similar to comment information. First, word segmentation is carried out and then word embedding vector transformation is carried out for the divided words. Through the above operations, other information is also transformed into vector form.

2.4. Network Construction. This chapter proposes to use the deep learning model of DNN for personalized POI recommendation. The network model is shown in Figure 3 and mainly includes two modules: feature extraction module and network learning module. Among them, the feature extraction module uses word embedding technology to extract and construct the features of location-based social networks. The network learning module includes network connection layer and network layer. The function of the connection layer in the network learning module is to fuse the extracted feature vectors. The function of the network layer is to train the proposed model and predict the score of users' preference for interest demand recommendation.

In the process of building effective features recommended by tourists' interest needs, the feature extraction module extracts topic feature vector, geographic factor feature vector, and user access feature vector through word embedding technology [26]. The network connection layer in the network learning module fully connects the vectors extracted by the feature extraction module and sends them into the deep neural network. The connection layer realizes the expansibility of the model. For example, if the relevant context information needs to be added in the application, it can be fully connected automatically through the full-connection layer to send the characteristics of the input layer into the network training. The network layer in the network learning module includes the following two functions:

(1) Training function. In the training stage, the network layer uses the deep neural network to learn and extract the implicit features, obtain the high-order interaction between the features, and then take the high-order feature results extracted from the implicit layer as the input of the softmax layer to learn the classification task. The proposed personalized tourist interest demand recommendation model based on deep learning using word embedding technology transforms the tourist interest demand recommendation into a binary classification task, in which the user's check-in record is defined as a positive sample, and the tourist interest demand points not visited by the user are regarded as a negative sample. The output result of the softmax layer in the model is a two-dimensional probability vector $P = [A, B]$, where A represents the user's preference probability for the tourist interest demand point and B represents the user's nonpreference probability for the tourist interest demand point. Then, the cross entropy loss function is selected, and the gradient descent method is used to optimize the function.

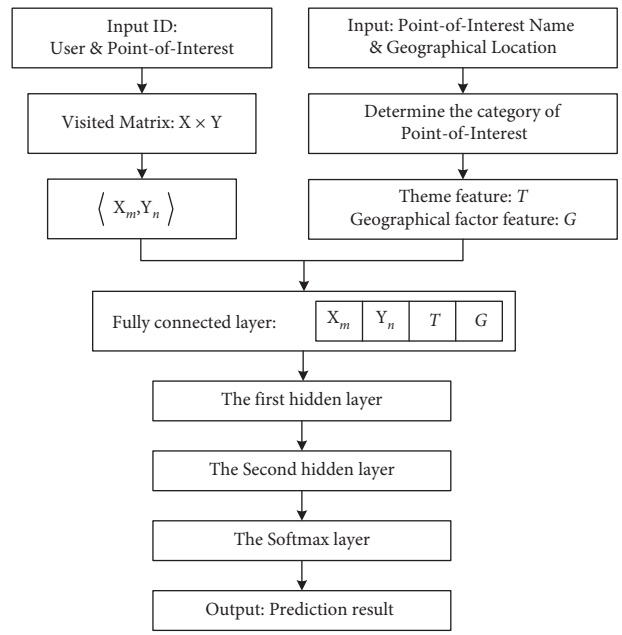


Figure 3: Basic architecture of the proposed network model.

(2) Prediction function. In the prediction stage, input the interest and demand point information of users, the network outputs a probability vector P, and recommend to users according to the probability ranking A_{top-i} of A.

2.5. Network Training. Taking the text features constructed in the previous stage as the input of the model, the final prediction results of tourists' interest and demand points can be obtained after training.

The connection layer of DLM learning module fully connects the existing features. The vector representation of the full-connection of any user-tourist interest demand point $\langle X_m, Y_n \rangle$ is shown in the following equation (5).

$$Q_0 = \mathrm{Me}\left[\langle X_m, Y_n \rangle, T, G\right], \tag{5}$$

where, T and G represent the user subject features and geographical factor features represented by the potential vector $\langle X_m, Y_n \rangle$ of user tourist interest demand points, respectively. Me represents the process of connecting all feature vectors into a one-dimensional vector and sending it into the model, which can be calculated in the hidden layer of the model according to the following equation (6).

$$Q_1 = Dr\left[Af\left(W_\sigma \cdot Q_0 + \theta\right)\right], \tag{6}$$

where, σ represents the number of hidden layers in the model. W represents the weight value of the hidden layer. θ represents the offset value of the hidden layer. Af indicates the activation function.

In addition, the dropout layer is added in the training process of each hidden layer to prevent over fitting problem.

In the output layer of the model, the predicted probability of the user's interest demand for tourists can be obtained, as shown in formula (7) below.

$$I = soft\,max\left(W_O \cdot Q_1 + \theta_O\right), \tag{7}$$

where, W_O represents the weight value of the output layer. θ_O represents the offset value of the output layer. soft max represents the normalized exponential function, and its output is two probability values, representing the probability that the user visits the tourist's interest demand and the probability that the user does not visit.

The proposed model uses a combined regularization, and the adjusted loss function is shown in formula (8):

$$L = smooth\left(u_p - x\right) + \alpha\left(\left\|\gamma_1\right\|_2^2 + \left\|\gamma_2\right\|_2^2\right), \tag{8}$$

where, u_p represents the final desired prediction result. α represents the regularization parameter. γ_1 represents the learning parameters of the aggregation layer. γ_2 represents the learning parameters of the layer of interest, and x represents the input data. smooth $(u_p - x)$ can be calculated by formula (9).

$$smooth\left(u_p - x\right) = \begin{cases} \frac{1}{2}\left(u_p - x\right)^2, & \left|u_p - x\right| < 1, \\ \left|u_p - x\right| - \frac{1}{2}, & \left|u_p - x\right| \geq 1. \end{cases} \tag{9}$$

The model is optimized through formula (8), and the sorted $A_{\text{top}-i}$ is output as the recommendation result, where i represents the number of recommendation results of tourist interest and demand.

Finally, according to the prediction results obtained by the model, the recommendation results are determined according to the user access probability. The higher the probability, the more likely the user is to access. When determining the value i, select the corresponding first i prediction results according to the ranking results of probability to recommend to users.

3. Experiment and Analysis

The software environment of the experiment is 64 bit Ubuntu 16 Python 3 and tensor flow deep learning platform installed under 04 operating system. The computer hardware configuration used in the experiment is Intel (R) core (TM) i7 CPU and single NVIDIA GTX Titan GPU.

3.1. Evaluating Indicator. In order to verify the effectiveness of the proposed method, two evaluation indexes, accuracy HR@N and average reciprocal ranking MRR@N, are used. The calculation of accuracy HR@N and average reciprocal ranking MRR@N are shown in equation (10) and (11), respectively.

$$HR@N = \frac{Number\,Of\,Hits@N}{GTGT}, \tag{10}$$

$$MRR@N = \frac{1}{N}\sum_{i=1}^{N}\frac{1}{rank_i}, \tag{11}$$

Table 2: Performance comparison of three algorithms on self-built data sets.

Algorithm	HR@15	MRR@15	HR@30	MRR@30
This paper	0.489	0.206	0.625	0.317
Ref. [11]	0.365	0.158	0.486	0.205
Ref. [14]	0.327	0.139	0.437	0.194

Table 3: Performance comparison of three algorithms on Foursquare data set.

Algorithm	HR@15	MRR@15	HR@30	MRR@30
This paper	0.403	0.317	0.576	0.395
Ref. [11]	0.315	0.231	0.374	0.284
Ref. [14]	0.258	0.207	0.332	0.276

Table 4: Performance comparison of three algorithms on Tokyo data set.

Algorithm	HR@15	MRR@15	HR@30	MRR@30
This paper	0.383	0.301	0.513	0.358
Ref. [11]	0.295	0.212	0.328	0.252
Ref. [14]	0.226	0.194	0.295	0.231

where, Number Of Hits@N represents the correct quantity in the prediction results. GT represents the total number of prediction results. N represents the total number of nodes in the network model. rank_i represents the predicted result sequence.

3.2. Experimental Results and Analysis. The algorithm in Ref. [11, 14] are selected for comparative analysis. Firstly, the self-built data set is used as the training set, and the number of nodes in the network model is set to 15 and 30, respectively. The results are shown in Table 2.

Then, the Foursquare data set is used as the training set, and the number of nodes in the network model is set to 15 and 30, respectively, to verify the three different algorithms. The results are shown in Table 3.

Finally, based on the data set of Tokyo, the number of nodes in the network model is set to 15 and 30, respectively, to verify three different algorithms. The results are shown in Table 4.

As can be seen from Tables 2–4, the accuracy HR@N and average reciprocal ranking MRR@N of the proposed algorithm on three different data sets are the best. This is because the algorithm proposed in this paper uses the Skip-gram model to improve the extraction performance of user access feature vector, uses deep neural network to process tourists' interest demand recommendation, which can better capture the sequence relationship, and better track users' footprints and users' points of interest, which greatly improves the recommendation performance. In the process of intelligent recommendation, the method in Ref. [11] does not consider the sequential process of tourists in the tourism process and does not divide the user access sequence into different interest segments according to time. Although the

method in Ref. [14] can model local sequence behavior, it has some limitations in capturing high-order sequence behavior. For tourists' complex interest point selection, feature extraction cannot be carried out.

4. Conclusion

This paper proposes a personalized tourist interest demand recommendation model based on deep learning and compares the proposed algorithm model with the other two algorithms through simulation experiments. The proposed model uses the deep neural network to transform the recommendation of tourists' interests and needs into the task of binary classification, which improves the ability of extracting the features of the original data and effectively enhances the performance of predicting users' ratings.

The future work will deeply study the method of real-time personalized travel route recommendation by improving the algorithm speed and the personalized route recommendation method for group travel.

Data Availability

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

Conflicts of Interest

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

References

[1] A. C. Ma, Z. W. Meng, and X. R. Ding, "Performance review of intelligent guidance robot at the outpatient clinic setting," *Cureus*, vol. 13, no. 8, pp. 45–52, 2021.

[2] Z. Z. Yin, D. P. Wang, and J. H. Liu, "A method of constructing robotics service platform for assisting handicapped or elderly people," *Journal of Robotics*, vol. 2020, Article ID 4259175, 6 pages, 2020.

[3] H. L. Qiu, M. L. Li, B. Y. Shu, and B. Bai, "Enhancing hospitality experience with service robots: the mediating role of rapport building," *Journal of Hospitality Marketing & Management*, vol. 29, no. 3, pp. 247–268, 2019.

[4] S. B. Kelley, B. W. Lane, and J. M. DeCicco, "Pumping the brakes on robot cars: current urban traveler willingness to consider driverless vehicle," *Sustainability*, vol. 11, no. 18, pp. 167–176, 2019.

[5] H. X. Lin, O. H. X. Chi, and D. Gursoy, "Antecedents of customers' acceptance of artificially intelligent robotic device use in hospitality services," *Journal of Hospitality Marketing & Management*, vol. 29, no. 5, pp. 530–549, 2020.

[6] H. Wu, X. J. Wu, Q. Ma, and G. Tian, "Cloud robot: semantic map building for intelligent service task," *Applied Intelligence*, vol. 49, no. 2, pp. 319–334, 2019.

[7] I. P. Tussyadiah, F. J. Zach, and J. X. Wang, "Do travelers trust intelligent service robots?" *Annals of Tourism Research*, vol. 81, no. 14, pp. 125–134, 2020.

[8] S. Park, "Multifaceted trust in tourism service robots," *Annals of Tourism Research*, vol. 81, no. 7, pp. 168–175, 2020.

[9] W. Shuai and X. P. Chen, "KeJia: towards an autonomous service robot with tolerance of unexpected environmental changes," *Frontiers of Information Technology & Electronic Engineering*, vol. 20, no. 3, pp. 307–317, 2019.

[10] J. X. Wang and Z. S. Shi, "A speech interaction system based on cloud service under ROS," in *Proceedings of the 38TH Chinese Control Conference (CCC)*, pp. 4721–4725, IEEE, Guangzhou, China, July 2019.

[11] S. Yamamoto and N. Kubota, "Development of smart device interlocked robot partners for information support and smart recommendation," in *Proceedings of the 2020 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE)*, pp. 338–344, Electr Network, Glasgow, UK, July 2020.

[12] C. Ruch, J. Gachter, J. Hakenberg, and E. Frazzoli, "The +1 method: model-free adaptive repositioning policies for robotic multi-agent systems," *IEEE Transactions on Network Science And Engineering*, vol. 7, no. 4, pp. 3171–3184, 2020.

[13] J. B. Yi, T. Kang, D. Song, and S.-J. Yi, "Unified software platform for intelligent home service robots," *Applied Sciences-Basel*, vol. 10, no. 17, pp. 354–361, 2020.

[14] Z. P. Shan, Y. Q. Lei, D. F. Zhang, and J. Zhou, "NASM: nonlinearly attentive similarity model for recommendation system via locally attentive embedding," *IEEE Access*, vol. 7, no. 5, pp. 70689–70700, 2019.

[15] L. Hu, Y. G. Jiang, F. X. Wang, K. Hwang, M. S. Hossain, and G. Muhammad, "Follow me robot-mind: cloud brain based personalized robot service with migration," *Future Generation Computer Systems-The International Journal of Escience*, vol. 107, no. 6, pp. 324–332, 2020.

[16] S. R. Pan, "Design of intelligent robot control system based on human-computer interaction," *International Journal of System Assurance Engineering And Management*, vol. 12, no. 3, pp. 11–18, 2021.

[17] L. K. Fan, X. C. Li, C. S. Guo, and B. Jia, "Path control of panoramic visual recognition for intelligent robots based-edge computing," *Computer Communications*, vol. 178, no. 8, pp. 64–73, 2021.

[18] Y. Ishida, T. Morie, and H. Tamukoh, "A hardware intelligent processing accelerator for domestic service robots," *Advanced Robotics*, vol. 34, no. 14, pp. 947–957, 2020.

[19] F. Lu, M. Huang, X. L. Li, G. Tian, H. Wu, and W. Si, "Learning and development of home service robots' service cognition based on a learning mechanism," *Applied Sciences-Basel*, vol. 10, no. 2, pp. 259–266, 2020.

[20] D. Yang, D. Zhang, V. W. Zheng, and Z. Yu, "Modeling user activity preference by leveraging user spatial temporal characteristics in LBSNs," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 45, no. 1, pp. 129–142, 2014.

[21] Y. G. Wang, X. M. Cai, C. L. Xu, and L. Jun, "Rise of the machines: examining the influence of professional service robots attributes on consumers' experience," *Journal of Hospitality And Tourism Technology*, vol. 12, no. 4, pp. 609–623, 2021.

[22] A. Onan, "Two-stage topic extraction model for bibliometric data analysis based on word embeddings and clustering," *IEEE Access*, vol. 7, pp. 145614–145633, 2019.

[23] A. Onan and S. Korukoğlu, "A feature selection model based on genetic rank aggregation for text sentiment classification," *Journal of Information Science*, vol. 43, no. 1, pp. 25–38, 2017.

[24] A. Onan, "Biomedical text categorization based on ensemble pruning and optimized topic modelling," *Computational and Mathematical Methods in Medicine*, vol. 2018, Article ID 2497471, 22 pages, 2018.

[25] D. Lian, C. Zhao, X. Xie, G. Sun, E. Chen, and Y. Rui, "GeoMF: joint geographical modeling and matrix factorization for point-of-interest recommendation," in *Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 831–840, New York, NY, USA, August 2014.

[26] R. Akabane and Y. Kato, "Pedestrian trajectory prediction based on transfer learning for human-following mobile robots," *IEEE Access*, vol. 9, no. 12, pp. 126172–126185, 2021.

Hindawi
Journal of Robotics
Volume 2022, Article ID 4163992, 12 pages
https://doi.org/10.1155/2022/4163992

Research Article

Support Vector Machine and Granular Computing Based Time Series Volatility Prediction

Yuan Yang and Xu Ma

School of Mathematics and Computer Science, Ningxia Normal University, Ningxia, Guyuan 756000, China

Correspondence should be addressed to Yuan Yang; sjyangyuan@nxnu.edu.cn

Received 17 December 2021; Revised 13 February 2022; Accepted 14 February 2022; Published 16 April 2022

Academic Editor: Shan Zhong

With the development of information technology, a large amount of time-series data is generated and stored in the field of economic management, and the potential and valuable knowledge and information in the data can be mined to support management and decision-making activities by using data mining algorithms. In this paper, three different time-series information granulation methods are proposed for time-series information granulation from both time axis and theoretical domain: time-series time-axis information granulation method based on fluctuation point and time-series time-axis information granulation method based on cloud model and fuzzy time-series prediction method based on theoretical domain information granulation. At the same time, the granulation idea of grain computing is introduced into time-series analysis, and the original high-dimensional time series is granulated into low-dimensional grain time series by information granulation of time series, and the constructed information grains can portray and reflect the structural characteristics of the original time-series data, to realize efficient dimensionality reduction and lay the foundation for the subsequent data mining work. Finally, the grains of the decision tree are analyzed, and different support vector machine classifiers corresponding to each grain are designed to construct a global multiclassification model.

1. Introduction

With the rapid development of internet technology and the improved performance of data storage devices in recent years, a large amount of data is generated and stored in various industries. Among these data, a large portion of them are time-tagged, that is, a series of observations recorded in chronological order, called time series. How to effectively analyze and process this time series of data to uncover potential and valuable knowledge and information to support more efficient production, operation, management, and decision-making activities of enterprises is one of the important tasks in today's big data era [1]. Computation is a new approach to simulating human problem-solving thinking and solving complex tasks with big data and is an emerging research direction in artificial intelligence in recent years. The main idea of the theory is too abstract and divides complex problems into several simpler problems (i.e., granulation), thus contributing to better analysis and problem-solving. The existing research on time-series information granulation is mainly divided into two aspects of research in the time axis and theoretical domain, that is, solving the problems of effective divisional representation of the time window and the theoretical domain. Support vector machines show many unique advantages in solving small sample, nonlinear, and high-form recognition problems, such as using this technology to avoid local minimum of knowledge and realize capacity control, and can be extended to other machine learning problems such as function fitting. Time-series time-axis information granulation research usually uses a fixed time interval to divide the time series, that is, hard division, and then represents the information grains obtained after the division, ignoring the changing characteristics of the time series on the time axis, which does not conform to the essential meaning of information grains, so it is necessary to design the time-series information

granulation method according to the changing characteristics of the time series on the time axis so that the information grains obtained internal structures are similar to each other and the information grains are distinct from each other [2]. Time-series domain information granulation studies usually cannot combine the requirements of both interpretability and prediction accuracy of the domain partition interval, so there is a need to design a time-series domain information granulation method that can have both strong interpretability and high prediction accuracy [3].

This paper introduces the granulation idea of granulation in time-series analysis, by granulating the time series with information, granulating the original high-dimensional time series into low-dimensional granulated time series, constructing information granules that can portray and reflect the structural characteristics of the original time-series data, thus realizing efficient dimensionality reduction and laying the foundation for subsequent data mining work, studying the time-series information granulation problem-oriented to clustering and prediction, addressing the shortcomings in existing research methods, proposing three different time-series information granulation methods in terms of both time axis and thesis domain, and applying them to stock time-series data for clustering and prediction analysis. Because of the shortcomings in existing research methods, the study of time-series information granulation for clustering and prediction proposes three different time-series information granulation methods from both time axis and theoretical domain and puts them into stock time-series data for clustering and prediction analysis. To address the problems of long training time and low efficiency of existing support vector machines for solving multiclassification problems, the idea of granular computing is introduced to construct support vector machine multiclassification models, and the learning algorithm for improving the construction of decision trees is investigated to achieve the purpose of improving its training efficiency and classification accuracy.

2. Related Work

Combining other more mature theories and methods with SVM has become a research topic with great potential for development. However, at the same time, it faces problems such as difficult classification and inaccurate prediction. The current research on granularity support vector machines mainly focuses on the combination of specific models: SVM with rough sets, decision trees, clustering, quotient spaces, association rules, and so on. These results only preprocess the data, but these models are important for the theoretical study of machine learning and support vector machines, as well as for the exploration of problems such as intelligent information processing.

Egrioglu E [4] studies rough lower approximation and rough upper approximation on the space of grain approximations from the perspective of rough set theory. Subsequently, the concept of grain-logic (G-logic) is given in the study of literature [5], where a similar inference system is built based on rough logic, while instance verification and analysis are implemented in medical diagnosis problems. Many results have also been achieved in terms of practical applications. The importance of attributes, as elaborated in the literature [6], was added to the granular computation of knowledge while used in solving the minimal attribute approximation, among others. In subsequent research, fuzzy quotient space theory was created by literature [7], improved by literature [8], perfected in the context of data mining, and so on. He Y [9] dealt with word computation and language dynamics and proposed a language dynamics system. The subsequent literature [10] elaborates a grain computation model based on tolerance relations, giving a grain operation criterion for incomplete information systems, a grain representation, and a grain decomposition method. At the same time, in connection with the attribute simplification of rough sets, the determination conditions are given, and the problems such as the acquisition of attribute necessity rules for incomplete information systems are addressed. Luo C [11] applies the compatible granularity space model in the field of image segmentation. Kim S T [12] combines granularity with neural networks and is applied to efficient knowledge discovery. Dong G [13] elaborates on the connection between concept description and concept hierarchy transformation based on the similarity of the concept lattice and granularity partitioning in the process of concept clustering. Su W H [14] grain vector space and artificial neural network, which improves the timeliness and comprehensibility of knowledge representation of the artificial neural network. Literature [15] decomposed copper and wheat prices based on EMD and EEMD methods, respectively, based on multiscale perspective analysis, and finally, BP neural network, SVM model, and ARIMA were used for prediction and integration, and the prediction results showed that the combined model prediction is better. Although the prediction model integrated by decomposition is better, there are some defects, such as the wavelet decomposition method has problems of weak adaptability of itself and poor robustness of network training in the process of data decomposition, while the EMD method has problems of modal overlap and lack of theoreticality in the decomposition process. Moreover, for price series with multiscale and high noise, the number of each component after decomposition of these methods is high, which is not conducive to the subsequent forecasting work. After that, the literature [16] constructed a new sequence decomposition method, the empirical wavelet transform method, based on wavelet transform and combined the advantages of EMD. The literature [17] and others used EWD and EWT to decompose the wind power sequences and then combined the with the neural network method for cross-combination prediction, and after comparison, it was found that the sequences decomposed by EWT had a better prediction effect. The basic idea of a rough set is to form concepts and rules through analytical induction and study target equivalence relations as well as categorical approximation knowledge discovery. Zhao Y [18] combines multilevel and multiperspective granularity methods by defining the division sequence product space and using nested division sequences to define different granular layers over the theoretical

domain. Finally, a granulation model based on the division order is given using the division order product space. Chen W [19] proposes a neighborhood granulation method by introducing inter- and intraclass thresholds to construct a supervised neighborhood-based rough set model and gives the rough approximation quality and conditional entropy monotonicity change theorems for this model by analyzing the neighborhood particle change law under double thresholds. After studying the operation mechanism of data information particles in literature [20], the nonstandard analysis is used as the operation rule of information particles; the accompanying binary relation is proposed; and the coarse and fine particle layer division of information particles in the binary relation is analyzed in-depth, and the algorithm can realize the merging and decomposition of particle layer space, which can effectively reduce the data calculation intensity and simplify the data analysis process.

3. Support Vector Machine Based Algorithms for Particle Computation

Set theory is the foundation of modern mathematics, and fuzzy set theory is one of the new mathematical tools and theories. Once the concept of fuzzy sets and the problem of granularity of fuzzy information were introduced, it rapidly expanded the scope of its use and extended the theory of fuzzy logic, followed by the "theory of word computation", which aims to use language for fuzzy computation and reasoning to achieve fuzzy intelligent control. At the same time, the integration of fuzzy set theory and quotient space, using fuzzy equivalence relations, completed the study of the expansion of the quotient space model and grain computation and was able to accurately map and solve uncertainty problems. Therefore, a proper hierarchical progressive granularity structure can solve the problem effectively. However, this theory does not have the means and technical algorithms to complete the transformation, including between granularity and granularity world, between granularity and granularity, and between granularity world and granularity world. If this problem can be solved, it will improve and promote the theory and the scope of use of the quotient space [21].

The fusion of the three models in turn produces fuzzy rough sets, fuzzy quotient spaces, and so on so that the three models are both distinct and related. First, between rough sets and fuzzy sets, the former are processed later, and the latter are preexisting. However, both describe and generalize the incompleteness and inaccuracy of information grains, and there are significant differences in the processing of information grains. Rough sets focus on the coarseness of information grains, describe grains by upper and lower approximation operators, and emphasize indistinguishability and the classification of different equivalence classes. Fuzzy sets focus on fuzziness, describe and emphasize the indistinguishability of boundaries using affiliation and affiliation functions, and study only the degree of affiliation of the same equivalence class. Figure 1 shows the framework of the algorithm flow of the support vector machine-based granular computing.

Solving for an estimate of a system sample based on a known training sample, in a dependency between the output inputs of a system, machine learning makes a relatively accurate estimate of the unknown data as possible. Then it is possible to model the problem of machine learning as the existence of some unknown dependency between input variables and output variables [22]. The basic idea of the support vector machine is to get a high-dimensional space, use nonlinearity to transform the space of the input, and then solve the optimal linear classification and finally define the appropriate inner product function to complete this nonlinear transformation. The triadic theory of granular computing includes multiperspective, multilevel granular structures, and granular computing triangles. The methodology of granular computing is a structured problem-solving, and the computational model of granular computing is structured information processing. The triad emphasizes the mutual support of the philosophical, methodological, and computational models of that computation. The study of granular computing attempts to organize, abstract, and combine granular processing ideas from various disciplines to obtain a higher-level, systematic, and discipline-specific knowledge-independent principle of granular computing.

The traditional algorithm steps are as follows:

(1) Select the number of grains to be divided p.

$$f_{\phi}(p,x) = d_p\left(x - \frac{\lambda_{\theta_1}}{d_p(x)}\right) + d_p\left(x - \frac{\lambda_{\theta_2}}{d_p(x)}\right). \quad (1)$$

$f_{\phi}(p,x)$ is the overall feature function of the data set.

(2) Determine its objective optimization function.

$$\nabla_{b^t} J(w,p;x_i^*,y) = \delta^{l+1}\frac{\partial J(w,p)}{\partial b_{ij}^{\ l}} + \lambda\alpha_{ij}^{l}. \quad (2)$$

In objective optimization, the function w is the penalty parameter, and the classified samples will appear as nonseparable regions and may also appear to belong to one class of samples or multiple classes of samples.

(3) Generate k decision functions as follows:

$$X_k = T_k + b\sum_{i=1,j=1}^{n} X_{ij}^2. \quad (3)$$

(4) The radial basis kernel function is then obtained as follows:

$$G(J) = j\frac{\partial\gamma}{\partial j} + \frac{1}{n}\sum_{i=1}^{n} X_i Y_i. \quad (4)$$

(5) The final if-then form and the main fuzzy constraint propagation rules are proposed.

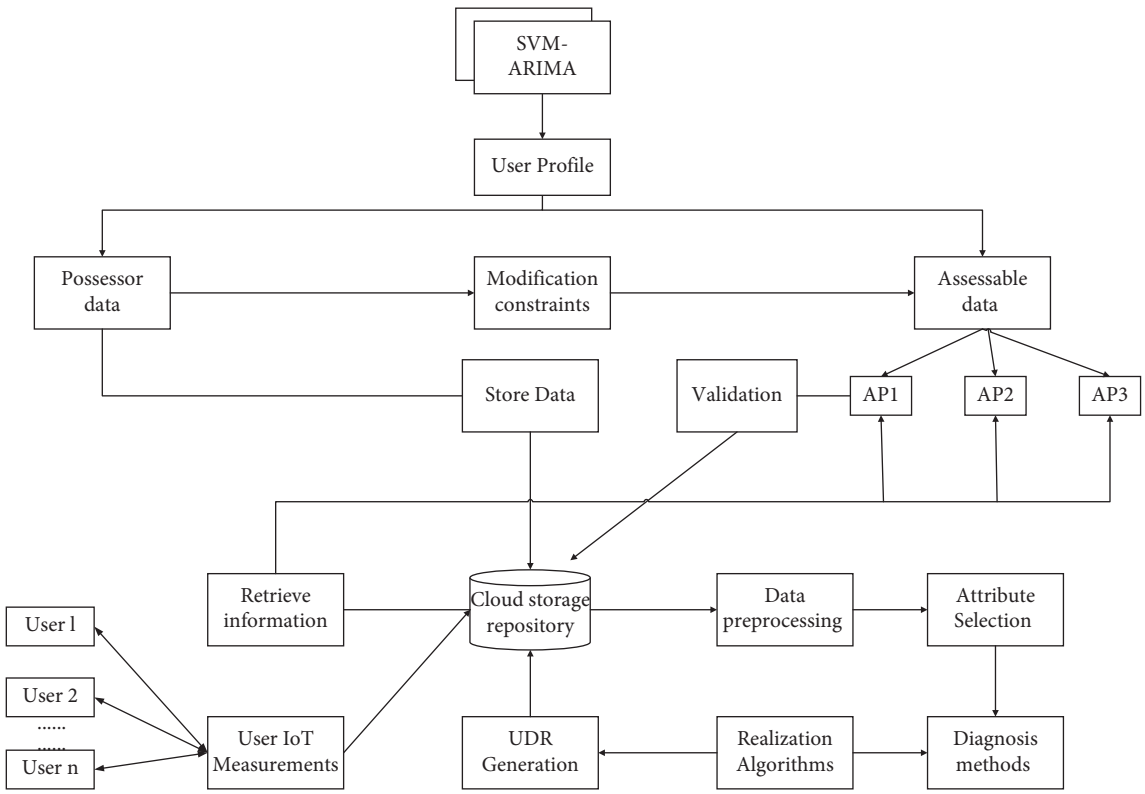


Figure 1: Support vector machine-based algorithm flow framework for granular computing.

$$T(i,j) = \frac{\overline{g}(i,j) + \overline{g}\max}{\overline{g}\max - \overline{g}\min} + |X - Y|^2. \quad (5)$$

It follows that cluster analysis can be considered as a concrete implementation of the idea of granulation, which is another layer of abstraction on the idea of cluster analysis. Granular computing is a concrete implementation of the ideas of granularity and hierarchy in the solution of machine problems. The core concept of granular computing is multilevel and multiview granular structure [23]. The fundamental framework of granular computing consists of particles, granular layers, and granular structures. The most fundamental element of granular computation is called a particle, which is composed of a collection of individuals described by internal properties and a whole of external properties. An abstracted description of the space of problems or samples of operations is called a granule layer, and the whole of particles obtained based on some required granulation criterion constitutes a granule layer of internal particles with some identical (or similar) granularity or properties. Granularity comes from the way people perceive the world. The observation, definition, and transformation of practical problems are granular calculations for different problems measured from different perspectives or levels, and in different applications, granularity can be interpreted as size, abstraction, complexity, and so on. Different grains can be ordered by their granularity. Each grain provides a local, partial description, and all grains in that layer combine to provide a global, complete description. Grain calculations are often solved on different grain layers. A multilevel grain structure is a description of the relationships and connections between grains and grains, grains and layers, and layers and layers. The grain structure is a relational structure consisting of interconnections between grain layers. There are three modes: top-down, bottom-up, and center-out, which are three common modes of information processing for humans and computer information processing.

4. Support Vector Machine and Granular Computing Based Time Series Volatility Prediction

4.1. Empirical Modal Decomposition of Time Series Fluctuation Algorithms. Empirical modal decomposition (EMD) is a signal decomposition processing method, the principle of which is to decompose the originally complex original price signal sequence into a finite number of simpler eigenmode functions (IMFs). Each IMF represents the information contained in the original price series at different scale levels, which can effectively reflect the embedded characteristics of the original price series at a low level. EMD decomposition method is a kind of data processing method that can smooth the complex signal series, which decomposes the original signal series into several component series according to the frequency level; the first component series has the highest frequency, then decreasing in order; and the last sequence has the lowest frequency so that many component sequences at different feature scale levels can be obtained, and the decomposed component sequence corresponds to an IMF eigenmode component of the relevant frequency. The

frequency fluctuations of the general sequence can indicate the degree of price series fluctuations; the higher the frequency of the component corresponding to the price series fluctuations more violent, the lower the frequency of the component corresponding to the price series fluctuations more moderate, so the last decomposition of the component out of the fluctuations of the most moderate general can represent the price series fluctuations towards, generally referred to as the residual component [24].

EMD decomposition is simple, and the decomposition results are more accurate because the process of mathematical processing is adaptive without human interference and is automatically generated decomposition results. EMD wil automatically generate different basis functions and the most appropriate number of components according to the fluctuation of different price series. In contrast, the wavelet decomposition method requires the selection of the basis function in advance when processing the original sequence and then requires several training trials to determine the most appropriate number of components. Time-series information granulation based on a support vector machine is to introduce the idea of granular computing support vector machine into time-series analysis, which is a new research direction of time-series analysis [25]. The idea of the concept of information granulation is to decompose a whole into small parts and then study the decomposed parts, and each part is divided into a particle. In other words, information granules are elements that are similar and indistinguishable or have a certain function. Information granulation of time series is the basis for compressing the scale of time-series data and using it for subsequent time-series analysis, interpretation, and modeling [26]. Therefore, compared with the wavelet decomposition method, the EMD decomposition method highlights obvious advantages in the operation of the decomposition process. However, there are some drawbacks in the application process of EMD, such as the easy occurrence of component stacking situations and endpoint contamination situations. This is because EMD needs to construct the upper and lower envelopes of the sequence using the cubic spline method in the decomposition process, but the cubic spline method will disperse near the boundary points of the original sequence, and with the decomposition of EMD, the endpoint effect will gradually spread inward to pollute the whole sequence, resulting in interference with the final decomposition effect. The simplest way to cope with this problem is to keep discarding the nonextreme part at the endpoints during the decomposition, but this will cause data waste and thus affect the latter prediction effect. If the data at the boundary points of the sequence is not deleted, generally, it is only possible to add data to each end by various methods, and this extension process will be disturbed by human factors, which will eventually affect the decomposition effect. Figure 2 shows the process of variational modal decomposition.

The empirical modal decomposition algorithm can be understood as a set of adaptive filters that sieve the data layer by layer according to its essential scale characteristics in a process where the characteristic time scales are separated in order from small to large. After such decomposition, the superimposed waves in the original signal can be removed, and symmetric modal waveforms can be obtained. In the EMD algorithm, the sequence IMF1 contains the component of the original sequence with the smallest period. The residual term after subtracting IMF1 from the original sequence contains the part of the vibration signal whose period is larger than IMF1, so the average period of IMF2 is generally larger than the average period of IMF1. By analogy, the IMF sequence filtered by the EMD algorithm has to decrease the signal frequency, decrease fluctuation intensity, and increase the average period, and the final residual term is a constant or monotonic function, which reflects the long-term trend in the sequence. The problem of unbalanced data classification has also become an important direction in the field of data mining and machine learning to study the classification problem. A better solution to the classification problem of unbalanced data distribution is required to deal with the data classification problem in a more comprehensive way. Usually, the total number of components obtained by EMD decomposition is $\log_2 N$, where N is the number of data samples in the original series. However, since the actual time series are composed of both real signals and noise, the empirical modal decomposition is processed for data containing noise, and for some time-series data with signal jump changes, the jump signals may cause scale loss, making the decomposed results have the problem of modal confounding.

When there is a jump in the scale of the original time-series signal, the decomposition result of EMD may have the problem of data modal mixing. The so-called modal overlap is manifested in the decomposition results because there should be only one scale feature in the scale feature, and the subsequence of the scale feature is not unique, and the signals of multiple scale features are mixed in the sequence. Particularly, influenced by the signal in many aspects such as collection frequency, frequency components, and signal amplitude, the phenomenon of modal blending can easily occur when empirical modal decomposition is performed directly, and modal blending mainly refers to the following two aspects:

(1) A single IMF contains components of the full heterogeneous scale

(2) Signal components of the same scale appear in different IMFs

To obtain a relatively stable speed of the vehicles within the cluster, we use the average speed of the vehicles within the cluster to characterize the stability of the cluster and filter the vehicle nodes within the above set of pairs of neighboring nodes by motion consistency to remove the vehicle nodes within the cluster with a large difference from the average speed, to ensure that the cluster can travel on the road in a relatively stable manner. Specifically, the average speed of the vehicles within the cluster at time t can be expressed as follows:

$$\overline{v_i} = \frac{\delta x}{\delta t}\left(\frac{n!}{r!\,(n-r)!}x^{\gamma} + \mu\right), \tag{6}$$

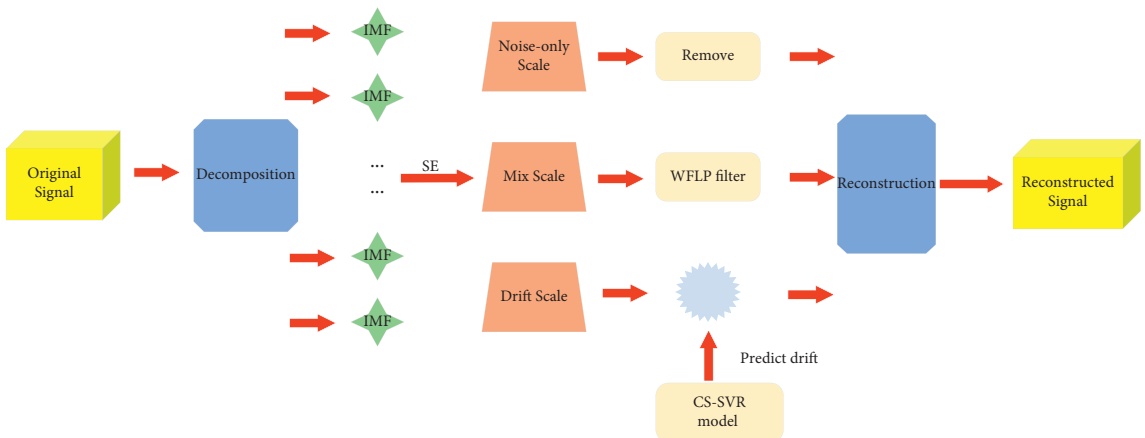


Figure 2: Variational modal decomposition flow.

where $N(t)$ denotes the number of elements in the set of neighboring nodes of V_i at time t and V_{it} represents the n-th element within the set N_{vi} of neighboring nodes of $_{Vi}$ at time t. If the velocity of V_{jn} satisfies the following equation, it will be removed:

$$N(t) = \sum_{i=1}^{n} (N_{vi} \cdot V_{in}) + A. \tag{7}$$

The set of neighbor nodes of vehicle V_i at moment t can be expressed as follows:

$$N_{vi} = \prod (V_i t + \theta) + \eta. \tag{8}$$

The average end-to-end delay reflects the effectiveness of the protocol and can be determined using the following formula:

$$\ddot{x} = \frac{\delta y}{\delta x}\left(\gamma\sqrt{y^2 + x^2} + y^3 + c\right). \tag{9}$$

When the input signal of the network is the k-th training sample X_k, its output value of the j-th neuron after the nonlinear transformation of the i-th hidden layer neuron is as follows:

$$X_k = T_k + b \sum_{i=1,j=1}^{n} X_{ij}^2. \tag{10}$$

To address the phenomenon of modal confusion in the empirical modal decomposition algorithm, Huang proposed the method of interruption detection; the specific idea is: after each decomposition, the final decomposition result is analyzed and judged, and if the phenomenon of modal confusion is found, it is filtered by selecting the appropriate interruption scale from it and then decomposed again. However, the interruption detection is a posteriori means of considering the judgment; then it may lead to the following situation: the scales potentially contained in the signal are incorrectly filtered out, and the adaptiveness of the empirical modal decomposition itself is thus greatly weakened, so in some specific cases, this interruption detection method has certain shortcomings, which will have an impact on its inspection effect. The approach in the direction of classification algorithms, that is, based on the flaws and deficiencies found in previous algorithms for solving imbalance problems, the algorithms are appropriately improved and extended to improve the ability to handle imbalance classification problems. The use of a pseudosignal has also been proposed to solve the modal aliasing problem by introducing a pseudosignal to avoid the inclusion of too wide a band in the IMF, but this method is also an approach that requires human subjective judgment to intervene and suffers from the same problem of weakened adaptivity.

4.2. Time Series Volatility Prediction Model Based on Support Vector Machine Grain Calculation. Support vector machine-based time-series information granulation is a new research direction of time-series analysis by introducing the idea of granular computing support vector machine into time-series analysis. The concept of information granulation whose idea is to decompose a whole into small parts and then study the parts obtained from the decomposition, and each part that is divided is a grain. In another way, an information grain is some elements that are similarly close and indistinguishable or have some function and are combined. Information granulation of time series is the basis for compressing the size of time-series data and using it for subsequent time-series analysis, interpretation, and modeling, and its research framework is shown in Figure 3.

When granulating information on time series, it mainly includes two steps: information grain division and information grain description. Information grain division is to divide the time series into several small subsequences, and each subsequence is called an information grain; information grain description is to construct a description method to effectively characterize the information grain on the information grain obtained from the division. Through the information granulation operation of time series, the research object can be abstracted from the low-level, fine-grained original high-dimensional time-series granulation to the high-level, coarse-grained low-dimensional grain time series, and the constructed information grain can portray and reflect the local features of the original time-series data, which achieves efficient dimensionality reduction and lays

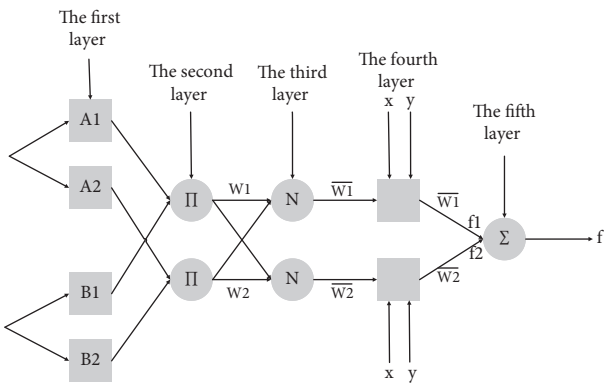


Figure 3: A research framework for granulation of time-series information.

the foundation for the subsequent data mining work. For the information granulation of time series, some scholars have conducted relevant research and achieved certain research results.

By combing the published research results, the existing research on time-series information granulation can be divided into two aspects: (1) time-axis information granulation of time series, that is, to solve the problem of an effective division of time window representation of time series, and (2) research on time series' domain information granulation, that is, to solve the problem of an effective division of time series' domain representation.

Time-axis information granulation of time series is to divide the time series into some time windows according to its change characteristics on the time axis according to some method, and the subsequence on each time window is regarded as an information grain, and then the subsequence on the divided time window is characterized effectively. The resulting interval information grain can achieve full coverage for the data samples on the time windows, as shown in Figure 4.

The theoretical domain information granulation of time series is to divide the time series into several theoretical domain intervals according to its variation characteristics on the theoretical domain according to some method, and each theoretical domain interval is regarded as an information grain, and then the divided theoretical domain intervals are characterized effectively. The research on the theoretical domain information granulation of time series is mainly divided into four types: the first is the equal interval theoretical domain division method; the second is the equal frequency theoretical domain division method; the third is the clustering-based theoretical domain division method; and the fourth is the optimization theory-based theoretical domain division method. The main research methods of time-series information granulation in terms of time axis are interval-based time-axis information granulation, clustering-based time-axis information granulation, and fuzzy set-based time-axis information granulation. These methods usually use a fixed time interval to divide the time series, that is, hard division, and then represent the subsequence (information grain) obtained after the division, ignoring the

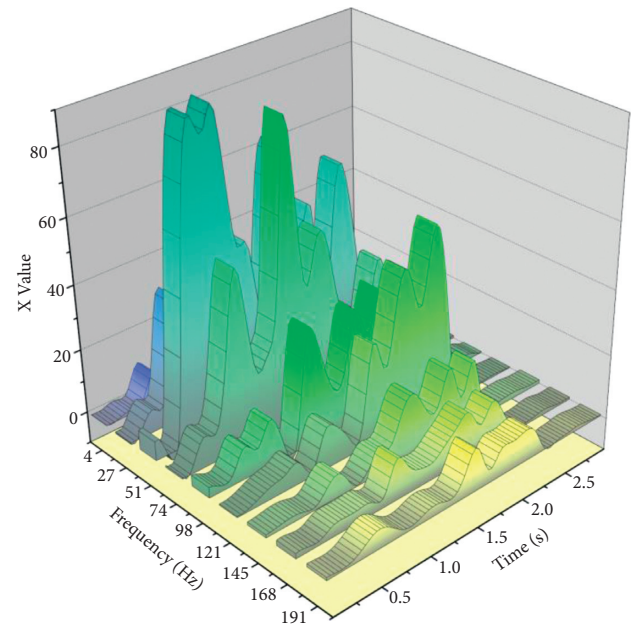


Figure 4: Interval information granulation of time series.

change characteristics (structural characteristics) of the time series on the time axis, which does not conform to the essential meaning of information grain. The problem of unbalanced data classification is a key problem in the field of machine learning and data mining. From a realistic point of view, the distribution of data sets in a large number of classification problems is unbalanced, and the importance of each category is different; usually, sparse categories of data are more worthy of study in a particular context. Therefore, it is necessary to design the information granulation method of time series according to the changing characteristics of time series on the time axis so that the obtained information grains have similar internal structures among themselves and distinguishable information grains from each other. Real-life time-series data are usually characterized by high dimensionality and high noise, so it is crucial to effectively perform information granulation operations on time series to reduce the data size of time series and reduce the impact of noise.

The information granulation operation on time series in terms of the time axis is essentially the same as the traditional time-series dimensionality reduction representation method; both are to reasonably compress the data size while keeping the important features of the original time series as much as possible. However, the traditional time-series dimensionality reduction representation method does not compress the time series to a high degree and does not reflect the structural characteristics of the time series well, thus affecting the effectiveness of the subsequent analysis work. The granular time series obtained after the information granulation operation cannot directly participate in the data mining work of time series yet and need to combine the characteristics of the information granular to propose the corresponding similarity measure before the subsequent analysis calculation. The most important decision tree in the

multiclassification problem is established by using the combination of grain calculation and Huffman tree, using its feature of the shortest length of the path with weight so that it can be attributed to the category in the shortest time, constructing the Huffman model of multiclassification, analyzing the grains of the decision tree, and using the granularity and decision tree to construct different multi-classifiers corresponding to each grain. Finally, the global model is constructed. In solving the multiclassification problem, the use of grain computing, time series, and support vector machines are combined, which not only inherit the advantages of each but also complement the disadvantages of each, and get the synergistic enhancement effect.

5. Experimental Verification and Conclusions

5.1. Time Series Domain Division. In multiclassification of textual problems, the huge amount of information is a major problem in solving multiclassification. First, the problem is granularized, and in the text background knowledge, the text content is categorized as environment, computer, transportation, education, economy, military, sports, medicine, art, politics, and so on. Then the different disciplines are considered as particles in the multilevel granular structure, and all these categories belong to the same granular layer. The particles in the upper granular layer of these particles are combinations with similar particle size characteristics, and these are coarse particles of the lower layer, and the lower layer is fine particles of the upper layer. The weights processed sum to 1. For better operation, the weights are multiplied by 1,000 times to become integers. And use VC++ 6.0 software programming to construct the decision tree, function CrtHuffTree (Huffnode ht[] and int n) to implement the tree construction, and function code (Huffnode ht[], Huffcode hcd[], Huffcode*ss, and int n) to implement the encoding. From Figure 5, it can be seen that the subintervals of the theoretical domain obtained based on the support vector machine class method are consistent with the distribution characteristics of the data, that is, the subinterval interval divided is smaller in regions with dense data distribution and larger in regions with sparse data distribution.

However, there is a great difference in the amount of data contained in the subintervals of the theoretical domain obtained by this method, that is, the amount of data contained in the subintervals with dense data distribution is large and the amount of data contained in the subintervals with sparse data distribution is small, so the subintervals are optimized using the information granulation method below. From the decomposition results, the multiscale characteristics of the original time series can be initially found: from IMF1 to IMF10, the IMF component gradually changes from high-frequency vibrations to low-frequency vibrations; the series with higher vibration frequencies, compared with the original series, remains the same in terms of fluctuation frequency, and there are some differences in terms of amplitude, which represent the short-term effects triggered by the normal fluctuations of the securities market and the occurrence of irregular events; the vibration series with

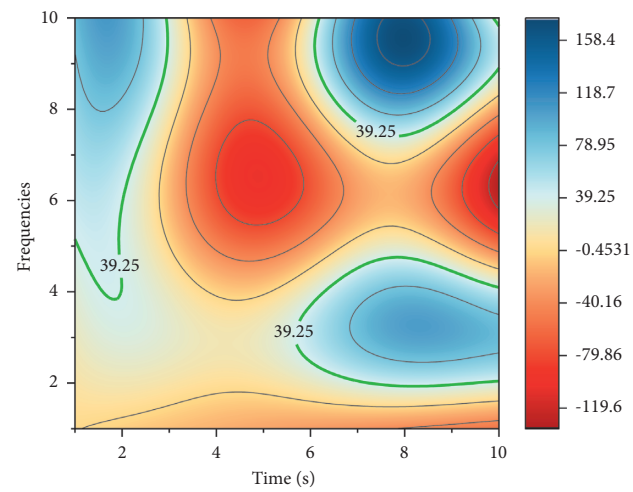


Figure 5: Time-series domain partitioning based on support vector machine class approach.

lower frequencies vary relatively flat, and each change in the series represents a long-term impact triggered by a major event; and the original series always fluctuates up and down around the residual term and shows an increasing trend in the long run.

5.2. Accuracy Comparison between Different Models. According to the results reflected in the volume-price relationship model, the model regression with the volume-price relationship as input is better, and the prediction effect is better for the more volatile prediction, and the volatility of the volume-price relationship has a closer relationship with the volatility of the closing price of the stock. For different step sizes, the loss function of the training set decreases and finally converges, while the loss function of the test set with step sizes 10 and 30 has a better effect, and the loss function of the test set with step size 50 decreases and converges poorly. According to the final mean absolute error MAE, the MAE of the training set with different step sizes is close to 0.04, and the best performance of the model with 30 step sizes is 0.011784855. Refining the classification decision for local neighborhood data distribution, adjusting the posterior probability estimates and using rough set approximation theory to handle extreme distribution cases, eliminates the uncertainty of lacking rare class data. After the reclassification decision based on the refined instance distribution, the dynamic mean-neighbor classification algorithm based on neighborhood rough sets can classify query instances into classes more accurately.

In terms of training time, the longer the step size window, the longer the training time. Based on the above conclusion, in the next model testing process, it was decided to choose a window with a step size of 30 for testing. According to the loss function change image of the technical indicator model, the loss function of both the training and test sets are under and finally converge. According to the results reflected by the technical indicator model, the loss function MAE of the test set with technical indicators as

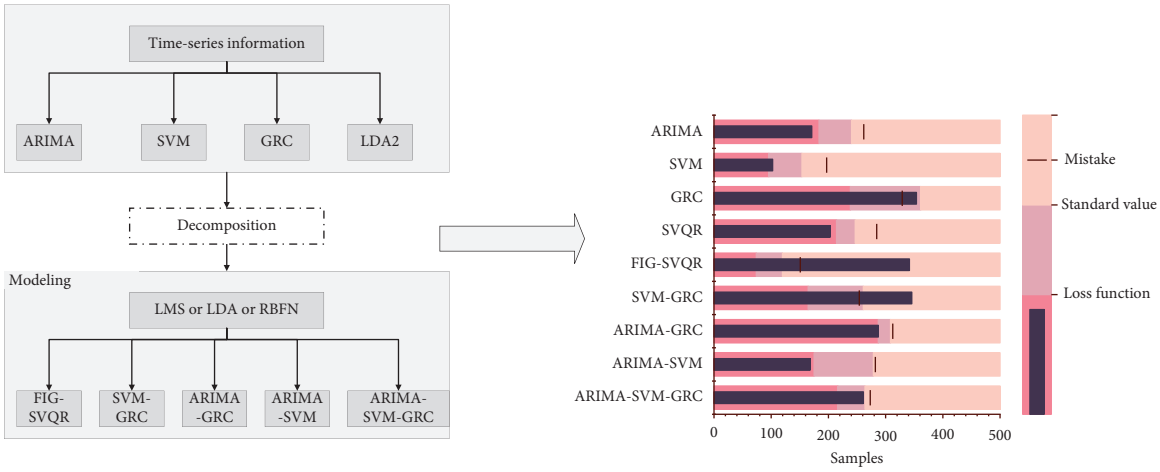


Figure 6: Variation of loss function values for each model.

input is 0.01071587, which is slightly smaller than that of the volume-price relationship model, and the overall regression result is slightly better than that of the volume-price relationship model.

However, based on the observation of the fitted curves, the technical indicator model has a better prediction for more volatile and weaker volatility realizations for less volatile. The ARIMA-SVM-GRC method is applied to reduce the normalized volume-price relationship and technical indicator data to extract the data features of volume-price relationship and technical indicators and reduce the six-dimensional volume-price relationship data to three-dimensional and technical indicator data to three-dimensional. According to the results of the principal component ratio, the first principal component of the volume-price relationship accounts for the largest share incorporates all the information of the sample, and the first and second principal components of the technical indicators account for about 50%. The change in the value of the loss function of each model is shown in Figure 6.

As shown in Figure 7, the ARIMA-SVM-GRC method increases the validity of the prediction results by performing deterministic prediction of runoff while also obtaining uncertain prediction results from experimental data. Usually, the prediction interval can reflect the fluctuation of runoff, and the ARIMA-SVM-GRC method can handle discrete and nonlinear relationships. To further demonstrate the predictive power of the proposed ARIMA-SVM-GRC method, we compare the ARIMA-SVM-GRC method with other traditional data-driven methods in a cross-sectional manner.

In this section, we conduct comparison experiments with BP, RBF, and SVQR methods. In the cross-sectional comparison analysis of these methods in terms of both point prediction and interval prediction, the BP and RBF algorithms have poor results for runoff prediction, with an average absolute percentage error above 6%, while SVQR and ARIMA-SVM-GRC have smaller average absolute percentage error and both below 4%, with the plurality-based SVQR method having the best average absolute percentage error of 2.7%. For the relative mean squared error, both BP and RBF results are above 3%, while both SVQR and ARIMA-SVM-GRC methods are stable below 2%. The mean absolute error is greater than 30 for both BP and RBF methods and below 30 for SVQR and ARIMA-SVM-GRC methods.

For the interval prediction results, we compared the SVQR method and the ARIMA-SVM-GRC method, and the coverage probabilities of both methods were above 95%, and the bandwidths were less than 30%, indicating that both the SVQR method and the ARIMA-SVM-GRC method could better reflect the interval prediction results, and the FIG-SVQR method improved the coverage probability by 0.59% compared with the SVQR method and reduced the coverage probability by 0.59%. The FIG-SVQR method improves the coverage probability by 0.59% and reduces the bandwidth by 0.99% compared with the SVQR method, which indicates that the ARIMA-SVM-GRC method proposed in this section is significantly better than the SVQR method in terms of interval prediction.

5.3. Comparison of Model Error Scenarios. According to the model error, the prediction error of the decomposition integration model is lower than the prediction error of the model under multiple factors, but the overall deviation situation is not significant. Moreover, the decomposition integration model can improve the prediction accuracy of the neural network model as much as possible. So, when the difference between the prediction results of the multifactor under model and the decomposition integration model under time series is applied to solve the more complex problem of fluctuation is not large, it is necessary to consider the prediction results of the reference two models together. When the difference between the forecasts of the two models is large, the model with the smaller error can be considered for forecasting. From the comparison in Figure 8, it can be seen that CMS takes the longest time, which is mainly because CMS needs to consider both intra- and interattribute similarity in the similarity of two objects, and from the introduction in Section 3, it can be seen that the algorithm has a high time complexity and therefore takes the longest

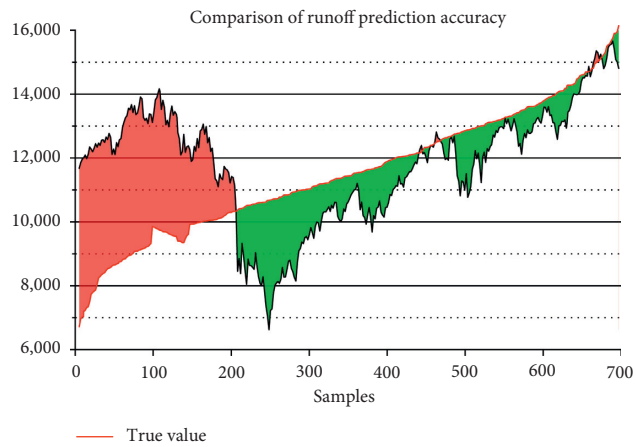


Figure 7: Comparison of runoff prediction accuracy.

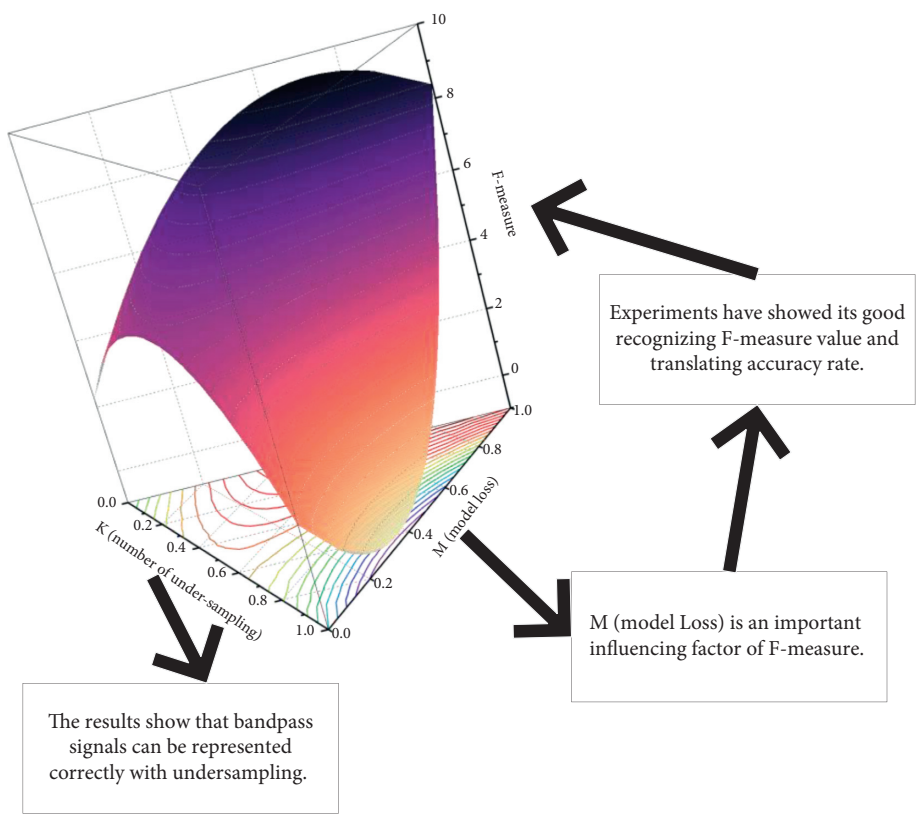


Figure 8: Comparison of model error scenarios.

time. HM, OF, IOF, Eskin, and k-modes algorithms belong to the more classical algorithms.

And after algorithm optimization, the running time of the ARIMA-SVM-GRC algorithm mainly depends on the baseline scale clustering results and the scale conversion method and is affected by the number of clusters in the baseline scale clustering results; the more the number of clusters obtained, the longer the time required for scale up-projection, and the less the number of clusters, the shorter the time required for scale up-projection, independent of the original data set independent of the size of the set. Since the number of clusters obtained from the benchmark scale clustering is much smaller than the sample size of the original data set, the ARIMA-SVM-GRC algorithm requires much less running time than other comparison algorithms. According to the dynamic mean query neighborhood, a more scientific as well as more rigorous method is needed to calculate the local confidence intervals and global confidence intervals in the query neighborhood to determine the actual situation of the sample distribution of the minority category data in the neighborhood. Since the algorithm running time is affected by the experimental environment and the degree

of algorithm optimization, the CMS and Eskin methods in the experimental algorithm are obtained according to the formula reduction in the journal without algorithm optimization, so the running time is relatively long.

6. Conclusions

With the development of technology and network, multiclassification problems occupy an increasingly important position in people's lives. The multiclassification problem is becoming more and more difficult to solve along with the increase of disturbing factors and the massive amount of data.

By studying the multiclassification problem, a support vector machine multiclassification model based on granular computing is proposed and combined with a time-series fluctuation prediction model to analyze and deal with the multiclassification problem. In this paper, granular computation is incorporated into the data preprocessing model. The problems of large data processing and low training speed can be effectively solved by using ideas such as granulation and hierarchy of granular computing triad, which are analyzed and transformed from the perspective of granular computing. In the multiclassification problem, whether the decision tree is constructed reasonably is crucial, using the granularity of granular computing combined with the Huffman tree to construct the optimal decision binomial tree, so that it can obtain the classification in the shortest time and solve the problem of uneven samples within the class, which provides a practical method for the analysis of multiclassification problems. In terms of the time axis of time series, the method of granulating time-series information based on fluctuation points is proposed for the structural characteristics of low-frequency time series. The key to the research of this method is the definition and identification of fluctuation points in the time series, identifying fluctuation points by operating on the original time series, dividing the information grain of the original time series using the fluctuation points, and then describing the information grain after the division using a linear function to complete the information granulation operation of the time series and transforming the original time series into a grain time series. Since the number of information grains and the corresponding time window size in the granular time series are different, a new similarity measure based on linear information granulation is proposed to facilitate the subsequent data mining work on the granular time series. Firstly, to ensure the one-to-one correspondence between linear information grains of different time series, the segmentation matching algorithm of linear information grains is proposed; secondly, for the matched linear information grains, the corresponding similarity metric algorithm is proposed.

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 in this article.

Acknowledgments

This work was supported by the projects of Ningxia Natural Science Foundation (No. 2022AAC03315, time-series analysis and method research based on granular computing; No. 2022AAC03328, research on vegetable supply chain traceability based on combined RFID technology and service-oriented architecture (SOA); No. 2021AAC03235, fusion method of interest target detection in low illumination visible image and infrared image; No. 2022AAC03314, high-precision numerical simulation for the incompressible magnetohydrodynamic problems; and No. 2022AAC03301, high-order difference schemes on an adaptive algorithm for convection-diffusion reaction equations).

References

[1] M.-H. Fan, M.-Y. Chen, and E.-C. Liao, "A deep learning approach for financial market prediction: utilization of google trends and keywords," *Granular Computing*, vol. 6, no. 1, pp. 207–216, 2021.

[2] E. Bas, U. Yolcu, and E. Egrioglu, "Intuitionistic fuzzy time series functions approach for time series forecasting," *Granular Computing*, vol. 6, no. 3, pp. 619–629, 2021.

[3] T. Wang, T. Ma, D. Yan et al., "Prediction of heating load fluctuation based on fuzzy information granulation and support vector machine," *Thermal Science*, vol. 25, no. 5, pp. 3219–3228, 2021.

[4] E. Egrioglu, U. Yolcu, and E. Bas, "Intuitionistic high-order fuzzy time series forecasting method based on pi-sigma artificial neural networks trained by artificial bee colony," *Granular Computing*, vol. 4, no. 4, pp. 639–654, 2019.

[5] Z. Han, J. Zhao, H. Leung, K. F. Ma, and W. Wang, "A review of deep learning models for time series prediction," *IEEE Sensors Journal*, vol. 21, no. 6, pp. 7833–7848, 2019.

[6] M. Bose and K. Mali, "Designing fuzzy time series forecasting models: a survey," *International Journal of Approximate Reasoning*, vol. 111, pp. 78–99, 2019.

[7] D. Yu, Z. Xu, and X. Wang, "Bibliometric analysis of support vector machines research trend: a case study in China," *International Journal of Machine Learning and Cybernetics*, vol. 11, no. 3, pp. 715–728, 2020.

[8] H. Wu, H. Long, Y. Wang, and Y. Wang, "Stock index forecasting: a new fuzzy time series forecasting method," *Journal of Forecasting*, vol. 40, no. 4, pp. 653–666, 2021.

[9] Y. He, Y. Yan, and Q. Xu, "Wind and solar power probability density prediction via fuzzy information granulation and support vector quantile regression," *International Journal of Electrical Power & Energy Systems*, vol. 113, pp. 515–527, 2019.

[10] K. Mateńczuk, A. Kozina, A. Markowska et al., "Financial time series forecasting: comparison of traditional and spiking neural networks," *Procedia Computer Science*, vol. 192, pp. 5023–5029, 2021.

[11] C. Luo, C. Tan, and Y. Zheng, "Long-term prediction of time series based on stepwise linear division algorithm and time-variant zonary fuzzy information granules," *International Journal of Approximate Reasoning*, vol. 108, pp. 38–61, 2019.

[12] S.-T. Kim, I.-H. Choi, and H. Li, "Identification of multi-concentration aromatic fragrances with electronic nose technology using a support vector machine," *Analytical Methods*, vol. 13, no. 40, pp. 4710–4717, 2021.

[13] G. Dong, R. Li, J. Jiang, H. Wu, and S. C. McClure, "Multigranular wavelet decomposition-based support vector regression and moving average method for service-time prediction on web map service platforms," *IEEE Systems Journal*, vol. 14, no. 3, pp. 3653–3664, 2019.

[14] W. H. Su, S. Bakalis, and D. W. Sun, "Chemometric determination of time series moisture in both potato and sweet potato tubers during hot air and microwave drying using near/mid-infrared (NIR/MIR) hyperspectral techniques," *Drying Technology*, vol. 38, pp. 806–823, 2020.

[15] M. Askari and F. Keynia, "Mid-term electricity load forecasting by a new composite method based on optimal learning MLP algorithm," *IET Generation, Transmission & Distribution*, vol. 14, no. 5, pp. 845–852, 2020.

[16] A. R. S. Parmezan, V. M. A. Souza, and G. E. A. P. A. Batista, "Evaluation of statistical and machine learning models for time series prediction: identifying the state-of-the-art and the best conditions for the use of each model," *Information Sciences*, vol. 484, pp. 302–337, 2019.

[17] C. H. Fajardo-Toro, J. Mula, and R. Poler, "Adaptive and hybrid forecasting models—a review," *Engineering Digital Transformation*, Springer, Berlin, Germany, 2019.

[18] Y. Zhao, T. Li, and C. Luo, "Spatial-temporal fuzzy information granules for time series forecasting," *Soft Computing*, vol. 25, no. 3, pp. 1963–1981, 2021.

[19] W. Chen, M. Jiang, W.-G. Zhang, and Z. Chen, "A novel graph convolutional feature based convolutional neural network for stock trend prediction," *Information Sciences*, vol. 556, pp. 67–94, 2021.

[20] T. Ouyang, Y. He, H. Li, Z. Sun, and S. Baek, "Modeling and forecasting short-term power load with copula model and deep belief network," *IEEE Transactions on Emerging Topics in Computational Intelligence*, vol. 3, no. 2, pp. 127–136, 2019.

[21] P. E. Puspita, T. İnkaya, and M. Akansel, "Clustering-based sales forecasting in a forklift distributor," *International Journal of Engineering Research and Development*, vol. 11, no. 1, pp. 25–40, 2019.

[22] P. S. G. de Mattos Neto, J. F. L. de Oliveira, D. S.D. O. Santos Júnior, H. V. Siqueira, M. H. N. Marinho, and F. Madeiro, "An adaptive hybrid system using deep learning for wind speed forecasting," *Information Sciences*, vol. 581, pp. 495–514, 2021.

[23] A. A. Nasser, M. Z. Rashad, and S. E. Hussein, "A two-layer water demand prediction system in urban areas based on micro-services and LSTM neural networks," *IEEE Access*, vol. 8, pp. 147647–147661, 2020.

[24] K. Singh, R. Tiwari, P. Johri, and A. A. Elngar, "Feature selection and hyper-parameter tuning technique using neural network for stock market prediction," *Journal of Information Technology Management*, vol. 12, pp. 89–108, 2020.

[25] W. Wang, W. Liu, and H. Chen, "Information granules-based BP neural network for long-term prediction of time series," *IEEE Transactions on Fuzzy Systems*, vol. 29, no. 10, pp. 2975–2987, 2020.

[26] Z. Nannan and L. Chao, "Adaptive online time series prediction based on a novel dynamic fuzzy cognitive map," *Journal of Intelligent & Fuzzy Systems*, vol. 36, no. 6, pp. 5291–5303, 2019.

Hindawi
Journal of Robotics
Volume 2022, Article ID 9742815, 11 pages
https://doi.org/10.1155/2022/9742815

Research Article

Fault Diagnosis Method of Distribution Equipment Based on Hybrid Model of Robot and Deep Learning

Shan Rongrong,[1] Ma Zhenyu,[2] Ye Hong,[3] Lin Zhenxing,[3] Qiu Gongming,[1] Ge Chengyu,[1] Lu Yang,[1] and Yu Kun[1]

[1]*NARI Group Co., Ltd, State Grid Electric Power Research Institute, Nanjing 210000, China*
[2]*Zhejiang Electric Power Corporation, Hangzhou 310013, China*
[3]*State Grid Wenzhou Power Supply Company Ouhai Power Supply Branch, Wenzhou 325000, China*

Correspondence should be addressed to Shan Rongrong; shanrongrong@sgepri.sgcc.com.cn

Received 8 February 2022; Revised 14 March 2022; Accepted 19 March 2022; Published 14 April 2022

Academic Editor: Shan Zhong

In view of the poor effect of most fault diagnosis methods on the intelligent recognition of equipment images, a fault diagnosis method of distribution equipment based on the hybrid model of robot and deep learning is proposed to reduce the dependence on manpower and realize efficient intelligent diagnosis. Firstly, the robot is used to collect the on-site state images of distribution equipment to build the image information database of distribution equipment. At the same time, the robot background is used as the comprehensive database data analysis platform to optimize the sample quality of the database. Then, the massive infrared images are segmented based on chroma saturation brightness space to distinguish the defective equipment images, and the defective equipment areas are extracted from the images by OTSU method. Finally, the residual network is used to improve the region-based fully convolutional networks (R-FCN) algorithm, and the improved R-FCN algorithm trained by the online hard example mining method is used for fault feature learning. The fault type, grade, and location of distribution equipment are obtained through fault criterion analysis. The experimental analysis of the proposed method based on PyTorch platform shows that the fault diagnosis time and accuracy are about 5.5 s and 92.06%, respectively, which are better than other comparison methods and provide a certain theoretical basis for the automatic diagnosis of power grid equipment.

1. Introduction

With the continuous improvement of social economy and people's living standards, the power demand is increasing day by day, and the scale of power system is growing day by day. It includes transmission and transformation networks of various voltage levels. Therefore, ensuring the safe and stable operation of complex power grid is an inevitable requirement to ensure the economic and social development. At the same time, the economic and social development has a great impact on the security and economy, and higher reliability is required [1, 2]. As the terminal of the whole power grid operation, distribution network is the part with the widest coverage and the largest scale in China's power system, and it is the key link to ensure that power can be "allocated and used" [3]. The distribution equipment will have some faults under long-term operation, resulting in abnormal temperature. Therefore, by detecting the temperature of the distribution equipment, the thermal fault diagnosis of the distribution equipment can be carried out quickly, which plays a great role in the safe operation of the power grid.

The infrared image of equipment is used for fault diagnosis with high efficiency, accurate judgment, safety, and reliability. At the same time, it is free from electromagnetic interference, fast detection speed, and no power failure of live equipment. Therefore, infrared diagnosis is widely used in the field of equipment fault monitoring and diagnosis technology [4]. However, due to the characteristics of large quantity and complex types of distribution equipment, if we

only rely on manual work in the process of data acquisition, analysis, and processing, the workload is relatively large, the efficiency is low, and the accuracy is relatively low due to the high dependence on manual experience. Therefore, automatic image acquisition and analysis of distribution equipment are of great significance to ensure the safety and stability of distribution network [5].

In recent years, the automatic inspection technology of power distribution room has been popularized, and various automatic robots and UAVs have made great progress in the original data acquisition stage. However, the accurate and efficient processing of collected image data is still in its infancy. How to extract the features of interest from infrared images for power distribution equipment recognition is a problem to be solved [6]. Among them, the deep learning algorithm has made great achievements in image processing, speech recognition, and text analysis. By establishing a deep-seated neural network, high-level features are extracted from low-level features layer by layer, so as to achieve the effect of target classification and recognition [7]. Compared with the manually designed feature extraction method, the distributed features obtained by deep learning network model can better express the essence of data [8, 9]. Therefore, in order to improve the efficiency of thermal fault detection of distribution equipment, improve the intelligence of power grid, reduce the labor cost of detection, and reduce the false detection rate, a fault diagnosis method of distribution equipment based on the hybrid model of robot and deep learning is proposed, which effectively ensures the safe and reliable operation of distribution equipment.

2. Related Research

At present, there are many researches on fault diagnosis of distribution equipment at home and abroad, which can be divided into traditional fault identification and classification methods and machine learning based identification and classification methods [10]. Among them, the traditional fault identification and classification methods mainly include fuzzy clustering, discrete wavelet transform, and chaotic algorithm. For example, [11] proposed an infrared image segmentation algorithm based on intuitionistic fuzzy clustering algorithm based on spatial distribution information, which is suitable for power equipment. It can well suppress the strong interference of nontarget objects in infrared image to image segmentation, but the method is more traditional and has poor segmentation effect for complex intelligent power grid equipment. Reference [12] proposed an anomaly detection method based on spatial clustering applied by auxiliary feature vector and density noise. The auxiliary feature vector of each conditional variable is constructed for clustering to identify normal data patterns and different types of anomalies. Reference [13] proposed a data mining driving scheme based on discrete wavelet transform to realize high impedance fault detection in active distribution network, but the universality of the method is not high. Reference [14] proposed a method to obtain the vibration characteristics of circuit breaker based on time-frequency and chaos analysis to realize circuit breaker fault identification. The image analysis process is complex, resulting in the reduction of fault identification efficiency. Reference [15] proposes a method based on feature model for single-phase grounding fault in active distribution network system, which transforms the solution of nonlinear feature model into single-objective optimization of feature entropy, which can well identify single-phase fault, but the identification effect of equipment with feature type is not ideal.

With the continuous development of computer technology and the rapid development of 5G communication technology, machine learning algorithm has been widely used this year, especially the deep learning algorithm has certain advantages in the field of fault identification and classification. Reference [16] proposes artificial neural network algorithm to identify the insulator state and uses single-layer and multilayer perceptron artificial intelligence algorithm to classify the conditions of distribution insulators. This technology can make the automatic inspection of electrical system more accurate and efficient, but it lacks high reliable database for support. Reference [17] proposed a Mask *R* convolution neural network method and used transfer learning and dynamic learning rate algorithm to realize efficient recognition of annotated image data sets, but it relied too much on graphics annotation and lacked practical application value. In [18], appropriate traveling wave time-frequency characteristic parameters of fault current are selected as the input of adaptive depth belief network model to obtain the fault type, but only considering the fault current characteristics as the basis, the reliability needs to be further improved.

Based on the above analysis, aiming at the problems such as the complexity and diversity of smart grid distribution equipment and the unsatisfactory effect of most existing image recognition methods, a distribution equipment fault diagnosis method based on robot and deep learning hybrid model is proposed. Its innovations are summarized as follows:

(1) In order to obtain the image information of distribution equipment more comprehensively, the proposed method introduces the robot to construct the corresponding image knowledge database, which provides the basis for fault classification and fault location.

(2) In order to locate the equipment defect area in the infrared image of distribution equipment, the proposed method performs threshold segmentation on the infrared image in hue saturation value (HSV) space and uses OTSU method to extract the equipment defect area, so as to improve the accuracy of subsequent fault diagnosis.

(3) Aiming at the problem that the deep learning algorithm is prone to gradient disappearance and gradient explosion, the proposed method uses the residual network to improve the region-based fully convolutional networks (R-FCN) algorithm and applies it to the learning of faulty equipment, so as to

obtain the fault type and location with high accuracy and further improve the safety of equipment.

3. Proposed Method

3.1. Construction of Image Information Database of Distribution Equipment Based on Robot Inspection. The traditional equipment status is usually determined by manual analysis. The workload is huge and error-prone, which affects the judgment of system status, resulting in potential safety hazards. Therefore, the robot is used for patrol inspection to obtain the status image of distribution equipment and build the corresponding information base for the analysis of equipment status, so as to find the faulty equipment in time and ensure the reliable operation of power grid [19]. The construction process of distribution equipment image information base based on robot inspection is shown in Figure 1.

The basic data sources of the database mainly include production system, online monitoring system, and robot background inspection system. The relevant data of the state quantity of power equipment mainly comes from the power production management system (PMS), which can provide the real-time operation condition, historical operation state, historical maintenance record, historical test data, equipment account, equipment parameters, and other information of the equipment. The online monitoring system mainly relies on various sensors on each power equipment for real-time monitoring. The robot background inspection system can not only provide the observation of some state quantities, but also carry out corresponding state evaluation and analysis for different equipment states according to the automatic state evaluation system. In addition, the data composition of the system includes infrared temperature measurement, visible light reading, and telemetry reading.

The robot inspection cycle generally refers to the inspection plan formulated by the distribution network operation inspection center, and two inspection robots complete the tasks of infrared temperature measurement and data transcription of equipment in the area [20]. At the same time, the robot background uses the threshold out of limit judgment method to automatically evaluate the equipment status. In order to ensure sufficient charging time of the robot and avoid the daily patrol and infrared temperature measurement period, the special patrol at night is set in the nonbusy working period of the robot every day, with the upper limit of one time. The data reports collected by the special patrol at night and infrared temperature measurement are included in the database for screening and preprocessing.

In addition, the background of the inspection robot is equipped with a system server, which includes data analysis software terminal, data exchange server, data storage server, data operation server, and other modules. The data exchange server is responsible for collecting and classifying the production system, online monitoring system, and robot patrol data into the storage server. There are three-party databases, fault information base and knowledge information base in the data storage server.

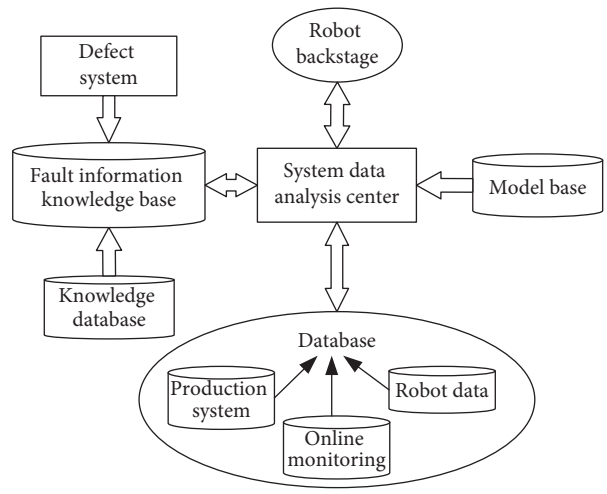


Figure 1: Construction process of image information database for distribution equipment.

(1) The data of the three-party platform includes the information required in the database structure table. After eliminating the redundant information, the integrated data in a unified format can be obtained, and the defect alarm data can be located and retrieved quickly.

(2) The fault information base is mainly taken from the defect system records and contains a large number of relevant equipment fault cases, including fault characteristics, solutions, expert opinions, and manufacturer records. At the same time, the maintenance record database and equipment account database are used to build a comprehensive database of fault information, so as to screen the fault inspection points.

(3) The knowledge information base is the engine for the system to evaluate the equipment status and judge the fault. The internal rules at all levels provide the logical basis for the system to judge the fault. The key is knowledge acquisition, that is, collecting and mining the knowledge at all levels to enrich the knowledge base.

3.2. Defect Feature Extraction of Distribution Equipment. When extracting the defect features of distribution equipment, it is necessary to perform threshold segmentation on the infrared image in HSV space, separate the infrared image background from irrelevant equipment and defective equipment, and then extract the equipment defect area [21].

3.2.1. OTSU Threshold Segmentation. OTSU is considered to be one of the best algorithms in image threshold segmentation. The threshold segmentation process of OTSU algorithm is as follows: firstly, the image is processed in gray level, the number of pixels in the whole image is counted, and the probability distribution of each pixel in the whole image is calculated; then, the gray level is traversed and

searched in the whole image, and the interclass probability of the image foreground and background at the current gray level is calculated; finally, the threshold corresponding to the variance between classes and within classes is calculated by the given objective function.

Suppose there are D gray levels in the image, in which the number of pixels with gray value of i is N_i and the total number of pixels in the image is N. Then, the average gray value of the whole image is

$$\mu_{\sum} = \sum_{i=0}^{D-1} i \frac{N_i}{N}. \tag{1}$$

According to the gray characteristics of the image, the image is divided into foreground B_0 and background B_1. Then, $p_0(T)$ and $p_1(T)$ represent the probability of occurrence of foreground B_0 and background B_1 when the threshold is T, respectively. The calculation is as follows:

$$\begin{aligned} p_0(T) &= \sum_{i=0}^{T} \left(\frac{N_i}{N}\right), \\ p_1(T) &= 1 - p_0(T). \end{aligned} \tag{2}$$

Then, the mean values of foreground B_0 and background B_1 are

$$\begin{cases} \mu_0(T) = \dfrac{\sum_{i=0}^{T} i(N_i/N)}{p_0(T)}, \\ \mu_1(T) = \dfrac{\mu_{\sum} - \sum_{i=0}^{T} i(N_i/N)}{p_1(T)}. \end{cases} \tag{3}$$

The interclass variance with threshold T in the gray histogram is calculated as follows:

$$\sigma_B^2(T) = p_0(T)\left[\mu_0(T) - \mu_{\sum}\right]^2 + p_1(T)\left[\mu_1(T) - \mu_{\sum}\right]^2. \tag{4}$$

The optimal threshold is defined as the T value corresponding to the maximum variance between classes, which is calculated as

$$\sigma_B^2(\widehat{T}) = \max_{0 \le T \le D-1} \left\{\sigma_B^2(T)\right\}. \tag{5}$$

3.2.2. Defect Region Extraction Based on HSV. In order to improve the accuracy of equipment fault image classification, the defect region and background in the infrared image of fault power equipment are separated by using the defect region segmentation algorithm based on HSV. Since it is impossible to determine the defect type only by analyzing the fault area, it is necessary to segment the defect area based on mathematical morphology according to the location of the defect area. Through this method, the defective power equipment and the background area in the infrared image are separated, so as to reduce the interference of the background area in the infrared image on the defect type classification [22]. The flow of HSV based defect region extraction algorithm is shown in Figure 2.

When processing the infrared image of defective equipment, first merge the similar pixels corresponding to the area with the same temperature, and segment the image according to the threshold of the three components of the defective area in the HSV color space to extract the defective area. Then, the discrete defect regions are connected through the closed operation in mathematical morphology, and the threshold segmentation of the original image is carried out by OTSU method to separate the power equipment and the background region. Finally, the defect area is found in the binary image separated by OTSU method; that is, the defective power equipment is separated from other areas, so as to achieve the purpose of extracting defective power equipment and facilitate the identification and diagnosis of power equipment types and fault types.

3.3. Fault Diagnosis of Distribution Equipment Based on Deep Learning Hybrid Model

3.3.1. Defect Training Based on Deep Learning Hybrid Model. R-FCN algorithm architecture mainly includes backbone network, region proposal network (RPN), and region of interest (ROI) subnet [23]. When fault diagnosis of power distribution equipment is carried out, first input the collected infrared image of power equipment into convolution neural network and extract the convolution feature map of infrared image. In this process, deeper and more abstract image features can be extracted by using a larger backbone network (ResNet 101) to improve the recognition accuracy [24]. Then, the feature map is sent to the RPN network to generate anchors, which are marked with foreground and background, and the foreground area with high score is selected as the recommended area ROIs. These ROIs are sent to the ROI subnet for further training, and 300 recommended windows are generated for each infrared image of power equipment. At the same time, the characteristic map of the full convolution layer is calculated with the multilayer convolution kernel to generate a position sensitive score map. The ROI and Score Maps are input into the later Softmax layer for vote. Through the Softmax layer for classification, the ROI with the highest score is finally obtained, that is, the location and type of the object located and recognized. The architecture of R-FCN algorithm is shown in Figure 3.

(1) Residual Network. When the depth of the deep learning network reaches a certain degree, the problems of gradient disappearance and gradient explosion often appear during training. In order to solve this problem, the residual network (ResNet) is used to improve the R-FCN algorithm; that is, the residual network is selected as the backbone network. The residual element is essentially the mapping residual required for fitting through these stacked layers. Suppose that the network mapping is $H(x)$ and the residual mapping function of the network is $F(x)$, $F(x) = H(x) - x$. The so-called residual is the difference between the observed value

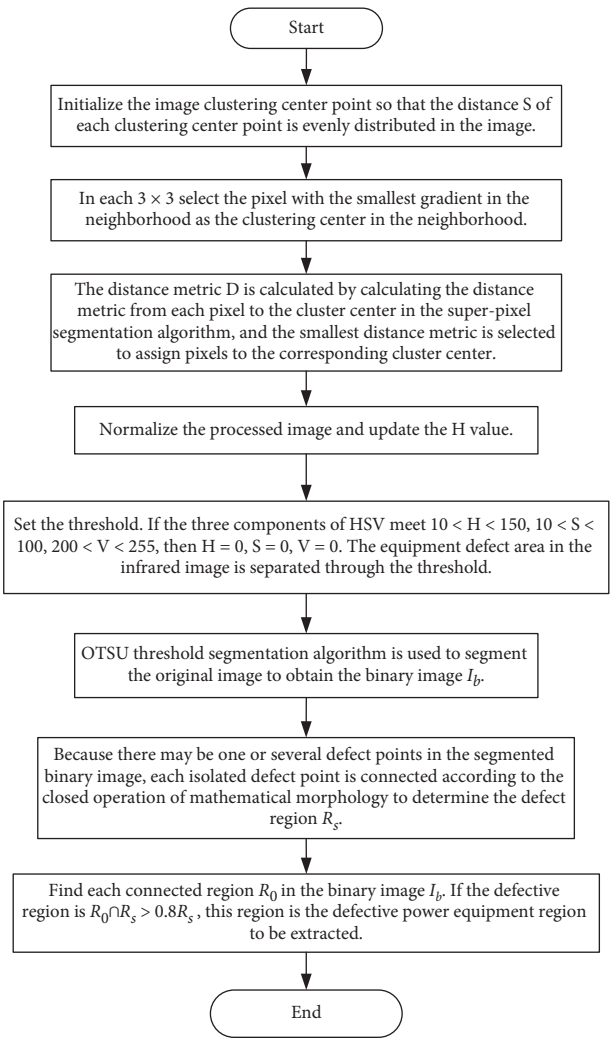


Figure 2: Defect region extraction process based on HSV.

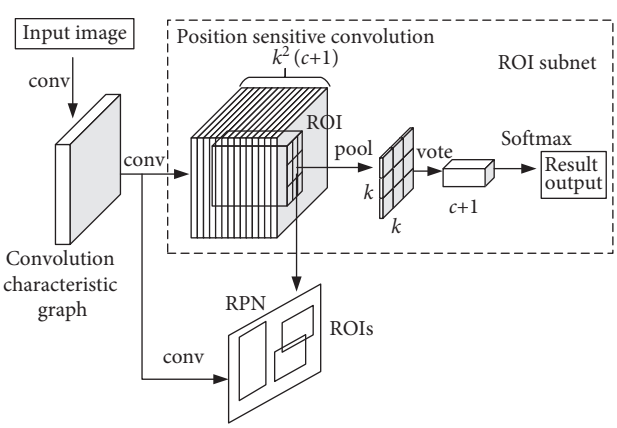


Figure 3: Architecture of R-FCN algorithm.

$H(x)$ and the estimated value x. The advantage of ResNet network is that it uses the stacking layer to fit $H(x)$ to get the mapping $H(x) = F(x) + x$. The advantage of this representation is that if the model has been fitted to the best state, it only needs to make $F(x) = 0$ to get $H(x) = x$, so as to avoid the disappearance and explosion of gradient.

The input x_m and output x_{m+1} of the m-th residual unit are expressed as follows:

$$\begin{aligned} x_{m+1} &= f\left(h\left(x_m\right) + \delta\left(x_m, \omega_m\right)\right), \\ x_M &= x_m + \sum_{i=1}^{M-1} \delta\left(x_i, \omega_i\right), \end{aligned} \tag{6}$$

where δ is the ReLu activation function, ω is the weight between each unit, m, M respectively represent the shallow residual unit and deep residual unit, $h(\cdot)$ represents the identity mapping, and x_M is the final output of the residual unit.

In order to learn more and more abstract image features, the proposed method selects ResNet 101 network, and its configuration is shown in Table 1.

(2) RPN Network. The input of RPN network is image feature graph. According to anchor mechanism, 9 rectangular boxes with different sizes are generated for each point. When training the RPN network, compare the anchor with the manually calibrated true value area in the data set, mark the anchor frame with the largest overlap ratio as the foreground, and mark the anchor frame with an overlap ratio greater than 0.7 as the foreground sample. Mark the anchor box whose overlap ratio is less than 0.3 as the background sample. Select the positive and negative samples of anchor in proportion, use the maximum suppression method (NMS) and other methods to screen the top 250 ROIs with the highest score, and send these preliminarily screened preselected frames to the ROI subnet.

In addition, RPN network adopts anchor mechanism, which not only solves the problem of translation invariance, but also enables R-FCN algorithm to identify and locate targets with different overall dimensions. In the actual process of infrared image recognition of power distribution equipment, due to different equipment with different shape and structure, different sizes and variable aspect ratio, in order to ensure that there are targets in the receptive field corresponding to each sliding window on the feature map, multiscale anchor is required to ensure that the candidate frame is as complete as possible to select the target [25]. In the implementation of RPN network anchor, multiscale anchor can be obtained by setting the area of reference window (base_size), different area multiples and anchor aspect ratio, so that RPN can give more accurate foreground recommendation area.

(3) ROI Subnet. The role of ROI subnet is to correct the ROIs location obtained from RPN network, so as to obtain a more accurate target location of power equipment, so as to identify ROI. A position sensitive convolution layer is added after the last layer of the full convolution network, which can realize the translation variability of the algorithm and output the $k^2(c+1)$-dimensional position sensitive fractional graph. The RPN network filters out the characteristic map of ROI with size $a \times b$, which is divided into $k \times k$ parts (bin) and convoluted with $k^2(c+1)$ convolution cores; that is, each

Table 1: Structure table of ResNet 101.

Layer name	101-Layer
conv1	7*7, stride 2
conv2_x	3*3max pool, stride 2
	$\begin{bmatrix} 1\times1 & 64 \\ 3\times3 & 64 \\ 1\times1 & 256 \end{bmatrix}\times 3$
conv3_x	$\begin{bmatrix} 1\times1 & 128 \\ 3\times3 & 128 \\ 1\times1 & 512 \end{bmatrix}\times 4$
conv4_x	$\begin{bmatrix} 1\times1 & 256 \\ 3\times3 & 256 \\ 1\times1 & 1024 \end{bmatrix}\times 23$
conv5_x	$\begin{bmatrix} 1\times1 & 512 \\ 3\times3 & 512 \\ 1\times1 & 2048 \end{bmatrix}\times 3$
Average pool, 1000-D FC, softmax	

part is mapped to a score map, in which $(c+1)$ is the number of categories plus the background to obtain the $k^2(c+1)$-layer position sensitive score map. The number of channels of bin in different positions is $a \times b \times (c+1)$, and the score of class $(c+1)$ in this position is stored. After the pooling process is completed, vote on the ROI. Sum the $k \times k$ parts bin to get the output of $(c+1)$ dimension, that is, the probability of category (score), and classify it through the softmax layer. At this time, the output result is the positioning coordinates and type of the object.

(4) OHEM. In RPN network, a large number of rectangular boxes are generated, and hundreds or thousands of regions will participate in the training of predicting target categories and locations. The proportion of power equipment is small, so the ratio of equipment background area to target area is too large, resulting in sample imbalance, which makes it difficult to identify distribution equipment. Therefore, online hard example mining (OHEM) method is used to train the network model [26]. During training, when there is a preselected area with large loss, the hard example can be trained and classified again, which can solve the problem of imbalance between positive and negative sample categories and improve the accuracy of infrared image recognition model of power equipment. The training framework of OHEM method is shown in Figure 4.

When OHEM carries out specific training, firstly, ResNet 101 is used to extract the image features of training samples, train the image classification and positioning branches, and calculate the classification loss L_{cls} and regression loss L_{reg} of each target. The loss function is calculated as follows:

$$L\left(s, d_{x,y,a,b}\right) = L_{cls}\left(s_{c^*}\right) + \zeta\left(c^* > 0\right)L_{\text{reg}}\left(d, d^*\right), \tag{7}$$

where $d_{x,y}$ is the upper left coordinate of the target area, d_a is the width of the target area, d_b is the height of the target area, ζ is the balance coefficient of classification loss and regression loss, and s is the image type.

Then, sort according to the loss value from high to low, and use the NMS method to select the first Ω samples with the largest loss value to screen out the difficult cases of this round of samples. Finally, these difficult samples are

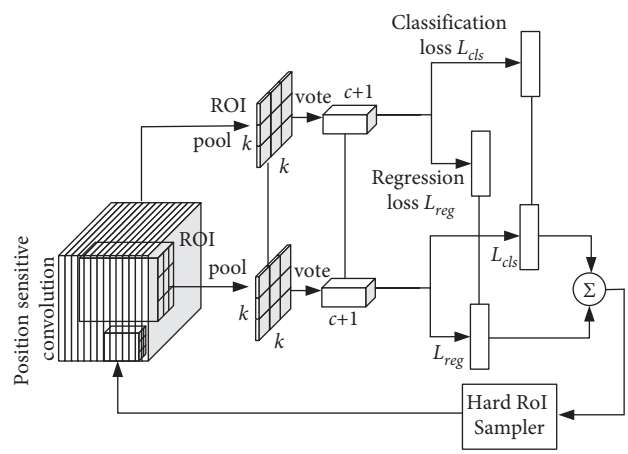


Figure 4: Training framework of OHEM method.

backpropagated to the network and trained again to update the weight of the whole network, and the penalty for low loss samples is ignored. Among them, the weight update is realized based on the random gradient descent method.

3.3.2. Fault Diagnosis of Distribution Equipment. There are many kinds of equipment in the distribution network, such as circuit breaker, potential transformer (PT), current transformer (CT), lightning arrester, and transformer. These equipment can be classified into current heating type, voltage heating type, and comprehensive heating type according to the heating factors. Different diagnostic methods are selected for different types of equipment, and different diagnostic methods have different diagnostic criteria. The defects of the equipment, such as general defects, serious defects, and critical defects, are determined through various diagnostic methods.

(1) The heating of current heating equipment is mainly due to the current thermal effect. Generally, the surface temperature judgment method and relative temperature difference judgment method are used for fault diagnosis. The relative temperature

difference ∇_T and the maximum surface temperature $T_{\max}$ of different defect degrees of each distribution equipment are shown in Table 2.

(2) Voltage heating equipment is mainly due to voltage effect. The main equipment categories include zinc oxide arrester, high-voltage bushing, and coupling capacitor. This kind of equipment generally adopts image feature judgment method, similar comparison judgment method, and comprehensive analysis judgment method. Using the thermal image characteristics of the equipment and the image feature judgment method, the fault can be found quickly. If the similar comparison discrimination method and comprehensive analysis judgment method are used, the temperature difference ΔT shall be taken as the fault diagnosis index, in which the ΔT of zinc oxide arrester is 0.5~1 K, and the ΔT of high-voltage bushing, coupling capacitor, oil immersed PT and CT are 2~3 K.

(3) Comprehensive heating equipment needs to be diagnosed in combination with the diagnosis methods of voltage heating equipment and current heating equipment, mainly including insulators, generators, and transformers. In the actual thermal fault diagnosis, the fault diagnosis indexes of various methods should be combined to improve the efficiency and accuracy of fault diagnosis [27].

According to different types of distribution equipment, the corresponding thermal fault diagnosis and judgment methods can be selected, the corresponding parameters can be calculated, and the thermal fault diagnosis of distribution equipment can be carried out according to the diagnosis criteria. The fault diagnosis process of power distribution equipment is shown in Figure 5.

Firstly, the infrared image of distribution equipment is input into the detection model for image preprocessing, and the defect area is extracted based on HSV to divide the structure of distribution equipment. Then, the trained depth learning hybrid model is used to obtain the type and location of the target equipment, and the temperature information of each structural area on the infrared image of the distribution equipment is read at the same time. Finally, according to the selected fault diagnosis method, the thermal fault state, thermal fault level, and thermal fault location of distribution equipment are determined by using the diagnosis criterion. Different from the existing fault diagnosis methods of artificial equipment, the proposed method can use the deep learning hybrid model to realize the intelligent classification and fault type diagnosis of infrared images of distribution equipment, greatly reduce the workload of inspectors, and improve the automation level of fault diagnosis of distribution equipment.

4. Experiment and Analysis

The experiment is carried out in the 64 bit operating system environment of Ubuntu 6.04.4 LTS, in which the deep learning hybrid model uses the deep learning framework PyTorch v1.3 version, the programming language is Python 3.6.0, and third-party dependent libraries such as Open CV 4.0 and NumPy 1.3 are used for batch processing of data. At the same time, the network model is trained on the GPU of dual card Tesla P100, with a total video memory of 8 GB. In this way, large batch data can be set during training to improve the convergence speed of the model. Other hardware environments are as follows: 512 GB of memory resources and 1 TB of hard disk. When processing data and I/O operations on a large scale, this can realize parallel processing and high speed and ensure the training requirements of network model.

4.1. Experimental Data Set. The research scenario is the power equipment in the distribution network, so it is necessary to collect the picture materials of equipment faults in the distribution system and then construct a fixed format data set for test training. Through safety training and professional leadership, use mobile phones, cameras, and other equipment to shoot at the site of power distribution equipment. The scene of abnormal power grid equipment is selected, and a large number of positive samples are collected in multiple directions according to the shooting angle of video monitoring. In order to ensure the rationalization of data distribution, all types and forms of equipment anomaly types are covered in the acquisition process. After that, the abnormal categories and areas of power grid equipment are marked through the open-source and free wizard marking software.

The data set contains 2580 on-site abnormal images of RGB power grid equipment during the day, 6300 mark boxes in total, and 2769 on-site abnormal images of infrared power grid equipment at night, 7000 mark boxes in total, all from the real power distribution room, power equipment plant, etc. After the data annotation is completed, the annotation file in XML format is generated, which corresponds to the real image.

4.2. Comparative Analysis of Training Speed. The proposed method combines the hybrid model of robot and deep learning for equipment fault diagnosis. ResNet network is used to optimize the R-FCN algorithm, and the defect area is extracted based on HSV to improve the diagnosis effect. In order to demonstrate the improvement effect of the proposed method, it is compared with the diagnosis methods of ResNet network, Otsu threshold segmentation, and OHEM training. The diagnosis accuracy and training time are shown in Table 3.

It can be seen from Table 3 that extracting deeper equipment fault features using ResNet network can greatly improve the diagnosis accuracy, which is 9.88% higher than that of the model. However, due to the deepening of network layers, the training time is also increased, more than 10 s. At the same time, by integrating OHEM training depth learning model, the diagnostic accuracy continued to improve by 4.41%. Due to the simple and easy implementation of the training process, the training time is only increased by 0.7 s. It can be seen that the diagnostic accuracy of the proposed

Table 2: Fault diagnosis criterion of current heating equipment.

	General defect	Serious defect	Emergency defect
Circuit breaker	$35\% \le \nabla_T < 80\%$	$80\% \le \nabla_T < 95\%$	$\nabla_T \ge 95\%$
	—	$55\% \le T_{\max} \le 80\%$	$T_{\max} > 80\%$
Disconnecting switch	$35\% \le \nabla_T < 80\%$	$80\% \le \nabla_T < 95\%$	$\nabla_T \ge 95\%$
	—	$90\% \le T_{\max} \le 130\%$	$T_{\max} > 130\%$
CT	$35\% \le \nabla_T < 80\%$	$80\% \le \nabla_T < 95\%$	$\nabla_T \ge 95\%$
	—	$55\% \le T_{\max} \le 80\%$	$T_{\max} > 80\%$
Capacitor	$35\% \le \nabla_T < 80\%$	$80\% \le \nabla_T < 95\%$	$\nabla_T \ge 95\%$
	—	$55\% \le T_{\max} \le 80\%$	$T_{\max} > 80\%$
High voltage bushing	$35\% \le \nabla_T < 80\%$	$80\% \le \nabla_T < 95\%$	$\nabla_T \ge 95\%$
	—	$55\% \le T_{\max} \le 80\%$	$T_{\max} > 80\%$

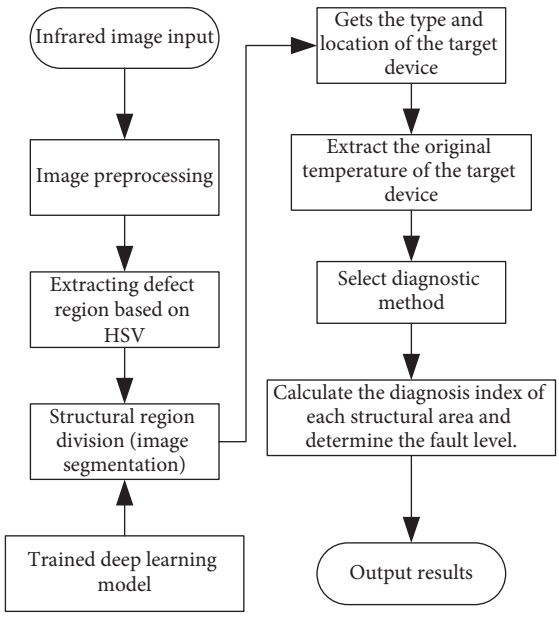


Figure 5: Fault diagnosis process of distribution equipment.

Table 3: Fault diagnosis criterion of current heating equipment.

Method	Accuracy/%	Training time/s
Original model	82.49	4.6
Original model + OTSU threshold segmentation	85.82	5.9
Original model + OTSU threshold segmentation + ResNet network	92.37	10.5
Original model + OTSU threshold segmentation + ResNet network + OHEM	96.78	11.2

method is higher than that of the basic method, but the training time increases, and the training speed decreases.

In order to demonstrate the performance of the proposed method in training speed, it is compared with reference [11], reference [13], and reference [16]. The results are shown in Figure 6.

As can be seen from Figure 6, the training time of reference [13] is the shortest, only about 5S. Because its intuitionistic fuzzy clustering algorithm based on spatial distribution information for image recognition is simple and easy to implement, the training speed is fast. Reference [13] combines discrete wavelet transform and support vector machine algorithm to complete fault diagnosis, and [16] uses artificial neural network algorithm to classify faults. Both methods are complex and take a long time to calculate, so the training time is about 10 s. By using the improved R-FCN algorithm for fault diagnosis, the proposed method uses OHEM method to train it, which can simplify the data processing process, the training speed is fast, and the training time is about 5.5 s. At the same time, the robot background is used for data analysis, which can reduce the transmission time of image information.

4.3. Comparative Analysis of Fault Diagnosis Accuracy. The accuracy of fault diagnosis is a key judgment index. The accuracy of the proposed method and the methods in reference [11], reference [13], and reference [16] for

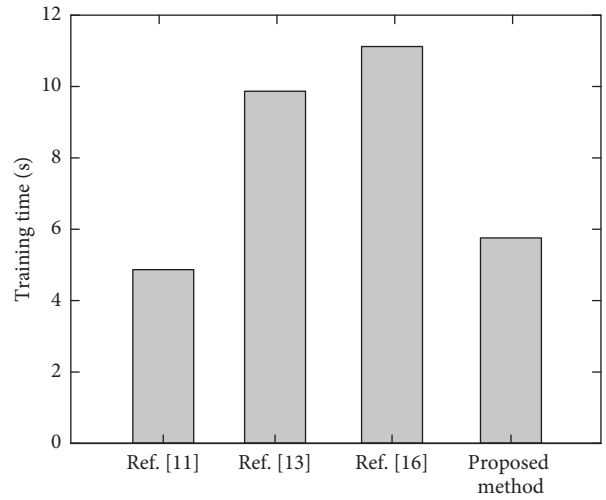


Figure 6: Training time of different methods.

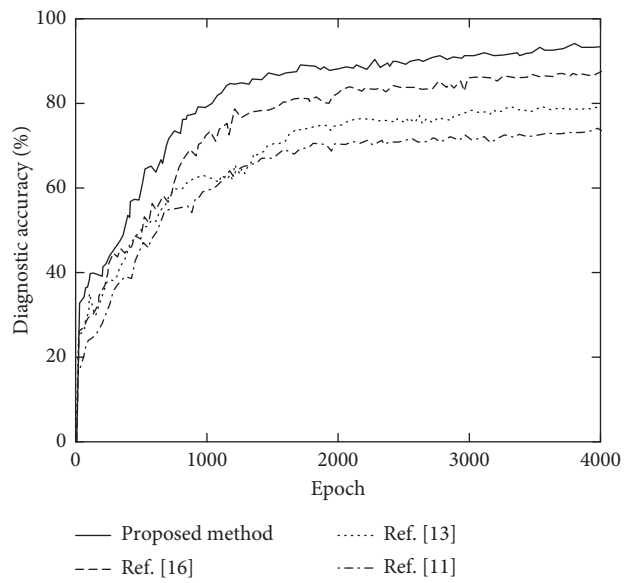


Figure 7: Diagnostic accuracy of different methods.

fault diagnosis of distribution equipment is shown in Figure 7.

As can be seen from Figure 7, compared with other methods, with the iteration of epoch, the fault diagnosis accuracy of the proposed method tends to be stable, about 92.06%. Due to its combination of robot and deep learning hybrid model, it deeply extracts the characteristics of various types of fault equipment for diagnosis, which further ensures the reliability of diagnosis results. Similarly, [16] uses artificial neural network algorithm for state recognition, but there is no efficient way to obtain the equipment state, and there is no complete database to support it. Therefore, the diagnosis accuracy is reduced by about 6% compared with the proposed method. Reference [13] adopts the improved support vector machine algorithm of genetic algorithm for fault detection, which has a good effect on high impedance fault diagnosis, but its universality is not high, so the diagnosis accuracy is about 80%. Reference [11] uses the traditional intuitionistic fuzzy clustering algorithm for graphic classification. The traditional method is difficult to apply to a large number of distribution equipment, so the diagnosis accuracy is low.

For the three fault types, the diagnostic accuracy of different methods is shown in Table 4.

It can be seen from Table 4 that the diagnostic accuracy of comprehensive heating equipment is generally lower than that of current heating equipment and voltage heating equipment. Taking the proposed method as an example, the diagnostic accuracy of comprehensive heating equipment is 89.31%, and the other two types are higher than 90%. Because the diagnostic criteria of comprehensive heating equipment are complex and easy to be confused, they affect the fault diagnosis. The recognition accuracy of current heating type defects is slightly higher, which may be due to

Table 4: Comparison results of diagnostic accuracy of each fault type.

Method	Current heating type (%)	Voltage heating type (%)	Comprehensive heating type (%)
Reference [11]	73.23	69.18	65.75
Reference [13]	82.84	80.36	79.04
Reference [16]	86.69	87.05	84.27
Proposed method	93.52	91.88	89.31

the obvious characteristics and large amount of data of infrared images of current heating type defects in the data set. However, the diagnosis accuracy of the proposed method is higher than that of other comparison methods. Taking the current heating equipment as an example, its diagnosis accuracy is as high as 93.52%, because it can well distinguish all kinds of faulty equipment by using the improved R-FCN algorithm to learn the equipment image features and evaluate the fault level according to the fault judgment. Other comparison methods only diagnose whether the equipment is faulty or not, but the diagnosis effect is poor for various specific fault types.

5. Conclusion

Nowadays, the construction of smart grid in China has entered a new stage of comprehensive and rapid development. The traditional manual detection methods have been difficult to deal with a large amount of infrared image data of power equipment. Therefore, a fault diagnosis method of distribution equipment based on the hybrid model of robot and deep learning is proposed. The image information database of distribution equipment based on robot inspection is constructed, and the OTSU method is used to extract the defect features of distribution equipment from the binary image of equipment based on HSV space. Then, the equipment defect characteristics are sent to the improved R-FCN algorithm for learning and analysis to obtain the fault type and location, and the fault level is obtained through the calculation of fault criterion. The experimental results based on PyTorch platform show that:

(1) Using the robot platform to build the image information database of distribution equipment can improve the accuracy of fault diagnosis of various equipment. The diagnostic accuracy of the proposed method for current heating equipment, voltage heating equipment, and comprehensive heating equipment is 93.52%, 91.88%, and 89.31%, respectively.

(2) Using OHEM method to train the improved R-FCN algorithm can shorten the model training time, improve the fault diagnosis efficiency, and further improve the diagnosis effect. The fault diagnosis time and accuracy are 5.5 s and 92.06%, respectively.

Since the extraction degree of the defective region will affect the recognition accuracy of the subsequent model, considering the subsequent semantic segmentation of the overall infrared image by using the convolution neural network optimized by conditional random field, the defective power equipment can be segmented in a more complex background to further improve the recognition accuracy of the subsequent model.

Data Availability

The data of the paper can be obtained from the corresponding author.

Conflicts of Interest

The authors declare no conflicts of interest.

References

[1] S. Gangolu, P. Raja, M. P. Selvan, and V. K. Murali, "Effective algorithm for fault discrimination and estimation of fault location in transmission lines," *IET Generation, Transmission & Distribution*, vol. 13, no. 13, pp. 2789–2798, 2019.

[2] N. Hu, H. Du, S. Liu, and Q. Lin, "Power equipment status information parallel fault diagnosis of based on MapReduce," *Journal of Computational Methods in Science and Engineering*, vol. 19, no. 88, pp. 1–6, 2019.

[3] E. Bashar, Q. Han, R. Wu, L. Ran, O. Alatise, and S. Jupe, "Analysis of DC offset in fault current caused by machines in a medium voltage distribution network," *Journal of Engineering*, vol. 2019, no. 17, pp. 3494–3499, 2019.

[4] H. Tian, P. Liu, S. Zhou et al., "Research on the deterioration process of electrical contact structure inside the ± 500 kV converter transformer RIP bushings and its prediction strategy," *IET Generation, Transmission & Distribution*, vol. 13, no. 12, pp. 2391–2400, 2019.

[5] W. Luo, H. Wang, L. Wang, Z. Zhu, and H. Gao, "Faulted line location method for distribution systems based on the equipment's information exchange," *Dianli Xitong Baohu yu Kongzhi/Power System Protection and Control*, vol. 47, no. 4, pp. 73–82, 2019.

[6] N. Narasimhulu, D. Kumar, and M. V. Kumar, "Detection and classification of high impedance fault in power distribution system using hybrid technique," *Journal of Circuits, Systems, and Computers*, vol. 29, no. 08, pp. 67–976, 2020.

[7] K. Chen, J. Hu, Y. Zhang, Z. Yu, and J. He, "Fault location in power distribution systems via deep graph convolutional networks," *IEEE Journal on Selected Areas in Communications*, vol. 38, no. 1, pp. 119–131, 2020.

[8] F. Wan, P. Madhika, J. Chwa, M. Mozumdar, and A. Ameri, "Automatic optimal synthesis of aircraft electric power distribution system," *International Journal of Computing and Digital Systems*, vol. 9, no. 3, pp. 363–375, 2020.

[9] K. Jia, Q. Zhao, T. Feng, and T. Bi, "Distance protection scheme for DC distribution systems based on the high-frequency characteristics of faults," *IEEE Transactions on Power Delivery*, vol. 35, no. 1, pp. 234–243, 2020.

[10] A. S. Alayande, I. K. Okakwu, O. E. Olabode, and O. K. Nwankwoh, "Analysis of unsymmetrical faults based on

artificial neural network using 11 kV distribution network of University of lagos as case study," *Journal of Advances in Science and Engineering*, vol. 4, no. 1, pp. 53–64, 2021.

[11] F. Hu, H. Chen, and X. Wang, "An intuitionistic kernel-based fuzzy C-means clustering algorithm with local information for power equipment image segmentation," *IEEE Access*, vol. 8, no. 6, pp. 4500–4514, 2020.

[12] H. Liu, Y. Wang, and W. Chen, "Anomaly detection for condition monitoring data using auxiliary feature vector and density-based clustering," *IET Generation, Transmission & Distribution*, vol. 14, no. 1, pp. 108–118, 2020.

[13] Youness, Mohammadnian, A. Turaj, and A. Soroudi, "Fault detection in distribution networks in presence of distributed generations using a data mining–driven wavelet transform," *IET Smart Grid*, vol. 2, no. 2, pp. 163–171, 2019.

[14] Q. Yang, J. Ruan, and Z. Zhuang, "Fault diagnosis of circuit breakers based on time-frequency and chaotic vibration analysis," *IET Generation, Transmission & Distribution*, vol. 14, no. 7, pp. 1214–1221, 2020.

[15] T. Zhang, H. Yu, P. Zeng, L. Sun, C. Song, and J. Liu, "Single phase fault diagnosis and location in active distribution network using synchronized voltage measurement," *International Journal of Electrical Power and Energy Systems*, vol. 117, no. 5, 2020.

[16] S. Frizzo-Stefenon, M. C. Silva, D. W. Bertol, L. H. Meyer, and A. Nied, "Fault diagnosis of insulators from ultrasound detection using neural networks," *Journal of Intelligent and Fuzzy Systems*, vol. 37, no. 5, pp. 6655–6664, 2019.

[17] B. Wang, M. Dong, M. Ren et al., "Automatic fault diagnosis of infrared insulator images based on image instance segmentation and temperature analysis," *IEEE Transactions on Instrumentation and Measurement*, vol. 69, no. 8, pp. 5345–5355, 2020.

[18] H. Liang, Y. Liu, G. Sheng, and X. Jiang, "Fault-cause identification method based on adaptive deep belief network and time-frequency characteristics of travelling wave," *IET Generation, Transmission & Distribution*, vol. 13, no. 5, pp. 724–732, 2019.

[19] M. Gholami, A. Abbaspour, M. Moeini-Aghtaie, M. Fotuhi-Firuzabad, and M. Lehtonen, "Detecting the location of short-circuit faults in active distribution network using PMU-based state estimation," *IEEE Transactions on Smart Grid*, vol. 11, no. 2, pp. 1396–1406, 2020.

[20] W. Hu, C. Ruan, H. Nian, and D. Sun, "Simplified modulation scheme for open-end winding PMSM system with common DC bus under open-phase fault based on circulating current suppression," *IEEE Transactions on Power Electronics*, vol. 35, no. 1, pp. 10–14, 2020.

[21] K. Zhu and P. W. T. Pong, "Fault classification of power distribution cables by detecting decaying DC components with magnetic sensing," *IEEE Transactions on Instrumentation and Measurement*, vol. 69, no. 5, pp. 2016–2027, 2020.

[22] L. Romero, J. Blesa, V. Puig, G. Cembrano, and C. Trapiello, "First results in leak localization in water distribution networks using graph-based clustering and deep learning," *IFAC-PapersOnLine*, vol. 53, no. 2, pp. 16691–16696, 2020.

[23] W. Wang, N. Yu, Y. Gao, and J. Shi, "Safe off-policy deep reinforcement learning algorithm for volt-VAR control in power distribution systems," *IEEE Transactions on Smart Grid*, vol. 11, no. 4, pp. 3008–3018, 2020.

[24] D. A. León-Vargas, V. A. Bucheli-Guerrero, and H. A. Ordoez, "Solar radiation prediction on photovoltaic systems using machine learning techniques," *Revista Facultad de Ingeniería*, vol. 29, no. 10, pp. 1–20, 2020.

[25] S. Yamane and K. Matsuo, "Adaptive control by convolutional neural network in plasma arc welding system," *ISIJ International*, vol. 60, no. 5, pp. 998–1005, 2020.

[26] M. Nabati, H. Navidan, R. Shahbazian, S. A. Ghorashi, and D. Windridge, "Using synthetic data to enhance the accuracy of fingerprint-based localization: a deep learning approach," *IEEE Sensors Letters*, vol. 4, no. 4, pp. 1–4, 2020.

[27] M. S. Erbnescu, N. C. Manea, L. Streba et al., "Automated gleason grading of prostate cancer using transfer learning from general-purpose deep-learning networks," *Romanian Journal of Morphology and Embryology*, vol. 61, no. 1, pp. 149–155, 2020.

Hindawi
Journal of Robotics
Volume 2022, Article ID 3640851, 11 pages
https://doi.org/10.1155/2022/3640851

Research Article

Design Method of Intelligent Ropeway Type Line Changing Robot Based on Lifting Force Control and Synovial Film Controller

Jiazhen Duan,[1] **Ruxin Shi**,[2] **Hongtao Liu**,[1] **and Hailong Rong**[3]

[1]*State Grid Changzhou Power Supply Company, Transmission and Distribution Engineering Company, Changzhou, Jiangsu 213000, China*
[2]*State Grid Changzhou Power Supply Company, Office, Changzhou, Jiangsu 213000, China*
[3]*School of Mechanical Engineering and Rail Transit, Changzhou University, Changzhou, Jiangsu 213164, China*

Correspondence should be addressed to Hailong Rong; rhle_16@163.com

Received 25 February 2022; Revised 25 March 2022; Accepted 30 March 2022; Published 12 April 2022

Academic Editor: Shan Zhong

Aiming at the problems of low efficiency, reliability, and safety of manual construction for demolition of old lines, a design method of an intelligent ropeway type line changing robot based on lifting force control and synovial film controller is proposed. First, the mechanical model of robot load and line sag is established, and the sag of the overhead line where the robot is located is used to calculate the jacking force that the jacking device needs to provide to the robot. Then, by introducing the radial basis function (RBF) neural network adaptive algorithm into the synovial controller, an adaptive sliding mode position control algorithm based on the RBF neural network is designed to achieve high-precision motion control of the robot in complex operating environments. Finally, based on the compactness, weight, and reliability of the robot, the optimal design is carried out from four aspects of topology, size, shape and morphology, and the design scheme of the robot for wire removal is proposed, and the robot is produced. The developed robot and the other three robots are compared and analyzed under the same conditions through simulation experiments. The results show that the maximum operating time, maximum climbing angle, and maximum traveling speed of the robot developed in this study are all optimal, which are 45 min, 10°, and 1 m/s respectively, and the performance is better than the other three comparison algorithms.

1. Introduction

With the continuous growth of the national electricity demand, more and more transmission lines are put into use. At the same time, there are a large number of old lines. Old lines may affect the safe operation of the power system, personal safety, and property safety [1–3]. At present, the demolition of the old lines of the power grid is basically completed by manual work, and there are many problems such as long construction time, high construction cost, and wide influence range [4, 5]. In view of the above problems, using robots to replace humans to complete construction work such as overhead line demolition and replacement has broad application scenarios and development space and is currently a research hotspot in the industry [6–8].

2. Related Work

Using robots instead of manual construction of the foot frame and completion of the loosening of wire and guide wire back pumping procedures after upping the tower can improve the efficiency of demolition construction and reduce the construction difficulty and cost. It plays a significant role in improving the overall technical level of power line construction [9, 10]. dos Santos [11] developed a rope-climbing robot that can move on distribution lines and the corresponding motion planning to avoid collisions with insulators and other devices. It designs a geometric motion planning control method by using the quintic polynomial interpolation method, so that the robot articulated suspension can retract when approaching obstacles and expand

after crossing obstacles. However, this method does not analyze the accurate control of robot motion speed. In view of the difficulty in calculating the drop of power lines, Zengin et al. [12] proposed a new method to accurately measure the drop of transmission lines based on the power line inspection robot and by using the sensors carried by the robot to collect data and send it remotely. However, this method does not study the motion control method of the robot itself. In view of the problem of poor battery life of the live working robot, Jiang et al. [13] constructed a method to optimize the motion energy consumption of the robot arm. It adapts the genetic algorithm and selects the appropriate algorithm parameters to solve the optimal motion planning of the robot energy consumption, which improves the operation efficiency of the robot. However, this method only takes the lowest energy consumption as the objective function and has limitations. Nguyen et al. [14] developed a robot for cleaning solar panels and simplified it. By establishing the control motion equation of the robot during driving and combining with the linear quadratic regulator, a scheme is proposed to ensure the stable movement of the system and track the desired trajectory. However, this method only focuses on the control method of the robot motion path and cannot guarantee the accuracy of speed control. Aiming at the difficulties and high risks of power grid fault detection, Zhao et al. [15] proposed a patrol robot that can detect equipment faults and identify infrared images of faulty equipment based on infrared imaging technology and support vector machine technology. Although the detection efficiency is enhanced, the accuracy of the method using robots for fault identification is not significantly improved. Xie et al. [16] proposed an integrated 3D printing tube-climbing robot composed of an ordered new soft bending mechanism. The finite element method is used to predict the maximum bending angle of the module, and the output torque and recovery torque are obtained by building a torque test bench. On this basis, the model of the whole robot is established. However, the calculation process of this method is complicated, and the cost is too high to apply in practical engineering. Song et al. [17] proposed an automatic inspection mechanical diagnostic robot based on the fuzzy search method of acoustic signals. The intelligent control of the inspection path of the mechanical diagnostic robot is realized through rough sets, fuzzy neural network (FNN), and self-positioning azimuth correction, and it obtains navigation and search method by using the extracted fault sound signal for fuzzy reasoning. However, this method does not analyze the control complexity and motion accuracy of the robot itself.

Based on the above analysis, in view of the low efficiency and high cost of the current old line dismantling work, a design method of intelligent ropeway type line changing robot based on lifting force control and synovial film controller is proposed. The basic idea is as follows: the mechanical model of robot load and line sag is established to analyze its force and the calculation method of lifting force required by the robot, and by introducing the RBF neural network into the synovial controller, an adaptive algorithm which can accurately control the robot motion is proposed. Compared with traditional detection methods, the innovations of the proposed method are listed:

(1) A new method is proposed to calculate the lifting force of the robot. The lifting force required by the robot is calculated based on the vertical radian of the position of the overhead line where the robot is located and the attitude sensor

(2) The RBF neural network is introduced into the synovial controller, which greatly improves the precision of robot motion control in complex operating environment

(3) On the basis of not increasing the weight of the robot, the reliability of the robot is improved by optimizing and adjusting the shape, position, and quantity of different structures

3. The Proposed Method

3.1. Automatic Control of Lifting Force Based on Vertical Radian of Overhead Line. When the robot moves on the wire, it needs certain climbing ability to ensure that the wheel will not slide at any time as far as possible, which means that the wheel surface and overhead line are relatively static. This requires that the friction force provided by the driving wheel and the fixed wheel is enough to overcome the influence of the heavy torque and the pulling torque of the rope, so that the robot can remain relatively static [18].

Since the load of the robot is closely related to the sag of overhead line, the mechanical model of the load and sag of the robot should be established first. The statics analysis of the robot is the same in ascending and descending stages. The following is an example of robot climbing. The static model of the robot when it goes uphill is shown in Figure 1.

In Figure 1, M represents the mass of the robot, g represents the acceleration of gravity, v represents the speed of the robot, d_0 represents the distance between the front and rear wheels of the robot, N_f and N_b, respectively, represent the positive pressure of the front and rear wheels of the robot on the line, F_{fs} and F_{bs}, respectively, represent the static friction forces of the front and rear wheels of the robot, M_{s1} and M_{s2}, respectively, represent the braking torque applied by the front and rear wheels of the robot, δ_{fs} and δ_{bs}, respectively, represent the static friction coefficients of the front and rear wheels of the robot, and M_{fs} and M_{bs} are, respectively, the rolling friction moments suffered by the front and rear wheels of the robot, whose values are very small and negligible. It can be regarded as $M_{fs} = M_{bs} = 0$.

Combined with the studies [19, 20] on the dynamic modeling of the shock contact phenomenon in a closed-loop robot chain and the kinetic models based on kinematic control of ellipsoids and cubic nanoparticles, according to Figure 1, the static analysis equation of the robot when it is uphill can be obtained as follows.

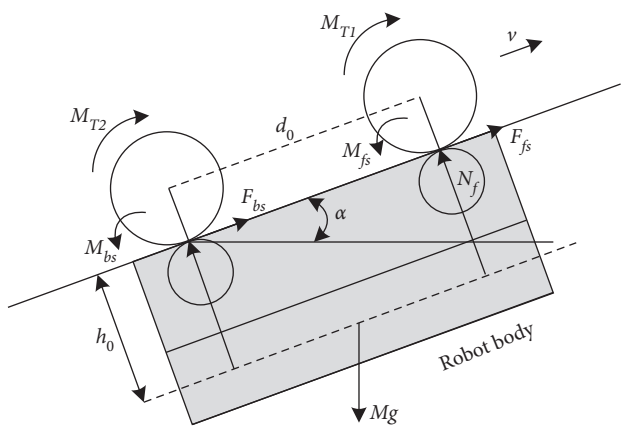


Figure 1: Statics model of the robot going uphill.

$$\begin{cases} M_{T1} + M_{T2} - N_f d_0 + Mgh_0 \sin\alpha + \frac{1}{2} Mgd_0 \cos\alpha = 0, \\ F_{fs} + F_{bs} - Mg \sin\alpha = 0, \\ N_f + N_b - Mg \cos\alpha = 0, \\ M_{T1} = F_{fs} r_w, \\ M_{T2} = F_{bs} r_w. \end{cases} \quad (1)$$

In (1), r_w is the angular velocity of the robot wheel.

It can be seen from (1) that in the process of wire climbing, when the driving torque is constant, the running state of the robot is mainly affected by friction. The dip angle of the transmission wire and the support force provided by the lifting device to the driving wheel are two main factors affecting the friction. Therefore, the lifting force required by the lifting device can be calculated by detecting the vertical radian of the position of the overhead wire where the robot is located.

For the robot that adopts the lifting grasping principle with fixed stroke, the friction between the driving wheel and overhead line is constant and cannot be changed according to the change of load, so the adaptability is poor [21, 22]. When the load of the robot is large, because the friction force is not enough to keep the driving wheel and the wire relatively still, it will affect the traveling speed or even cannot climb the slope. In addition, when downhill or under a small load, excessive friction will lead to waste of power and can affect the duration of continuous operation of the robot. Because there are certain differences in the sag of different overhead lines, it is necessary to detect the inclination of the wire before calculating the power required for the robot to move and control the pressure output of the lifting mechanism to adapt to the current load condition.

In order to solve the above problems, the attitude sensor is used to sense the own attitude of robot so as to obtain the inclination of the wire indirectly [23, 24]. The outputs of these sensors are calibrated by the low-power processor in the sensor, and then, these outputs are fused by the complementary filtering algorithm or extended Kalman algorithm, and the attitude quaternion characterizing the inclination of the transmission wire is obtained. Finally, the attitude quaternion is transformed into Euler angle (pitch angle, roll angle, and azimuth angle) to lay a foundation for the subsequent control of lifting force. The torque to be overcome by the driving wheel is shown in Figure 2.

It can be seen from Figure 2 that the lifting force required to be provided by the two driving wheels is different. After measuring the inclination of the wire, the friction force needed by the driving wheel and the fixed wheel can be calculated combining with the weight of the robot and the real-time drag force of the tow rope, and the lifting force needed by the robot can be calculated on this basis.

The specific design scheme of the lifting force control system is shown in Figure 3. The entire jacking force control system adopts a single closed-loop control method. The control system takes the torque calculation result at the contact point between the driving wheel and the power line as feedback, takes the torque exerted by the driving wheel as the reference, and the difference between the feedback and the reference as the input of the controller, and the output of the controller directly controls the brushless DC. The output of the motor finally achieves the purpose of indirectly adjusting the force of the driving wheel through the electric cylinder. The two driving wheels share a set of torque calculation modules, but the control systems are different. The design of the controller is one of the key points. First, the mathematical model of each part of the brushless DC motor and electric cylinder is constructed, and then, the mathematical model of the controller is constructed by using the automatic control theory, and then, the stability of the controller is analyzed by the Lyapunov method, and other related control theories are used. The robustness and adaptability of the controller are analyzed, and finally, a robust controller is designed.

The entire lifting force control system adopts a single closed-loop control method. The control system takes the torque calculation result at the contact point between the driving wheel and the power line as the feedback and takes the torque exerted by the driving wheel as the input, and the difference between the feedback and the input is the input of the controller. Since the output of the controller can directly control the output of the brushless DC motor, the electric cylinder can indirectly adjust the force applied by the driving wheel.

The two driving wheels share a set of torque calculation modules, and the difference is that their control systems are different. The design process of the controller consists of 4 steps:

(1) Build the mathematical model of each part of the brushless DC motor and electric cylinder
(2) Use automatic control theory to build the mathematical model of the controller
(3) Use the Lyapunov method to analyze the stability of the controller
(4) Use other related control theories to analyze the robustness and adaptability of the controller

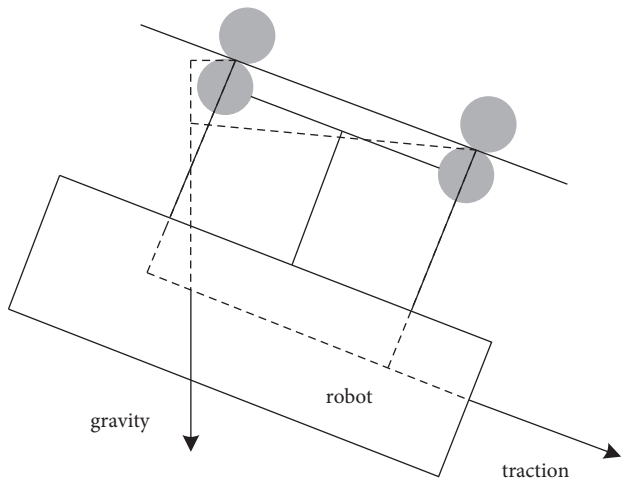


Figure 2: The torque to be overcome by the driving wheel.

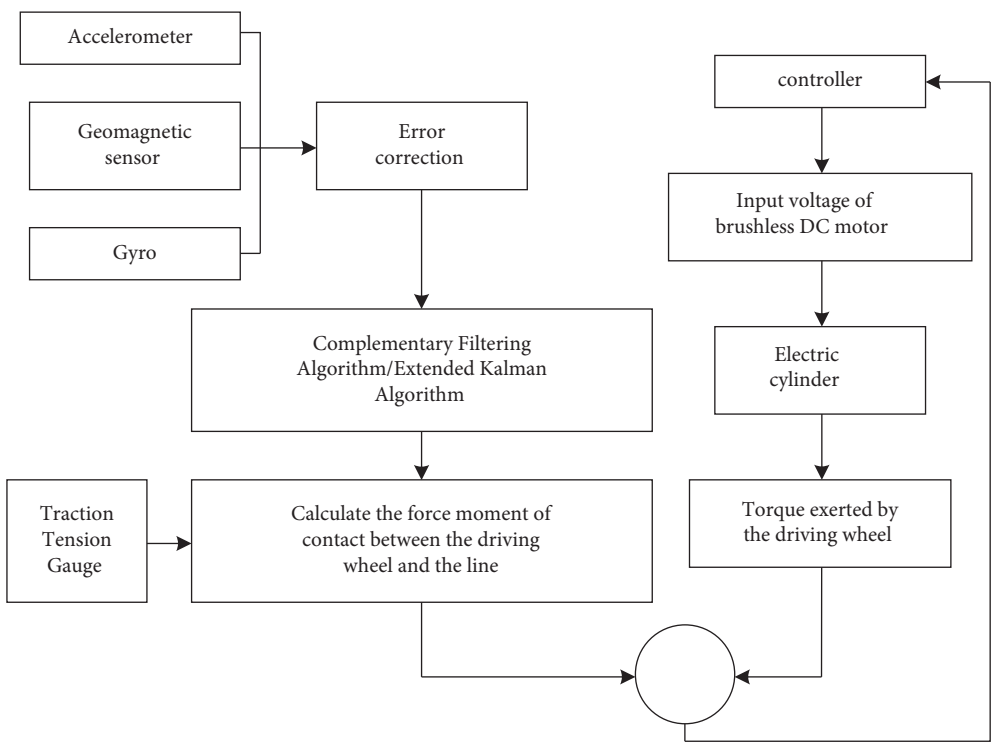


Figure 3: The design scheme of the lifting force control system.

3.2. Adaptive Speed Control Based on Neural Network Synovial Controller. When the robot moves along the overhead line, the vertical radian of the overhead line and the length of the traction rope change all the time. In addition, the operating environment is very complicated due to the slippery circuit caused by rain or the swing caused by wind. The complex environment will cause the robot load to change constantly. For the robot controller that realizes speed control through the PID controller, it cannot achieve high-precision speed control in a time-varying environment of load due to its low degree of freedom: when the load of the robot is small, its traveling speed will be fast, which will affect the fastening installation of the hook. When the load of the robot is large, its traveling speed will be very slow, which will affect the construction efficiency [25, 26]. Therefore, it is necessary to design a more accurate controller to control the walking speed of the robot.

In this study, the speed of the brushless DC motor is controlled by using a synovial controller, so as to achieve accurate control of robot walking speed. The synovial control method for system parameter variations and external disturbances has good robustness and complete adaptability. In practical applications, due to the back-and-forth switching of the control action, the inertia and delay of the system, and the measurement error and other influencing factors, the structure control will appear high-frequency

chattering in the sliding mode, which seriously affects the control performance of the system. It is difficult to solve the above problems only by improving the synovial control method. There will be a static error, and the implementation process of the high-order sliding mode control algorithm is very complicate. So, it is difficult to apply in practice.

This study considers the application of adaptive control idea. Combining the radial basis function (RBF) neural network adaptive algorithm with sliding mode variable structure control, it designs the corresponding RBF neural network adaptive sliding mode position control algorithm and finally realizes the high-precision motion control of the robot in the above complex operation environment. The RBF neural network is an advanced intelligent control algorithm, which has strong self-learning, self-adaptation, and self-organization functions, and has a good application prospect in dealing with nonlinear and uncertain problems of control systems. In addition, the RBF neural network has good approximation ability, simple network structure, and fast learning ability.

The main body of the synovial controller based on the RBF neural network is the synovial controller. By introducing the RBF neural network into the synovial controller, the synovial surface switching function of the controller is adjusted, and the external load disturbance component is added to the switching function. Therefore, the synovial controller generated by modifying the switching function with the neural network becomes an adaptive synovial controller. Once the external disturbance due to the external environment change, the external disturbance will make a sharp change in the switch function of the synovial surface, so that the adaptive synovial controller can respond to the external disturbance quickly and adjust the input current of the brushless DC motor timely; thus, the speed control accuracy is greatly improved.

In order to improve the position control accuracy of the brushless DC motor, the following torque balance equation is given, considering the changes of internal parameters and external loads.

$$\dot{\omega} = -(a + \Delta a)\omega + (b + \Delta b)i - (z + \Delta z). \tag{2}$$

In (2), $a = B/J$, $b = KT/J$, $z = TL/J$. Δa, Δb, and Δz, respectively, represent the disturbance variation caused by the disturbance of the internal parameters of the system and the disturbance of the external load.

In order to make the corresponding angle θ of the position controller track the set angle θ_d faster, the position tracking error of the controller can be expressed as

$$e(t) = x_1 = \theta_d - \theta. \tag{3}$$

At this time, the following equation is established.

$$\begin{cases} e(t) = x_2 = \dot{\theta}_d - \dot{\theta}, \\ \dot{e}(t) = \dot{x}_2 = \ddot{\theta}_d - \ddot{\theta}, \\ = \ddot{\theta}_d + (a + \Delta a)\dot{\theta}_d - (a + \Delta a)x_2 - (b + \Delta b)i + (z + \Delta z). \end{cases} \tag{4}$$

It can be seen from (4) that there is $\dot{e} \longrightarrow 0$ when $e \longrightarrow 0$, so the position controller meets the design requirements. At this time, the sliding mode surface switching function is set as

$$s = x_2 + cx_1, \quad c > 0. \tag{5}$$

In (5), $x = [x_1, x_2]^T$ represents the input of the neural network, and c is a constant.

The neural network adaptive sliding mode controller system structure can be divided into three parts: sliding mode variable structure controller, RBF network, and adaptive law.

The input of the neural network will continuously change the size of the weights after learning by the neural network, so that the output function $\hat{f}(x)$ approximates the ideal nonlinear function $f(x)$. The output $\hat{f}(x)$ of the RBF network and the ideal nonlinear function $f(x)$ are shown in the following equations, respectively.

$$\hat{f}(x) = \hat{W}^T h_f(x). \tag{6}$$

$$f(x) = (a + \Delta a)\dot{\theta}_d - (a + \Delta a)x_2 + (z + \Delta z). \tag{7}$$

In (6), $h_f(x)$ is the Gaussian function of the RBF neural network and $\hat{W}$ represents the weighted vector.

In order to further improve the chattering problem of the adaptive sliding mode of the RBF network, the reaching law is optimized, and the optimized reaching law is shown as

$$\dot{s} = -\mu s^2 |s| \operatorname{sgn} s - \beta s, \quad \mu > 0, \beta > 0. \tag{8}$$

In (8), the reaching law can be divided into power part and exponential part. When the distance between the moving point of the control system and the sliding mode surface is large, the value of s is large. In this case, the exponential part and the power part work simultaneously and the approaching speed is fast. When $s \longrightarrow 0$, the power part tends to zero, and only the exponential part plays a role. The jitter caused by the sign function $\operatorname{sgn} s$ will diminish with the decrease of the power part. Therefore, the design of approach law not only ensures the convergence speed but also makes the dynamic response of the control system more stable.

Finally, the control law of the system can be obtained as

$$i = \frac{1}{b}\left[\ddot{\theta} + \hat{f}(x) + cx_2 + \mu s^2 |s| \operatorname{sgn} s + \beta s\right]. \tag{9}$$

Here, the gradient descent method is used to learn the RBF neural network. If the learning rate η is set to a fixed value in the process of weight $\hat{W}$ update, it will cause problems such as low learning efficiency and slow convergence speed. Therefore, the adaptive learning rate is used to adjust the learning rate online, which can speed up the learning rate, while ensuring the stability of the system and the stability of the learning process. The recursive error is

$$E(k) = \frac{1}{2}y(k)^2 - \frac{1}{2}y(k-1)^2. \tag{10}$$

The rules for adjusting the learning rate according to the size of the recursive error are as

$$\begin{cases} \gamma_1 \times \eta(k-1), & (k) < E(k-1), \\ \gamma_2 \times \eta(k-1), & \gamma_3 \times \eta(k-1) < E(k), \\ \eta(k-1), & \text{else}. \end{cases} \tag{11}$$

In (11), γ_1, γ_2, and γ_3 are the proportional constants.

Thus, the specific design scheme of the controller can be obtained as shown in Figure 4.

In Figure 4, SMC represents the synovial controller, BLDCM represents the brushless DC motor, the RBF network represents the radial basis function neural network, and adaptive law represents the law of adaptive adjustment.

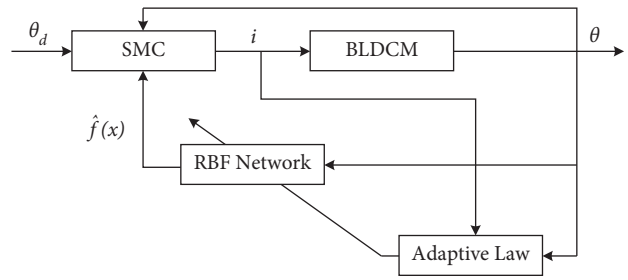


Figure 4: Design scheme of synovial controller based on the neural network.

3.3. Design of Line Changing Robot. The design of the line changing robot needs to ensure the following three technical indicators:

(1) Compact structure: the uncompact structure of the robot affects the flexibility of its movement and reduces the work efficiency

(2) Weight size: the weight of the robot will increase the time and energy consumption of the tower up and down, and the high energy consumption during the wiring, and prolong the continuous working time.

(3) Reliability: improving the reliability of robot operation is very important for safety and maintenance cost reduction in engineering applications

(4) Mobility: the mobility of a robot affects its adaptability to the operating environment

(5) Load capacity: in order to ensure the working ability of the robot, the robot also needs to bear the maximum load ≥5 kg in addition to carrying its own weight

(6) Speed requirements: the robot should improve its maximum moving speed as far as possible under the premise of ensuring stability and reliability and reduce the influence of external factors such as wind on work efficiency

(7) Self-protection ability: in order to prevent the robot from falling, it should have certain safety self-protection measures

Aiming at the above seven indexes, the optimized design is mainly carried out from four aspects of topology optimization, size optimization, shape optimization, and morphology optimization.

Topology optimization: for the critical path of load, aluminum alloy polypropylene composite lightweight laminate is used. For the noncritical path, the polymer plate is used to minimize the bodyweight while ensuring the payload. Designing based on the payload transfer path and optimal material distribution can improve the overall structure and reduce the overall design cost.

Dimension optimization: the volume of robot parts is set as the objective function, and the combination of optimal design parameters is calculated based on dimension parameters of robot parts such as thickness of plate and section area of pillar.

Shape optimization: based on the designed robot topology, the geometry of different parts is optimized to improve the strength of parts.

Morphology optimization: based on the weight of the designed robot, after fixing the weight, adjust the shape, position, and quantity of different concave-convex structures and optimize the stiffness and mode of sheet metal structural parts, so as to eliminate potential weak links and improve the robot the purpose of reliability.

The main process of robot design includes software and hardware design. The hardware design includes the four aspects:

(1) Hardware design of robot ontology structure

(2) Hardware design of the robot control system

(3) The design of the main control module and the peripheral basic hardware circuit

(4) Hardware circuit design

The specific design content is given in Table 1.

The hardware design drawing and the line changing robot obtained are shown in Figures 5 and 6, respectively.

The software design mainly includes the realization of the robot control system, the display of robot running state, and the realization of robot remote control. The control flowchart of the line changing robot is shown in Figure 7. The workflow of robot mainly includes the following steps:

(1) First initialize the robot before work, start the subroutine receiving instructions, and connect it with the ground control device

(2) The robot starts the synovial controller based on the RBF neural network to realize the adaptive speed control at work

(3) The robot starts the subroutine that collects and sends its own state and the subroutine that collects and sends video and sends the above data information to the ground receiving device

(4) The robot receives the instruction from the ground and judges the type of the instruction and executes it until it stops working after receiving the disconnected instruction

Table 1: The design content of the line changing robot.

	Content	Purpose
1	Simplify	Reduce structural complexity
2	Redundancy	Redundancy of key structural parameters, functional realization
3	Derating	Reduce the failure rate of components
4	Components outline	Control and manage electronic components and mechanical parts
5	Critical and important	Utilize existing resources to improve the reliability of key and important parts
6	Environmental protection	Choose appropriate materials or solutions to requirements of waterproof, electromagnetic protection, ambient temperature, and humidity.
7	Software reliability	By adopting N-version programming and implementing software engineering and specifications to improve critical reliability
8	Packaging, shipping, storage	Determine packaging protection measures, meet shipping, storage requirements
9	Ergonomics	Make the equipment easy to operate and maintain

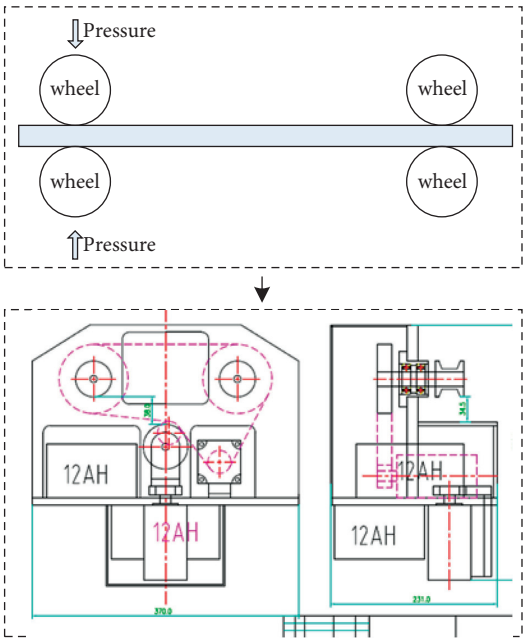


Figure 5: Hardware design of line changing robot.

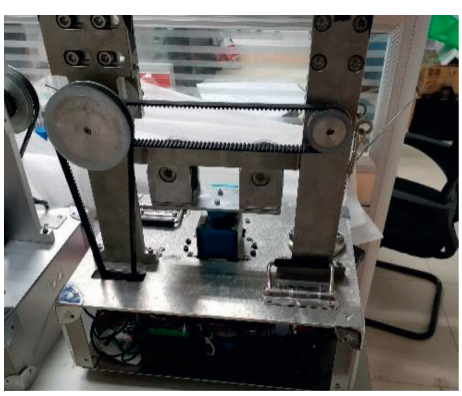


Figure 6: Actual object of line changing robot.

4. Experiment and Analysis

4.1. Practical Application Experiment. Some experiments are carried out on the 220 kV overhead transmission line for the designed line changing robot. It is shown in Figure 8.

In the process of practical application, a number of performances of the robot are tested. It measures and analyzes the indicators including the operation success rate of the line changing robot, the online moving speed, the speed control accuracy, the climbing angle, the continuous operation time, and the environmental adaptation.

The experimental results are given in Table 2.

The experimental results show that the robot can move up to 1 m/s, remote control distance up to 1 km, climbing angle up to 10°, continuous operation time up to 45 minutes, and waterproof grade up to 6 on 220 kV and below overhead line. The adaptability of the environment is as follows: it can work normally in the environment, in which the wind is level 5 or less, and the precipitation is 8 ml or less per hour, and the temperature range is −10–40°C. In addition, the proposed intelligent ropeway type line changing robot based on lifting force control and synovial film controller does not need to build scaffolding and cannot hinder the normal operation of the facilities to be crossed. Compared with on-site construction, the time cost is shortened from days to hours, and the process is fast, efficient, simple, safe, and reliable. It can be reused after a one-time investment, which greatly reduces the cost.

4.2. Performance Comparison Analysis. Since the main external factor affecting the high-altitude operation of robot is the wind, the robot developed in this study and the robot developed in [11, 13, 14] are, respectively, compared and analyzed in terms of operation duration, climbing angle, and traveling speed of the robot under different winds.

The relationship between the maximum operating time, maximum climbing angle, and maximum traveling speed of different robots and wind power is shown in Figures 9–11, respectively.

It can be seen from Figures 9–11 that the developed robot has the largest working time, forward speed, and climbing angle compared with other robots, no matter in no wind or in different wind levels. It indicates that the developed robot is better able to work outdoors at high altitude. The reason is that the sliding mode controller is used to control the robot, and the RBF neural network is introduced in it, which can self-learn, self-adapt, and self-organize to achieve high-precision motion control of the robot in the complex

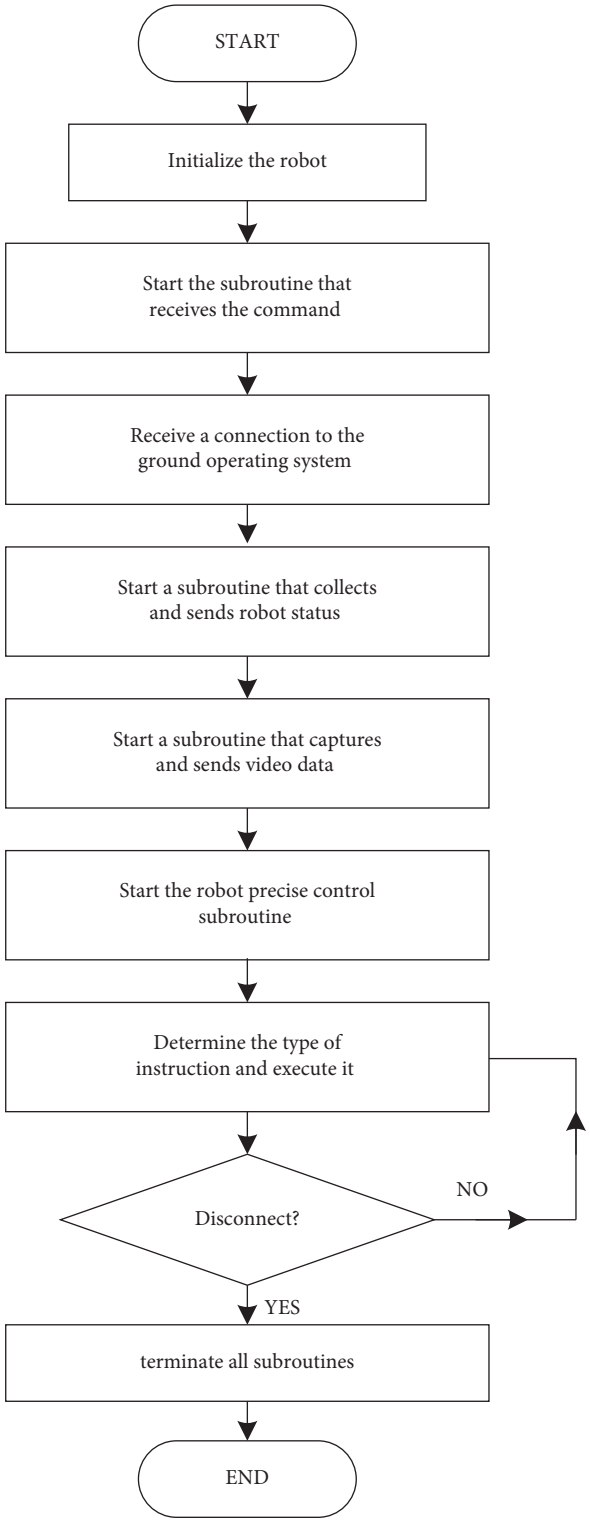


Figure 7: Control flowchart of line changing robot.

operation environment. However, the robot developed in [11] does not take into account the accurate control of the robot movement speed. The robot developed in [13] has good endurance without considering the external environment, but does not consider the performance of other aspects of the robot. The robot developed in [14]

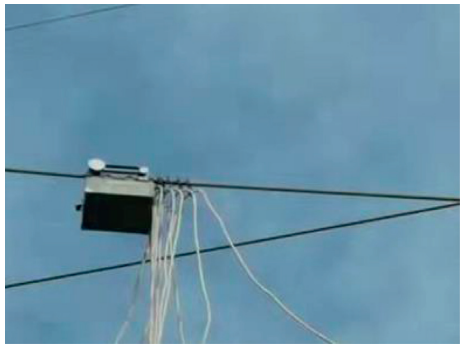


Figure 8: Practical application experiment of line changing robot.

Table 2: The experimental data of the line changing robot.

Number	Runtime (min)	Success rate (%)	Velocity (m/s)	Climbing angle	Wind	Temperature
1	28	100	0.5	5	2	−3°C
2	30	100	0.6	6	2	−5°C
3	32	100	0.8	5	2	−10°C
4	34	100	1.0	7	3	30°C
5	36	100	0.7	8	4	40°C
6	37	100	0.6	8	5	−7°C
7	40	100	0.8	9	5	−8°C
8	42	100	1.0	10	4	20°C
9	44	100	1.0	10	5	−10°C
10	45	100	1.0	10	3	40°C

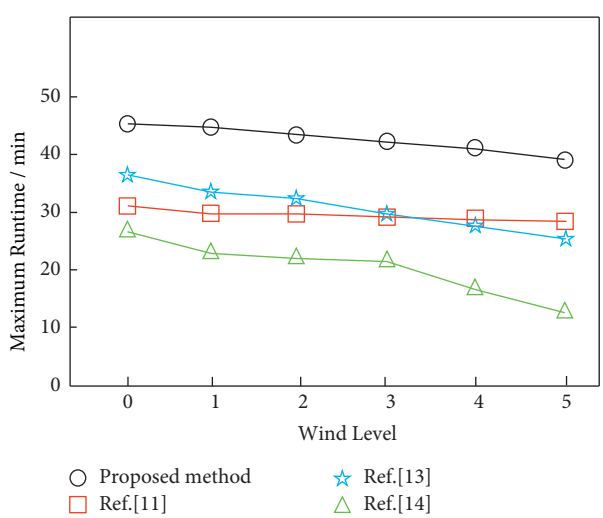


Figure 9: The relationship between the maximum runtime of different robots and the wind level.

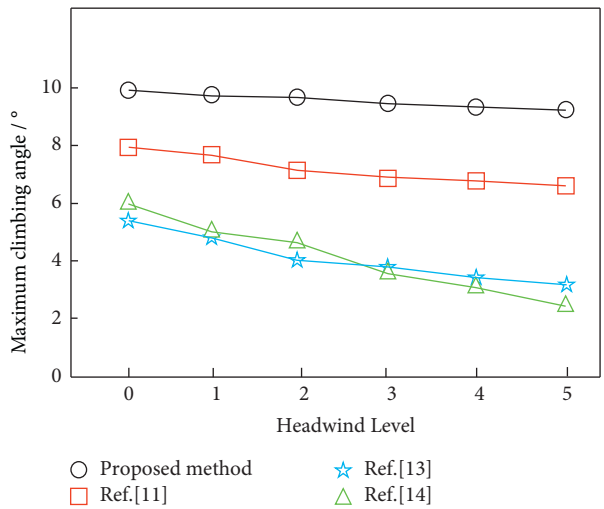


Figure 10: The relationship between the maximum climbing angle of different robots and the headwind level.

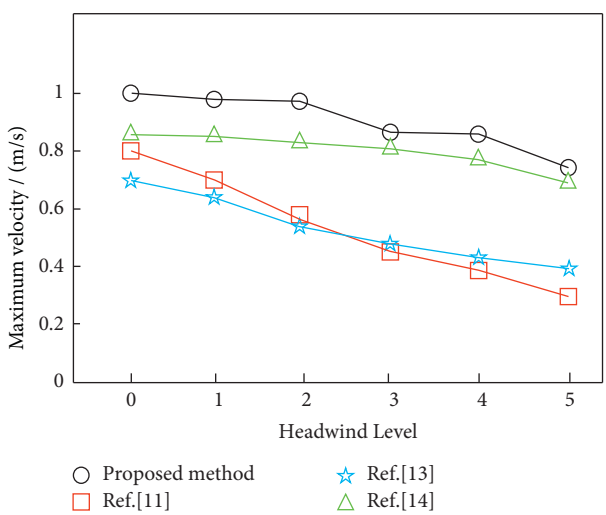


Figure 11: The relationship between the maximum velocity of different robots and the headwind level.

overemphasizes the importance of path control while neglects the optimization of the climbing ability of robot speed.

5. Conclusion

To solve the problem of low efficiency in dismantling old lines, an intelligent ropeway line changing robot design method based on lifting force control and synovial controller is proposed. The simulation results are compared between this method and other three methods. The results show that the lifting force required to be provided to the robot can be accurately calculated by using the vertical radian of the overhead line where the robot is located. By introducing the radial basis function (RBF) neural network into the synovial controller, the self-learning and adaptive robot high-precision motion control can be realized. Optimizing the shape, position, and number of different structures without changing the weight of the robot can effectively eliminate potential weak links and improve the reliability of the robot. In the future, further research will be conducted on the battery module of the line changing robot. Therefore, in the future research on related robots, the RBF neural network can be introduced into the control module to improve the control stability of the robot, and the reliability of the robot can be improved through reasonable structural adjustment and optimization. On the basis of ensuring the good function of the robot, the endurance time of the robot can be extended as far as possible to enhance the ability of its continuous work.

Data Availability

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

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work was supported by State Grid Jiangsu Electric Power Co., Ltd. (Incubate project: Research on the Intelligent Cable Method-Based Robot Used for Power Line Removing and Changing under grant JF2021022).

References

[1] J. Yerramsetti, D. S. Paritala, and R. Jayaraman, "Design and implementation of automatic robot for floating solar panel cleaning system using AI technique," in *Proceedings of the 11th International Conference of Computer Communication and Informatics (ICCCI)*, pp. 105–110, Sri Shakthi Institute of Engineeing and Technology, Coimbatore, India, 2021.

[2] D. Hrubý, D. Marko, M. Olejár, V. Cviklovič, and D. Horňák, "Comparison of control algorithms by simulating power consumption of differential drive mobile robot motion control in vineyard row," *Acta Technologica Agriculturae*, vol. 24, no. 4, pp. 195–201, 2021.

[3] W. Liao, K. Nagai, and J. Wang, "An evaluation method of electromagnetic interference on bio-sensor used for wearable robot control," *IEEE Transactions on Electromagnetic Compatibility*, vol. 62, no. 1, pp. 36–42, 2020.

[4] W. Q. Zou, X. Shu, Q. Q. Tang, and S. Lu, "A survey of the application of robots in power system operation and maintenance management," in *Proceedings of the 2019 Chinese Automation Congress (CAC)*, pp. 4614–4619, Hangzhou, China, November 2019.

[5] M. B. Khan, T. Chuthong, C. D. Do et al., "iCrawl: an inchworm-inspired crawling robot," *IEEE Access*, vol. 8, no. 32, pp. 200655–200668, 2020.

[6] N. B. David and D. Zarrouk, "Design and analysis of FCSTAR, a hybrid flying and climbing sprawl tuned robot," *IEEE Robotics and Automation Letters*, vol. 6, no. 4, pp. 6188–6195, 2021.

[7] T. He, Y. Zeng, and Z. Hu, "Research of multi-rotor UAVs detailed autonomous Inspection Technology of Transmission lines based on route planning," *IEEE Access*, vol. 7, no. 3, pp. 114955–114965, 2019.

[8] Z. P. Wang, B. He, Y. M. Zhou, K. Liu, and C. Zhang, "Design and implementation of a cable inspection robot for cable-stayed bridges," *Robotica*, vol. 39, no. 8, pp. 1417–1433, 2021.

[9] J. Katrasnik, F. Pernus, and B. Likar, "A survey of mobile robots for distribution power line inspection," *IEEE Transactions on Power Delivery*, vol. 25, no. 1, pp. 485–493, 2010.

[10] A. Papadimitrio, G. Andrikopoulos, and G. Nikolakopoulos, "On path following evaluation for a tethered climbing robot," in *Proceedings of the 46th Annual Conference of the IEEE-Industrial-Electronics-Society (IECON)*, pp. 656–661, ELECTR Network, Singapore, October 2020.

[11] C. H. F. dos Santos, M. H. Abdali, D. Martins, and C. B. Aníbal Alexandre, "Geometrical motion planning for cable-climbing robots applied to distribution power lines inspection," *International Journal of Systems Science*, vol. 52, no. 8, pp. 1646–1663, 2021.

[12] A. T. Zengin, G. Erdemir, T. C. Akinci, and S. Seker, "Measurement of power line sagging using sensor data of a power line inspection robot," *IEEE Access*, vol. 8, no. 5, pp. 99198–99204, 2020.

[13] W. Jiang, G. C. Ye, D. Zou et al., "Dynamic model based energy consumption optimal motion planning for high-voltage transmission line mobile robot manipulator," *Proceedings of the Institution of Mechanical Engineers, Part K: Journal of Multi-Body Dynamics*, vol. 235, no. 1, pp. 93–105, 2021.

[14] T. P. Nguyen, H. Nguyen, V. H. Phan, and H. Q. Thinh Ngo, "Modeling and practical implementation of motion controller for stable movement in a robotic solar panel dust-removal system," *Energy Sources, Part A: Recovery, Utilization, and Environmental Effects*, vol. 23, no. 2, pp. 57–66, 2021.

[15] X. Zhao, Z. Peng, and S. Zhao, "Substation electric power equipment detection based on patrol robots," *Artificial Life and Robotics*, vol. 25, no. 3, pp. 482–487, 2020.

[16] D. Xie, J. Liu, R. Kang, and S. Zuo, "Fully 3D-printed modular pipe-climbing robot," *IEEE Robotics and Automation Letters*, vol. 6, no. 2, pp. 462–469, 2020.

[17] L. Song, H. Wang, and P. Chen, "Automatic patrol and Inspection method for machinery diagnosis robot-sound signal-based fuzzy search approach," *IEEE Sensors Journal*, vol. 20, no. 15, pp. 8276–8286, 2020.

[18] X. Liu, X. Lu, and S. Zhao, "Kinematics and singularity analysis of an electricity pylon climbing robot," *Journal of Machine Design*, vol. 33, no. 5, pp. 7–13, 2019.

[19] M. Ahmadizadeh, A. M. Shafei, and M. Fooladi, "Dynamic modeling of closed-chain robotic manipulators in the presence of frictional dynamic forces: a planar case," *Mechanics Based Design of Structures and Machines*, no. 1, pp. 1–21, 2021.

[20] M. H. Korayem and H. Khaksar, "Estimation of critical force and time required to control the kinematics and friction of rough ellipsoidal and cubic nanoparticles using mechanics of contact surfaces," *Tribology International*, vol. 137, pp. 11–21, 2019.

[21] X. Lu, S. Zhao, D. Yu, and X. Liu, "Pylon-climber: a novel climbing assistive robot for pylon maintenance," *Industrial Robot: An International Journal*, vol. 44, no. 1, pp. 38–48, 2017.

[22] A. T. Zengin, G. Erdemir, T. C. Akinci, F. A. Selcuk, M. N. Erduran, and S. S. N. Seker, "ROSETLineBot: one-wheel-drive low-cost power line Inspection robot design and control," *Journal of Electrical Systems*, vol. 15, no. 4, pp. 626–634, 2019.

[23] T. Q. Mao, K. Huang, X. W. Zeng et al., "Development of power transmission line defects diagnosis system for UAV inspection based on binocular depth imaging technology," in *Proceedings of the 2nd International Conference on Electrical Materials and Power Equipment (ICEMPE)*, pp. 478–481, Guangzhou, China, April 2019.

[24] M. U. Sumagayan, C. Premachandra, R. B. Mangorsi, C. J. Salaan, H. W. H. Premachandra, and H. Kawanaka, "Detecting power lines using point Instance network for distribution line Inspection," *IEEE Access*, vol. 9, no. 4, pp. 107998–108008, 2021.

[25] Z. Pu, Y. Xiong, H. Wang et al., "Design and construction of a new insulator detection robot for application in 500 kV strings: electric field analysis and field testing," *Electric Power Systems Research*, vol. 173, no. 7, pp. 48–55, 2019.

[26] Y. Li, S. He, L. Xu, X. Wang, and A. Zou, "Effect of patrol robot on electric field distribution of human body in 500 kV substation," *Advanced Technology of Electrical Engineering and Energy*, vol. 39, no. 9, pp. 74–80, 2021.

Hindawi
Journal of Robotics
Volume 2022, Article ID 3664663, 8 pages
https://doi.org/10.1155/2022/3664663

Research Article

Resource Optimization Technology Using Genetic Algorithm in UAV-Assisted Edge Computing Environment

Huijuan Sun and Hongqi Xi

College of Computer and Information Technology, Henan Finance University, Zhengzhou, Henan 450046, China

Correspondence should be addressed to Huijuan Sun; sunhuijuan@hafu.edu.cn

Received 16 February 2022; Revised 18 March 2022; Accepted 19 March 2022; Published 6 April 2022

Academic Editor: Shan Zhong

As fixed edge computing systems can hardly meet the demand of mobile users for massive data processing, a computational resource allocation strategy using the genetic algorithm in UAV-assisted edge computing environment is proposed. First, a UAV-assisted mobile edge computing (MEC) system is designed to help users execute computation tasks through the UAV or relaying to the ground base station. Then, a communication model and a computation model are constructed to minimize the total system energy consumption by jointly optimizing the UAV offloading ratio, user scheduling variables, and UAV trajectory. Finally, the minimization of total system energy consumption is modeled as a nonconvex optimization problem and solved by introducing an improved genetic algorithm, so as to achieve a rational allocation of computational resources. Based on the experimental platform, the simulation of the proposed method is carried out. The results show that the total energy consumption is 650 J when the execution time is 110 s and the execution time is 17.5 s when the number of users is 50, which are both better than other comparison methods.

1. Introduction

In recent years, as cloud computing are gradually moving to the edge of the network, mobile edge computing (MEC) has emerged and developed rapidly, which mainly utilizes the spare computing capacity and storage of the edge network to execute computation-intensive and delay-sensitive tasks [1, 2]. Currently, scholars are actively trying to integrate MEC-related technologies with a wireless communication theory to better address long propagation delays, and the deployment of mobile base stations on unmanned aerial vehicle (UAV) is a possible solution to extend the coverage of wireless systems to the areas with limited available infrastructure for wireless access points [3].

UAV is a non-manned vehicle operated by remote control, with advantages of good maneuverability and small size, which is easy to carry and transport. Meanwhile, with great stability during low altitude flight, UAV can better acquire images obscured by clouds and other objects, which cannot be captured by the satellite light-sensing remote controller and other aerospace photography [4]. Due to its high maneuverability and flexibility, UAV plays an increasingly prominent role in the communication system and is widely used in aerial surveillance, imaging, and traffic [5]. To maintain the safe operation of UAVs through real-time command/control and to enable new applications with artificial intelligence, it is becoming more important to enhance the communication and computing capabilities of UAVs [6, 7].

However, UAVs are limited by size, load capacity, and power consumption, which make the communication, computing capacity, and endurance time all suffer from a large impact. In order to settle this limitation, it is of great practical significance to design resource allocation strategies and optimal control methods for UAV trajectories to make the best use of limited energy of UAV, so as to better provide high-capacity, high-speed, and low-delay services for wireless communication edge computing systems [8].

At present, there have been some studies at home and abroad for computational resource allocation strategies in edge computing environment. According to [9], the mixed integer and nonconvex optimization problem was decoupled

into two sub-problems of task offloading decision and resource optimization, which were jointly optimized by using a designed iterative algorithm. The experimental results showed that the optimization performance is comparable to that of the exhaustive search. However, this algorithm takes a longer time for optimization and the efficiency needs to be improved. The authors of [10] proposed a dynamic multi-win game model based on incomplete information to solve the task offloading and edge resource allocation problems for multiple terminal users. Further optimization is needed for multiobjective task offloading and resource allocation in complex network situations. The authors of [11] proposed a computing resource allocation scheme for MEC scenarios based on DRL networks. The cost of task offloading still needs to be optimized. In order to investigate the relationship between the cost of computation tasks and time and energy consumption, a computation offloading strategy was proposed, which reduces the execution cost by collaboratively allocating computational resources between mobile devices and edge servers [12].

With the continuous development of machine learning and UAV communication technologies, deep learning and UAV assistance have been gradually applied in computational resource allocation. Authors in Ref. [13] proposed a multiobjective optimization algorithm based on the MOACO algorithm combined with reinforcement learning, which effectively improves the quality of service in edge computing by establishing an effective cognitive agent model to evaluate the resource allocation. However, with more iteration, the time efficiency needs to be improved. In the study [14], a multiagent reinforcement learning framework based on independent learners was proposed to achieve computation offloading and resource allocation for multiple private users in IoT edge computing networks. Authors in Ref. [15] introduced agents into task offloading and proposed a new framework for agent-enabled task offloading in UAV-aided MEC to help users, UAV, and edge clouds perform computation task offloading. UAV can improve wireless communication coverage to a certain extent, but the computation offloading strategy still needs further optimization.

Based on the above analysis, a computational resource allocation strategy using the genetic algorithm in the UAV-assisted edge computing environment is proposed to address the problem that the computing capacity of existing mobile devices (computers, cell phones) cannot fully meet the users' demand for communication quality. Based on the UAV-assisted edge computing system, a minimization model of total system energy consumption with jointly optimized UAV offloading ratio, user scheduling variables, and UAV trajectory is constructed. The minimization problem is solved using an improved genetic algorithm to achieve a reasonable allocation of computational resources.

2. System Model and Optimization Objective

2.1. System Model. As shown in Figure 1, (a) UAV-assisted MEC system consists of a cellular-connected UAV and multiple ground-requesting users.

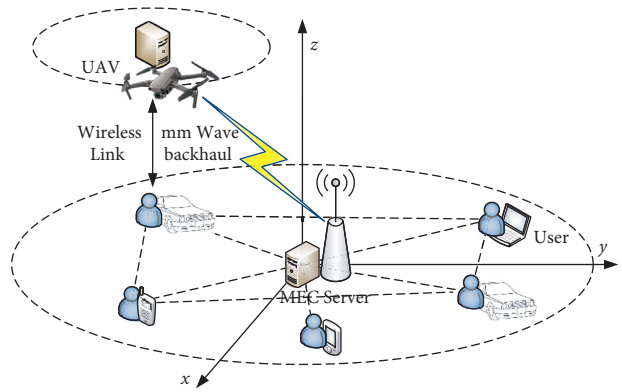


Figure 1: System model.

Both UAV and ground users have communication circuits and computing processors and are powered by batteries with a certain capacity. It is worth noting that different users have different computing capacitates. The ground base station is deployed with multiple built-in MEC servers with high computing capacity, which can be regarded as an access point [16, 17]. In this scenario, the user has delay-sensitive computation tasks that require UAV to assists in computation. The UAV is deployed with MEC servers to help users to execute the task on itself or relay to the ground base station for execution.

2.2. Communication Model. In the system model, users are fixed on the ground and the position can be denoted as q, $q = [x, y]$, where x and y are the horizontal coordinates of the user. UAV keeps flying at the same altitude $H\,(H > 0)$ and its position at i th timeslot can be expressed as $q_u = [x_u[i], y_u[i]]$, $0 \le i \le I$.

It is assumed that the wireless channel between the UAV and users is a Line of Sight (LOS) channel. Therefore, the channel power gain between the UAV and users can be expressed as follows:

$$h[i] = \delta_0 d^{-2}[i] = \frac{\delta_0}{H^2 + \left\|q_u[i] - q\right\|^2}, \quad i \in I, \tag{1}$$

where δ_0 is the channel power gain within unit distance ($d_0 = 1m$) and $d^{-2}[n]$ is the horizontal distance between the UAV and the user at i th timeslot,$i = 1, 2, \ldots, I$. $\|\cdot\|$ is the Euclidean norm.

2.3. Computation Model

2.3.1. Local Execution. When users execute the task locally, the execution time can be calculated as follows:

$$T_n^{loc}[i] = \frac{(1 - \eta_n[i]\vartheta_n[i])C_n D_n[i]}{f_n}, \tag{2}$$

where $\eta_n[i]$ is the scheduling variable of user n. $\vartheta_n[i]$ is the offloading ratio of user n at i th time slot. $D_n[i]$ denotes the bit size of the computation task of user n at ith timeslot. C_n denotes the CPU frequency required when user n executes a

task of 1 bit. f_n denotes the local computing capacity of user n.

Then, the energy consumption for local execution can be written as follows:

$$E_n^{loc}[i] = \gamma\left(1 - \eta_n[i]\vartheta_n[i]\right)D_n[i]C_n f_n^2, \tag{3}$$

where γ is a coefficient that depends on the chip architecture.

2.3.2. UAV-Assisted Computation. For UAV-assisted computation, it can be divided into two parts: the transmission process and the assisted computation. Thus, the UAV transmission time delay can be calculated as follows:

$$T_n^{tr}[i] = \frac{\eta_n[i]\vartheta_n[i]D_n[i]}{v_n[i]}, \tag{4}$$

where $v_n[i]$ is the transmission rate between user n and the UAV at i th time slot. Then, the transmission energy consumption can be written as follows:

$$E_n^{tr}[i] = P_0\frac{\eta_n[i]\vartheta_n[i]D_n[i]}{v_n[i]}, \tag{5}$$

where P_0 is the rated transmission power of the UAV.

For the UAV-assisted computation, in each timeslot, the UAV serves multiple ground users, the execution time of UAV serving user n at ith timeslot for the assisted computation can be calculated as follows:

$$T_n^{UAV}[i] = \frac{\eta_n[i]\vartheta_n[i]D_n[i]C_n}{f_n^{UAV}}, \tag{6}$$

where f_n^{UAV} is the computing capacity of the UAV when serving user n. Hence, the energy consumption for the UAV-assisted computation can be formulated as follows:

$$E_n^{UAV}[i] = \gamma'\eta_n[i]\vartheta_n[i]D_n[i]C_n f_n^{UAV2}, \tag{7}$$

where $\gamma\prime$ is a coefficient related to the chip architecture of UAV.

2.4. Optimization Objective. According to the aforementioned UAV-assisted computation model, the corresponding optimization problem is designed. It is assumed that user local execution and UAV-assisted computation are performed simultaneously, and the UAV needs to perform the computation and return the results at the same time slot. Because the size of the returned results is much smaller than that of the UAV-assisted computation task, the energy consumption for returning results can be neglected [18, 19]. The objective is to minimize the total system energy consumption by jointly optimizing the UAV offloading ratio, the user scheduling variables and the UAV trajectory. Thus, the optimization problem can be formulated as follows:

$$\Gamma = \min_{\{\eta_n[i],\vartheta_n[i],q[i]\}} \sum_{i=1}^{I} E^{\text{fly}}[i] + \sum_{i=1}^{I}\sum_{n=1}^{N}\left(E_n^{UAV}[i] + E_n^{tr}[i] + E_n^{\text{loc}}[i]\right), \tag{8}$$

$$\begin{aligned} \text{s.t.}\quad & q[1] = q[I], \\ & \frac{\|q[i+1] - q[i]\|}{\tau} \le V_{\max}, \quad i \in I, \end{aligned} \tag{9}$$

$$0 \le \vartheta_n[i] \le 1, \quad \forall n, i, \tag{10}$$

$$\begin{aligned} & \sum_{n=1}^{N} \eta_n[i] \le 1, \quad \forall i \\ & \eta_n[i] \in \{0, 1\}, \quad \forall n, i, \end{aligned} \tag{11}$$

$$v_n[i] \ge \eta_n[i] \cdot v, \quad \forall n, i, \tag{12}$$

$$\sum_{i=1}^{I}\sum_{n=1}^{N} T_n^{\text{loc}}[i]\eta_n[i] \le T, \quad \forall n, i, \tag{13}$$

$$\sum_{i=1}^{I}\sum_{n=1}^{N}\left(T_n^{tr}[i] + T_n^{UAV}[i]\right)\eta_n[i] \le T, \quad \forall n, i, \tag{14}$$

where $E^{\text{fly}}[i]$ is the energy consumption of the UAV during its flight. $V_{\max}$ is the maximum flight speed of the UAV. τ is the length of each timeslot. v is the flight speed of the UAV. Equation (9) shows the constraint of the UAV trajectory. Equation (10) defines the constraint of the offloading ratio. Equation (11) represents the user scheduling variables.

Equation (12) ensures the communication quality during offloading. Equation (13) represents the constraint of user local execution delay. Equation (14) represents the constraint of user transmission delay and UAV-assisted computation delay.

3. Computational Resource Allocation Strategy Using Improved Genetic Algorithm

Genetic algorithm is a relatively mature algorithm that searches for the global optimal solution by simulating the process of biological evolution in nature. Its parallelism is suitable for distributed edge computing. Moreover, the genetic algorithm has good global search capability, so that an appropriate solution set for resource allocation can be obtained by using the genetic algorithm with multiobjective optimization.

Therefore, this paper introduces an improved genetic algorithm, which can not only improve the search efficiency but also obtain the further optimization objective by establishing a dynamic adjustment model of crossover rate and mutation rate, while reflecting the advantages of distributed execution of tasks. In the following sections, the implementation of resource scheduling of the proposed model is addressed based on the genetic algorithm.

3.1. Basic Genetic Algorithm. The basic genetic algorithm was proposed by J. H. Holland of the University of Michigan, USA, which is the basis for other genetic algorithms. Moreover, other improved genetic algorithms are developed by adding new mechanisms. The flow chart of basic genetic algorithm is shown in Figure 2.

The basic genetic algorithm generates optimal individuals by simulating the selection, crossover, and mutation operations in biological genetic and evolutionary processes and finds the optimal solution through an adaptive search process [20, 21].

3.2. Improved Genetic Algorithm

3.2.1. Encoding and Population Initialization. The proposed method adopts binary encoding where a chromosome corresponds to a solution of the optimization problem. The defined chromosomes are encoded using real numbers. Therefore, it is necessary to convert the genes in the chromosome to be represented in binary. Moreover, the genes in the chromosome are all real numbers, which can be converted into a four-bit binary number and then be replaced in the corresponding genes. Moreover, this is the encoding method used in the improved genetic algorithm [22, 23]. For example, a chromosome sequence is (3, 5, 8) when encoded in real numbers, then the chromosome sequence is converted into (0011, 0101, 1000) after binary encoding.

However, the objective function value is calculated mainly based on real numbers, which requires reconverting the binary-coded chromosome sequence into a real-coded chromosome sequence [24]. Moreover, after crossover and mutation operations, the value of genes converted to real numbers may exceed the range of edge node numbers, so a remainder approach is used [25, 26]. For example, the total number of edge nodes is defined to be 7 and thus the binary code is {1101}, which will be converted to a decimal number of 13. The remainder of 13 divided by 7 is 6, then the corresponding edge node number of this binary code is 6.

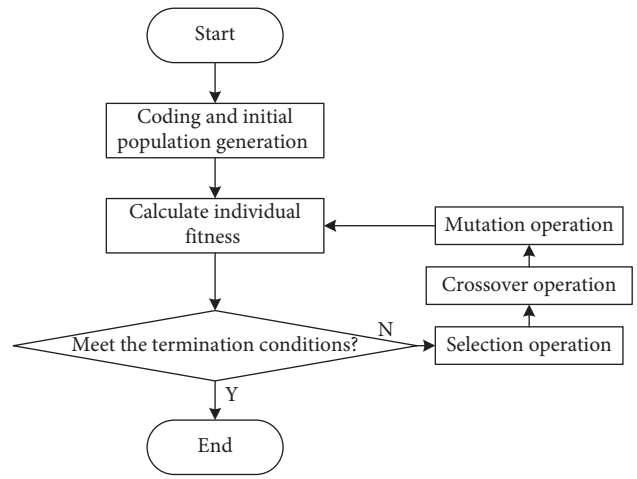


Figure 2: Flow chart of basic genetic algorithm.

Set the population size as S, the task sequence as S_T, the length as a, and the number of edge nodes in the random network topology graph as m. Each chromosome is composed of a random combination of binary numbers whose length is 4^*a, and every four binary numbers will be converted into a decimal number, which represents the number of edge nodes.

3.2.2. Fitness Function. The fitness function is the key to evaluate the direction of population evolution in task scheduling of terminal users in edge computing. In order to ensure that the fitness function can remain meaningful when the value of optimization objective Γ is equal to 0, it is defined as follows:

$$f(\Gamma) = e^{-\Gamma}. \tag{15}$$

3.2.3. Genetic Operations. The selection operation is performed by a roulette wheel. Assuming that the size of the population is S and the fitness value of an individual s is f_s, and p_s is the probability of an individual s being selected, the probability can be calculated as follows:

$$p_s = \frac{f_s}{\sum_{j=1}^{S} f_j}. \tag{16}$$

The crossover operator is used to generate new individuals by swapping the positions of chromosome genes of the parent individuals. The traditional single-point crossover method is used, in which a crossover point is set between any two adjacent genes in each chromosome sequence S_F, then a crossover point is arbitrarily selected in the first w^{-1} genes,

and all genes after that crossover point are replaced with each other [27].

In the mutation operation, each chromosome sequence S_F uses a basic position mutation, where one gene is arbitrarily selected in w number of genes and replaced by a random number between 1 to w. Thus, a new chromosome is generated and the diversity of the population can be ensured.

3.3. Steps of Task Scheduling. The steps of the improved genetic algorithm are as follows.

(1) Population initialization. The user submits a terminal request and randomly sets the initial population $g(0)$ based on the request. Set the iteration counter $t = 0$, and $g(t)$ indicates the tth generation of population.

(2) Calculate the fitness value. The fitness value of an individual is calculated according to the fitness function in the improved genetic algorithm.

(3) Judge the termination condition. When the number of evolutionary generations reaches the specified number of iterations Θ, the result should be given and an edge computation task scheduling strategy is found; otherwise, go to step (4).

(4) Selection operation. The selection operation of the populations is performed using the roulette wheel.

(5) Crossover operation. According to the crossover method in the designed improved genetic algorithm, two individuals are randomly selected among the individuals produced by the selection operation and are performed based on the crossover rate, so as to generate new individuals.

(6) Mutation operation. According to the mutation method in the improved genetic algorithm, new individuals are generated by crossover operation and the selected individuals among the new individuals are operated according to the mutation rate to generate new individuals.

(7) Compare the fitness value. If the fitness value of the mutated individuals is smaller than that of parent individuals, they will be replaced by the parent generation. If it is greater than or equal to the fitness value of the parent generation, the parent generation will be replaced. Then, a new generation of individuals $g(i+1)$ is generated.

(8) Update the iteration counter. Set the iteration counter as $t = t + 1$, then go to step (2).

4. Experiment and Analysis

The UAV trajectory and the performance of the proposed improved genetic algorithm are demonstrated by numerical simulations. The system simulation parameters are set as shown in Table 1.

The parameters of the improved genetic algorithm are set as follows. The number of iterations is 00, the crossover rate is 0.75, the mutation rate is 0.01, and the initial population size is 25.

4.1. UAV Trajectory. The UAV trajectory optimized by the improved genetic algorithm during different task execution time T is shown in Figure 3, where the red circles indicate the positions of ground users.

From Figure 3, it can be seen that when $T = 30$ s, the UAV flies along the closed-loop trajectory and the flight range is small, the UAV is far from users when offloading data, and the communication quality is poor. With the increase of T, and when $T = 60$ s and $T = 120$ s, the distance between the UAV and users decreases and the flight range expands, thus the communication quality improves. When T is large enough, the UAV continues to fly along the closed-loop trajectory and serves one user in several timeslots. In addition, it can be found that the UAV has a larger flight range and is closer to users when T is larger, and the communication quality is better.

4.2. Convergence Performance of Algorithms under Different Computation Tasks. In order to demonstrate the convergence of the improved genetic algorithm, it is compared with the traditional genetic algorithm, and the results are illustrated in Figure 4.

It is depicted from Figure 4 that the proposed improved genetic algorithm executes more tasks within the same number of iterations, and when the number of iterations exceeds 70, the amount of computation tasks tends to converge to 3450 M bit. Because it is optimized in encoding and genetic operations, the amount of computation increases. However, the process is slightly more complicated, resulting in longer processing time, and the number of iterations that tend to be stable is about 10 times higher than that of the original genetic algorithm. On the whole, the proposed algorithm can achieve effective convergence performance.

4.3. Relationship between Total System Energy Consumption with Task Execution Time. To demonstrate the performance of the total energy consumption of the proposed method, it is compared with those of the [9, 13, 15], and the results are shown in Figure 5.

As can be seen in Figure 5, with the increase of task execution time, the total energy consumption of the system also increases. Meanwhile, the proposed UAV-assisted computational resource allocation method requires lower energy consumption and has better performance than the other three methods. Moreover, the total energy consumption required by the proposed method is always the least while the task execution time is increasing. The total energy consumption is 650 J when the execution time is 110 s. As the proposed method executes the computation task with the assistance of UAV and optimizes the UAV trajectory using the improved genetic algorithm, it can dynamically and timely adjust the computation tasks so as to reduce the energy consumption. Reference [9] introduces

Table 1: System simulation parameters.

System parameters	Value
Number of users N	10
Flight altitude H	150 m
Channel power gain δ_0	−75 dB
CPU frequency C_n	1200 cycles/bit
$\gamma = \gamma'$	10^{-27}
UAV computing capacity f_n^{UAV}	1400 Hz
User local computing capacity f_n	[200, 500] Hz
User fixed transmission capacity P_0	50 dB·m
Maximum flight speed of UAV $V_{\max}$	60 m/s

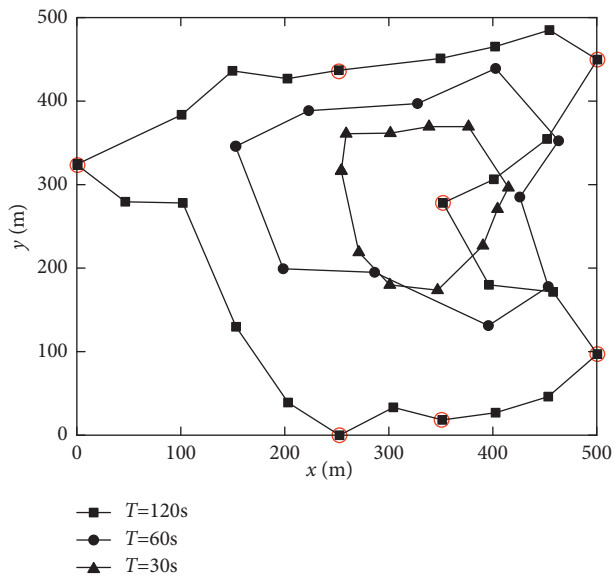


Figure 3: UAV trajectory.

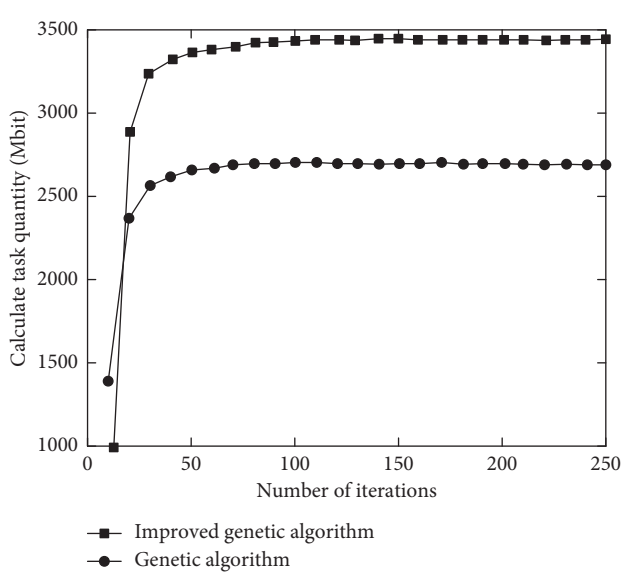


Figure 4: Convergence performance of genetic algorithms before and after improvement.

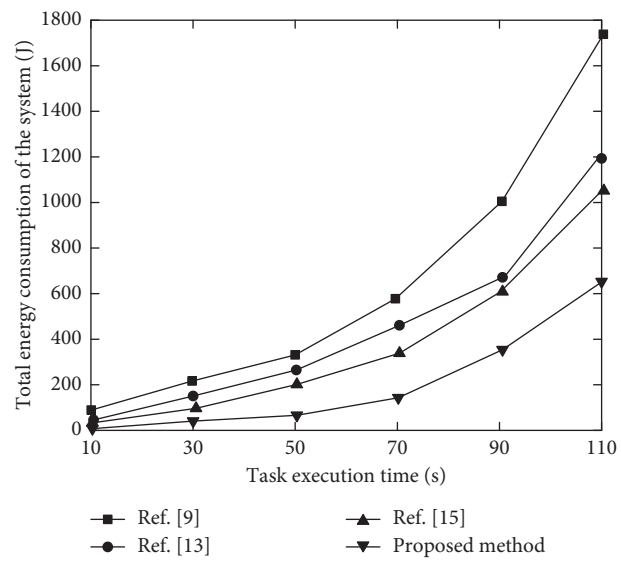


Figure 5: Total energy consumption of the system under different methods.

mixed integer and nonconvex optimization to obtain task offloading decision and resource optimization. However, the optimization method is traditional, the obtained resource allocation strategy is not reasonable enough, and the total energy consumption of the system is higher than 1600 J. Reference [13] combines MOACO algorithm and reinforcement learning for multiobjective resource allocation, which can effectively improve the quality of users' computing services, but the computational process is complex and the energy consumption is high. Similarly, [15] introduces UAV agents to assist in the offloading of computation tasks. However, it lacks an effective optimization algorithm to schedule its trajectory, so the energy consumption during UAV flight is high and its total system energy consumption is 62.5% higher than the proposed method.

4.4. Performance Comparison with Other Methods. In addition, to further validate the advantage of the proposed method in term of execution time, it is compared with [9, 13, 15], and the results are depicted in Figure 6.

It can be seen from Figure 6 that the task execution time increases continuously with more users, but the proposed method has the slowest increase and an obvious advantage in terms of execution time. The execution time is 17.5 s when the number of users is 50. The proposed method uses an improved genetic algorithm to optimize the UAV trajectory, it reduces energy consumption by lessening the blind flight time of the UAV and it offloads the tasks nearby to further reduce the transmission time. Compared with the proposed method, [15] lacks an effective algorithm to optimize the UAV trajectory, resulting in a longer flight time and a higher overall execution time, which is similar to that of [13]. Thus, the maximum execution time is more than 25 s. The optimization method in [9] is simple and easy to implement, and the execution time is shorter when the number of users is small. However, when the number of users increases, it is

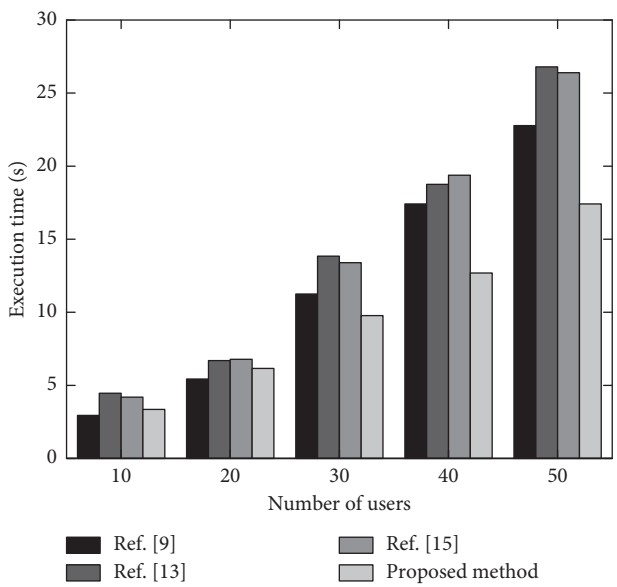


Figure 6: Execution time of different methods.

unable to handle the huge amount of data, resulting in the significant rise of the execution time.

5. Conclusion

With the explosive growth of IoT devices, cloud computing can no longer meet the demand of some time-sensitive computation tasks. For this reason, a computational resource allocation strategy using the genetic algorithm in the UAV-assisted edge computing environment is proposed. With the advantages of high mobility, low cost, and easy carrying, the UAV is introduced in the edge computing environment to assist in task offloading. The flexible scheduling of the UAV enables efficient allocation of user tasks and reduces execution time. In order to optimize the flight trajectory of the UAV, the problem of minimizing the total energy consumption of the system is solved by applying an improved genetic algorithm. Then, the optimal flight scheme can be obtained, that is, the optimal resource allocation strategy, which greatly reduces the system energy consumption and improves the transmission efficiency. The results based on the UAV simulation platform show as follows:

(1) By real number encoding, selection, crossover, and mutation operations, the performance of the genetic algorithm in finding the optimal solution is effectively improved, and the data size of computational tasks can reach 3450 M bit when the number of iterations exceeds 70.

(2) The introduction of the UAV allows dynamic response to users' computational demands and improves the efficiency of the computation offloading. When the number of users is 50, the execution time is 17.5 s, and at the same time, when the execution time is 110 s, the total energy consumption is 650 J, which effectively reduces the system energy consumption.

The UAV-assisted MEC system only considers the scenarios where UAV is connected with users or the ground. When the UAV only has limited computing capacity, the case that UAV collects users' data and offloads tasks to the ground base station for assisted computing is also worth studying.

Data Availability

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

Conflicts of Interest

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

References

[1] X. Liu, J. Yu, J. Wang, and Y. Gao, "Resource allocation with edge computing in IoT networks via machine learning," *IEEE Internet of Things Journal*, vol. 7, no. 4, pp. 3415–3426, 2020.

[2] H. C. Ke, H. Wang, H. W. Zhao, and W. J. Sun, "Deep reinforcement learning-based computation offloading and resource allocation in security-aware mobile edge computing," *Wireless Networks*, vol. 27, no. 5, pp. 3357–3373, 2021.

[3] J. Feng, L. Liu, Q. Pei, F. Hou, T. Yang, and J. Wu, "Service characteristics-oriented joint optimization of radio and computing resource allocation in mobile-edge computing," *IEEE Internet of Things Journal*, vol. 8, no. 11, pp. 9407–9421, 2021.

[4] L. Deng, P. Yang, W. Liu, L. Wang, S. Wang, and X. Zhang, "NAAM-MOEA/D-Based multitarget firepower resource allocation optimization in edge computing," *Wireless Communications and Mobile Computing*, vol. 2021, Article ID 5579857, 14 pages, 2021.

[5] Z. Song, Y. Liu, and X. Sun, "Joint task offloading and resource allocation for NOMA-enabled multi-access mobile edge computing," *IEEE Transactions on Communications*, vol. 69, no. 3, pp. 1548–1564, 2021.

[6] B. Baek, J. Lee, Y. Peng, and S. Park, "Three dynamic pricing schemes for resource allocation of edge computing for IoT environment," *IEEE Internet of Things Journal*, vol. 7, no. 5, pp. 4292–4303, 2020.

[7] Z. Wang, S. Zheng, Q. Ge, and K. Li, "Online offloading scheduling and resource allocation algorithms for vehicular edge computing system," *IEEE Access*, vol. 8, no. 9, pp. 52428–52442, 2020.

[8] W. Zhan, C. Luo, G. Min, C. Wang, Q. Zhu, and H. Duan, "Mobility-Aware multi-user offloading optimization for mobile edge computing," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 99, pp. 3341–3356, 2020.

[9] 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. 2, no. 1, pp. 1–16, 2021.

[10] B. Wang and M. Li, "Resource allocation scheduling algorithm based on incomplete information dynamic game for edge computing," *International Journal of Web Services Research*, vol. 18, no. 2, pp. 1–24, 2021.

[11] Y. Zhang, M. Zhang, C. Fan, F. Li, and B. Li, "Computing resource allocation scheme of IOV using deep reinforcement

learning in edge computing environment," *EURASIP Journal on Applied Signal Processing*, vol. 21, no. 1, pp. 1–19, 2021.

[12] S. Chen, Q. Li, M. Zhou, and A. Abusorrah, "Recent advances in collaborative scheduling of computing tasks in an edge computing paradigm," *Sensors*, vol. 21, no. 3, pp. 779–787, 2021.

[13] V. Shanmuganthan, M. Khari, N. Dey, R. G. Crespo, and Y. H. Robinson, "Enhanced resource allocation in mobile edge computing using reinforcement learning based MOACO algorithm for IIOT," *Computer Communications*, vol. 151, no. 4, pp. 355–364, 2020.

[14] X. Liu, J. Yu, Z. Feng, and Y. Gao, "Multi-agent reinforcement learning for resource allocation in IoT networks with edge computing," *China Communications*, vol. 17, no. 9, pp. 220–236, 2020.

[15] W. A. Rui, C. A. Yong, B. An, T. A. Alamoudi, and R. Nour, "Agent-enabled task offloading in UAV-aided mobile edge computing," *Computer Communications*, vol. 149, no. 2, pp. 324–331, 2020.

[16] S. Li, S. Lin, L. Cai, W. Li, and G. Zhu, "Joint resource allocation and computation offloading with time-varying fading channel in vehicular edge computing," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 3, pp. 3384–3398, 2020.

[17] J. Yan, S. Bi, and Y. Zhang, "Offloading and resource allocation with general task graph in mobile edge computing: a deep reinforcement learning approach," *IEEE Transactions on Wireless Communications*, vol. 19, no. 8, pp. 5404–5419, 2020.

[18] Z. Yang, C. Pan, J. Hou, and M. Shikh-Bahaei, "Efficient resource allocation for mobile-edge computing networks with NOMA: completion time and energy minimization," *IEEE Transactions on Communications*, vol. 67, no. 11, pp. 7771–7784, 2019.

[19] J. Zhao, Q. Li, Y. Gong, and Z. Ke, "Computation offloading and resource allocation for cloud assisted mobile edge computing in vehicular networks," *IEEE Transactions on Vehicular Technology*, vol. 68, no. 8, pp. 7944–7956, 2019.

[20] L. I. Shulei, D. Zhai, D. U. Pengfei, and T. Han, "Energy-efficient task offloading, load balancing, and resource allocation in mobile edge computing enabled IoT networks," *Science China Information Sciences*, vol. 62, no. 2, pp. 1–3, 2019.

[21] L. Yang, G. Xu, J. Ge, P. Liu, and X. Fu, "Energy-efficient resource allocation for application including dependent tasks in mobile edge computing," *KSII Transactions on Internet and Information Systems*, vol. 14, no. 6, pp. 2422–2443, 2020.

[22] Y. Guo, F. Liu, N. Xiao, and Z. Chen, "Task-based resource allocation bid in edge computing micro datacenter," *Computers, Materials and Continua*, vol. 58, no. 2, pp. 777–792, 2019.

[23] H. T. Chien, Y. D. Lin, C. L. Lai, and W. Chien-Ting, "End-to-End slicing with optimized communication and computing resource allocation in multi-tenant 5G systems," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 2, pp. 2079–2091, 2020.

[24] Y. Chen, Z. Li, B. Yang, K. Nai, and K. Li, "A Stackelberg game approach to multiple resources allocation and pricing in mobile edge computing," *Future Generation Computer Systems*, vol. 108, no. 3, pp. 273–287, 2020.

[25] M. Dai, Z. Su, Q. Xu, and N. Zhang, "Vehicle assisted computing offloading for unmanned aerial vehicles in smart city," *IEEE Transactions on Intelligent Transportation Systems*, vol. 22, no. 3, pp. 1932–1944, 2021.

[26] K. Wang, Z. Xiong, C. Lin, Z. Pan, and H. Shin, "Joint time delay and energy optimization with intelligent overclocking in edge computing," *Science China (Information Sciences)*, vol. 63, no. 04, pp. 154–169, 2020.

[27] M. Zeng and V. Fodor, "Energy minimization for delay constrained mobile edge computing with orthogonal and non-orthogonal multiple access," *Ad Hoc Networks*, vol. 98, no. 3, pp. 1–13, Article ID 102060, 2020.

Hindawi
Journal of Robotics
Volume 2022, Article ID 1739884, 8 pages
https://doi.org/10.1155/2022/1739884

Research Article

Robot Path Planning Using Improved Ant Colony Algorithm in the Environment of Internet of Things

Hongliu Huang,[1] **Guo Tan**,[2] **and Linli Jiang**[1]

[1]*School of Mathematics and Computer Science, Guangxi Science and Technology Normal University, Laibin, Guangxi 546199, China*
[2]*Experimental Training Centre, Guangxi Science and Technology Normal University, Laibin, Guangxi 546199, China*

Correspondence should be addressed to Guo Tan; 47462043@qq.com

Received 18 January 2022; Revised 16 March 2022; Accepted 17 March 2022; Published 4 April 2022

Academic Editor: Shan Zhong

It is a research topic of practical significance to study the path planning technology of mobile robot navigation technology. Aiming at the problems of slow convergence speed, redundant planning path, and easy to fall into local optimal value of ant colony algorithm in a complex environment, a robot path planning based on improved ant colony algorithm is proposed. First, the grid method is used to model the path environment, which marks each grid to make the ant colony move from the initial grid to the target grid for path search. Second, the ant colony is divided according to different planning tasks. Let some ants explore the way first, and carry out basic optimization planning for the map environment. The antecedent ants mark the basic advantage on a target value of the path with pheromone concentration so as to guide the subsequent route-finding operation of the main ant colony. Finally, in order to avoid the individual ants falling into a deadlock state in the early search, the obstacle avoidance factor is increased, the transition probability is improved, and the amount of information on each path is dynamically adjusted according to the local path information, so as to avoid the excessive concentration of pheromones. Experimental results show that the algorithm has high global search ability, significantly speeds up the convergence speed, and can effectively improve the efficiency of mobile robot in path planning.

1. Introduction

Nowadays, with the continuous integration of informatization and industrialization, the intelligent industry represented by robot technology is booming, which is also the key development field of scientific and technological innovation in various countries [1–4].

Mobile robots can be divided into three types: remote control, semiautonomy, and autonomy. Autonomous mobile robots have the ability to perceive, make decisions, and adapt to the environment, so they can meet a wider range of task requirements [5]. At present, the focus and difficulty of mobile robot research are how to make the robot complete a predetermined task autonomously in a complex environment. The process of a mobile robot completing a predetermined task autonomously is called mobile robot navigation [6–8]. The navigation of mobile robot refers to "based on the surrounding environment information perceived by the sensors carried by the mobile robot and the state information of the mobile robot, the target-oriented motion of the mobile robot can be realized safely in the environment containing a limited number of obstacles."

Only when the mobile robot accurately grasps its own position and the position of obstacles in its environment, can it safely realize the target-oriented movement and complete the navigation task [9–11]. Therefore, the research on path planning technology in mobile robot navigation technology is a research topic with practical significance.

After decades of development, researchers have proposed many effective path planning methods, such as a * algorithm, genetic algorithm, and fuzzy logic algorithm [9, 12–15]. Ant colony algorithm has fast convergence speed,

but it also has some problems, such as algorithm search stagnation, and is easy to fall into local optimization. In order to solve these problems, researchers have proposed many improvement strategies [16–19]. Reference [20] analyzed the influence of pheromone Volatilization Coefficient on the optimization ability of ant colony algorithm and proposed an adaptive pheromone Volatilization Coefficient to improve the global searchability. In order to avoid falling into local optimization, [21] considered the path length and time of each ant and proposed an adaptive ant colony optimization algorithm to balance the optimal path and planning time. Reference [22] proposed a free step ant colony algorithm and designed the corresponding local pheromone update rules. The simulation showed that the path found by this algorithm is shorter and the convergence is better. Reference [23] designed a path planner composed of heterogeneous ant colonies. Different kinds of ant colonies have different visual ranges and motion speeds and modified the state transition rules and pheromone update method. Reference [24] uses the genetic algorithm to generate the initial path to reduce the blindness of the ant colony algorithm in the initial search and then uses the ant colony algorithm to further optimize the initial path. Reference [25] obtained a series of suboptimal solutions by roughly searching the initial pheromone distribution of the ant colony algorithm and then obtained the optimal solution by the ant colony algorithm. Reference [26] studied the updating strategy of residual pheromone and proposed an improved ant colony algorithm to solve the time-dependent road network planning problem. Reference [27] used the search mode of the visual field to limit the search range, adopted the deadlock-free mechanism for the deadlock problem, and proposed an improved pheromone update algorithm. Reference [28] introduced the evaluation function and bending suppression operator of a * algorithm, improved the heuristic information of the ant colony algorithm, and speeded up the convergence speed. The results showed that the improved ant colony algorithm is more efficient. Most of the above improved algorithms are committed to improving the search efficiency of the ant colony algorithm and getting the shortest search path possible, but they have not studied other optimal factors of the path.

Based on the above analysis, aiming at the problems of slow convergence speed and path redundancy of ant colony algorithm, a robot path planning method based on an improved ant colony algorithm is proposed. In the proposed method, the working area of the mobile robot is modeled by the grid method, and each grid is marked to make the ant colony move from the initial grid to the target grid for path search. The contribution of the proposed method is summarized as follows:

(1) An adaptive strategy of ant population division is proposed, which divides the ant population according to different planning tasks, allows some ants to explore the way first, and makes basic optimization planning for the map environment so that the mobile robot can quickly find the optimal path.

(2) By increasing the obstacle avoidance factor and improving the transition probability, it greatly avoids the individual ant falling into a deadlock state in the early search. By dynamically adjusting the amount of information on each path according to the local path information, the excessive concentration of pheromones is avoided, which speeds up the speed of path planning.

2. Introduction to Problem Model

When the robot performs a task, its working area is divided into grids. In size 30×30 grid map as an example. The robot can pass through the white area normally, and black is an obstacle. When the robot does not encounter obstacles, it can move in any direction at its current position and cannot return. Grid coordinates are as follows:

$$\begin{aligned} x &= \mathrm{mod}(i, N) - 0.5, \\ y &= N + 0.5 - \mathrm{ceil}\left(\frac{i}{N}\right), \end{aligned} \tag{1}$$

where N is the number of rows and columns of the grid map and i is the grid sequence number, as shown in Figure 1.

In order to ensure that the mobile robot can effectively avoid obstacles, the obstacles are expanded according to the physical size. The expanded size of the obstacles is the sum of the radius of the mobile robot and the safe reserved distance so the mobile robot can be regarded as a particle. If there are irregular obstacles in the working environment, when it fills one or more grids, the additional part is still recorded as a grid.

3. Global Path Planning Based on Improved Ant Colony Algorithm

Ant colony algorithm is easy to fall into local optimization and slow convergence. In order to find the optimal path and avoid blind search, a global path planning method based on an improved ant colony algorithm is proposed in this article.

3.1. Ant Population Adaptive Strategy. In the classical ant colony algorithm, the task of each ant is roughly the same. In addition to the sharing of information, the behavior of ants is relatively independent. This method has brought many benefits in the single-objective programming problem, but it is insufficient if still used in the multiobjective programming problem. In the way of independent action, ant individual needs to measure multiple target values to release pheromones independently, so it needs to obtain an overall evaluation value in one way, which is the standard to release the corresponding concentration of pheromones. In this way, the advantages and disadvantages of various target values in the information left for subsequent ants are mixed together, and the information is relatively vague. It is unable to give the different advantages and disadvantages of the path in the two target values, which is not conducive to the distinction and selection of subsequent ants and reduces the

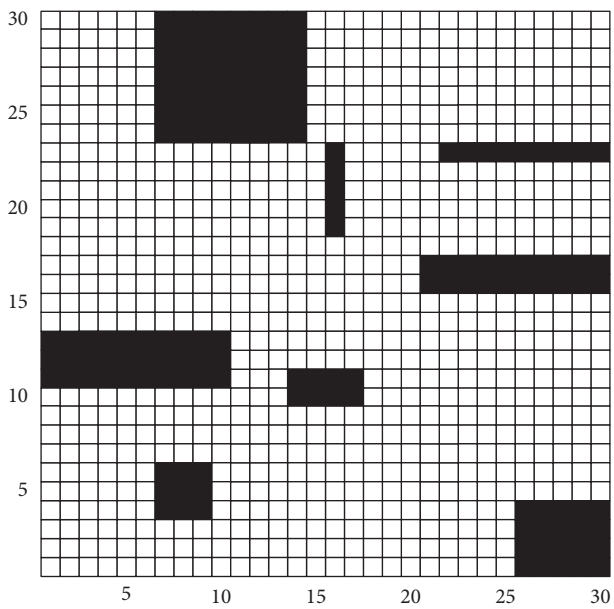


Figure 1: Grid map.

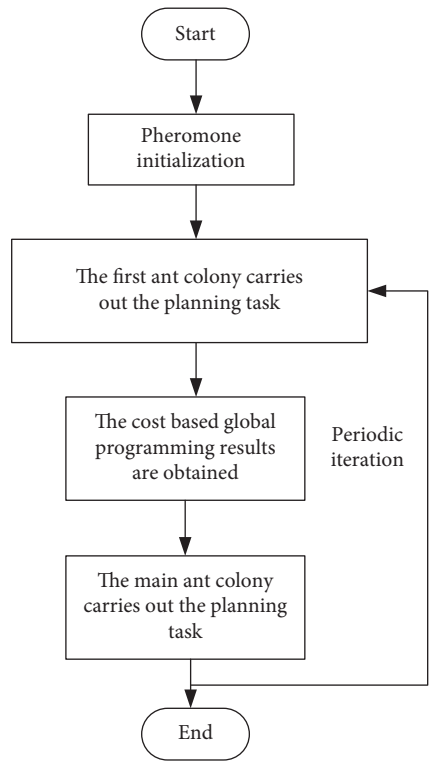


Figure 2: Algorithm flow.

diversity of understanding to some extent. At the same time, there is such a situation that in a map, the distribution of path cost is uneven, the cost values between different regions are very different, but the difference between the path lengths of different sections is almost the same. At this time, if one set of criteria is used to determine the distribution of pheromones, it will inevitably lose its adaptability. Therefore, in order to adapt to multiple target values, an adaptive strategy for ant population division is proposed. By dividing the ant population according to different planning tasks, let some ants explore the way first, carry out basic optimization planning for the map environment, and then let the main ant group carry out pathfinding operation based on the planning results obtained by the previous ant group. In other words, let the antecedent ants mark the basic advantage of a target value of the path with pheromone concentration so as to bring planning guidance to the subsequent pathfinding operation of the main ant group. As mentioned above, the proposed improved algorithm splits the ant colony as follows:

$$\text{AntNum}\begin{cases}\text{pre Antecedent ant colony}\\ \text{AntNum} - \text{pre Main pathfinding ant colony}\end{cases}, \tag{2}$$

where pre is the number of antecedent ant groups. The antecedent ant group will optimize according to the target value of path traffic cost to obtain the result of global cost planning. Then, based on the cost planning results, the main pathfinding ant group optimizes the path by comprehensive indicators. In this way, multiple target values have obvious advantages in their respective target fields, which can better reflect the advantages and disadvantages of different paths. Therefore, the pathfinding process of the algorithm is generally divided into two steps, as shown in Figure 2.

3.2. Transition Probability Improvement. Due to the limitation of the algorithm tabu table, ants can only move forward and cannot go back in the process of finding the path, which makes a large number of ants fall into a locked state in the early search, and finally "lose," that is, stop the search without reaching the end. As shown in Figure 3, for the three paths A, B, and C, because the ants fall into deadlock and stay at the last searched path node, they do not reach the target point safely and correctly, and the path search efficiency is low. Especially when the environment scale becomes larger and the terrain is complex, the deadlock phenomenon of the ant colony algorithm is more serious.

In the t-th iteration, the state transition probability of the ant k selecting the next node j from the current node i is as follows:

$$p_{ij}^{k}(t) = \begin{cases} \dfrac{\left[\tau_{ij}(t)\right]^{\alpha}\left[n_{ij}(t)\right]^{\beta}}{\sum_{j\in C_k}\left[\tau_{ij}(t)\right]^{\alpha}\left[n_{ij}(t)\right]^{\beta}}, & j \in C_k, \\ 0, & j \notin C_k, \end{cases} \tag{3}$$

where C_k represents the set of all reachable path nodes in the next step and α is the information heuristic factor. The larger the value α, the stronger the guiding role of the pheromone. β is the expected heuristic factor. τ_{ij} is the pheromone concentration of the path and n_{ij} is the heuristic function.

Deadlock ants are inevitable in the ant colony algorithm. The fundamental reason is that when searching the path, ants will inevitably encounter the situation that the surrounding nodes have passed or there are obstacles. Ants

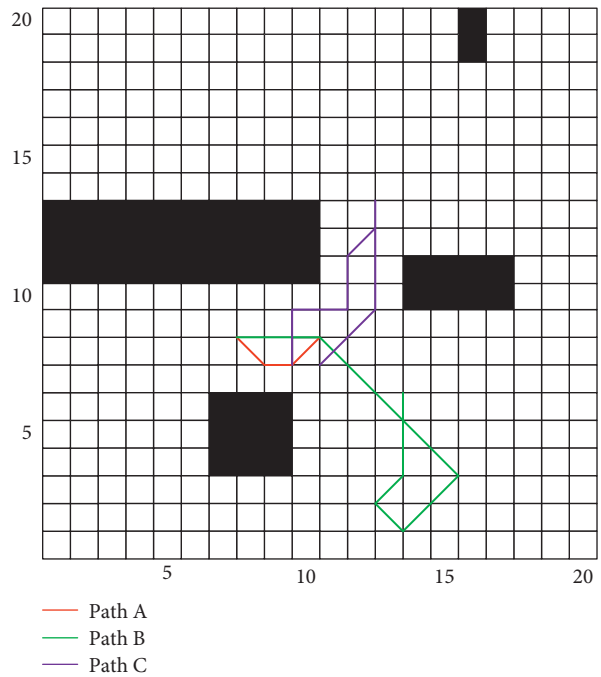


Figure 3: Algorithm flow.

cannot pass through and can only stay in place. Therefore, the obstacle avoidance strategy is adopted at the initial stage of search, and the obstacle avoidance factor s_j is introduced as follows:

$$p_{ij}^{k}(t) = \begin{cases} \dfrac{\left[\tau_{ij}(t)\right]^{\alpha}\left[n_{ij}(t)\right]^{\beta}\left[s_j\right]^{\varepsilon}}{\sum_{j\in C_k}\left[\tau_{ij}(t)\right]^{\alpha}\left[n_{ij}(t)\right]^{\beta}\left[s_j\right]^{\varepsilon}}, & j \in C_k, \\ 0, & j \notin C_k, \end{cases} \tag{4}$$

$$s_j = \frac{A_j - O_j - L_j}{A_j}, \tag{5}$$

where O_j is the total number of grids adjacent to the node j and with obstacles and L_j refers to the total number of grids adjacent to the node j and restricted by the tabu table. A_j is the total number of grids adjacent to the node j and ε is the obstacle avoidance coefficient, which takes a small positive number. By adding the obstacle avoidance strategy, the ants avoid the surrounding obstacles as much as possible every time they iterate to search the path, and the number of deadlock ants is significantly reduced. Parameter adaptive pseudorandom transfer strategy is adopted as follows:

$$j = \begin{cases} \arg\max\left(\left[\tau_{ij}(t)\right]^{\alpha}\left[n_{ij}(t)\right]^{\beta}\left[s_j\right]^{\varepsilon}\right), & q \le q_0, \\ p_{ij}^{k}, & \text{else}, \end{cases} \tag{6}$$

$$q_0 = \delta\frac{N_{\max} - N_m}{N_{\max}},$$

where $N_{\max}$ is the maximum number of iterations, N_m is the current number of iterations, and δ is the adjustment coefficient, the value range of which is $(0.5, 1)$.

3.3. Pheromone Update. In the standard ant colony optimization algorithm, set $\tau_{ij}(0) = C$, where C is a constant. In order to accelerate the early convergence speed of the algorithm, this article changes the method of selecting $\tau_{ij}(0)$ as a constant and proposes an initialization allocation of τ_{ij} according to the proportion of the local path length d_{ij} in the lengths of all connected paths between points i and j.

$$\begin{cases} \tau_{ij}(0) = Z \cdot M_j, \\ Z = \dfrac{d_{ij}}{\text{average}\sum_{i,j\in n} d_{ij}}, \end{cases} \tag{7}$$

where d_{ij} represents the distance between two nodes i and j, $\text{average}\sum_{i,j\in n} d_{ij}$ represents the average distance of all paths directly connected to nodes i and j, and M_j represents the number of paths connected to node j.

The initialization allocation method of τ_{ij} comprehensively considers the ratio Z of the length of a path to the average length of the paths in the local region and the number of paths M_j that ants can choose next. Figure 4 shows the path diagram of two points connected. The more the paths to be selected, the greater the diversity of algorithm solutions. The calculation formula of $\text{average}\sum_{i,j\in n} d_{ij}$ is as follows:

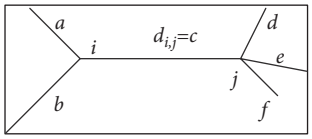


Figure 4: Path connected with two points.

$$\text{average} \sum_{i,j\in n} d_{ij} = \frac{a+b+c+d+e+f}{6}. \tag{8}$$

It is called an iterative process that all ants in the ant colony with population W move once. After n iterative processes, all ants complete a cycle. At this time, the pheromone on each path that the ants pass through needs to be adjusted. After the ants complete a cycle, it will produce an optimal solution l_{Best} and the worst solution l_{Worst} in the current cycle:

$$l_{s\longrightarrow g} \leq \frac{l_{\text{Best}} + l_{\text{Worst}}}{2}. \tag{9}$$

After determining the path that meets the update conditions, update the pheromone of each path according to formula (10). ρ is pheromone residue operator and $\Delta\tau_{ij}$ represents the amount of pheromone left by ants on the path $i \longrightarrow j$:

$$\tau_{ij}(t+1) = \rho\tau_{ij}(t) + \Delta\tau_{ij}. \tag{10}$$

Among them,

$$\Delta\tau_{ij} = \begin{cases} \varepsilon\dfrac{Q}{d_{ij}}, & \text{If the ant passes the path} \quad i \longrightarrow j, \\ 0, & \text{else}, \end{cases} \tag{11}$$

$$\varepsilon = \begin{cases} \dfrac{d_j^{\min}}{d_{ij}}, & M_j \geq M_i, \\ \dfrac{d_i^{\min}}{d_{ij}}, & M_j < M_i, \end{cases} \tag{12}$$

where d_{ij} represents the distance between nodes i and j, $d_j^{\min}$ represents the shortest path distance among all paths connected to node j, M_i represents the number of paths connected to node i, M_j represents the number of paths connected to node j, andM_i and M_j do not include path $i \longrightarrow j$. Formula (12) updates the pheromone of each path according to different conditions by using the local path information of the area around each path. It can be seen that this algorithm only allows some ants to find relatively short paths to update pheromones. According to the path length obtained by ants in this cycle, select to increase the amount of information of some better paths and dynamically adjust the amount of information on each path according to the local path information so as to avoid the excessive concentration of pheromones. This algorithm accelerates the convergence speed and optimization ability of the algorithm and avoids premature convergence of the algorithm.

3.4. Specific Implementation Steps of Improved Algorithm. The improved ant colony algorithm divides the ant colony into two parts. Each part has its own planning task. On the whole, the steps of the algorithm are as follows:

Step 1: Set the corresponding parameters required by the algorithm and read the grid map information, including the grid map scale and some content information. The required parameters include the following: grid size, number of ant populations, number of antecedent ant populations, number of iterations, initial amount of pheromone, and pheromone weight.

Step 2: Initialize all ants, including the ant taboo table, path table, path length, total traffic cost, and initial position of ants.

Step 3: The cycle iteration starts, and the antecedent ant group starts planning tasks.

Step 4: The antecedent ants calculate the transfer probability according to the formula, then select the walking node, and update the path table, path length, total communication cost, and other information. If the ant reaches the destination, the pheromone concentration on the path is updated and volatilized according to the pheromone update and volatilization rules.

Step 5: If all the antecedent ant groups reach the destination, continue the planning task of the main ant group; otherwise, turn to Step 4 to continue to find the way.

Step 6: The main ant colony carries out the planning task. According to the pheromone concentration left by the antecedent ant, the transfer probability is calculated according to formula (4). The path is selected by combining the two target values, and the relevant information table and variables are updated after selecting the next node. If the ant reaches the destination, the pheromone concentration on the path is updated and volatilized according to the pheromone update and volatilization rules.

Step 7: If all ants complete the planning task, retain the global optimal path information, compare the obtained paths, and give corresponding rewards and punishments according to the advantages and disadvantages of the paths.

Step 8: If all the iteration cycle has been completed at this time, the planning result is output and the algorithm ends. Otherwise, turn to Step 3, reset the ant information and continue the iteration.

4. Example Verification and Discussion

4.1. Simulation Environment Setting and Parameter Selection. In this article, MATLAB R2016a simulation software is used to simulate the path planning process of the mobile robot on the ground with obstacles. Because Roulette is used in the simulation process of the ant colony algorithm, the results will be different each time. In the simulation, the method of averaging multiple experiments is adopted. The simulation

Table 1: Initialization parameters.

Parameter	Value
ρ	0.2
C	15
σ	1
y	100
l	1
W	80
Q	100

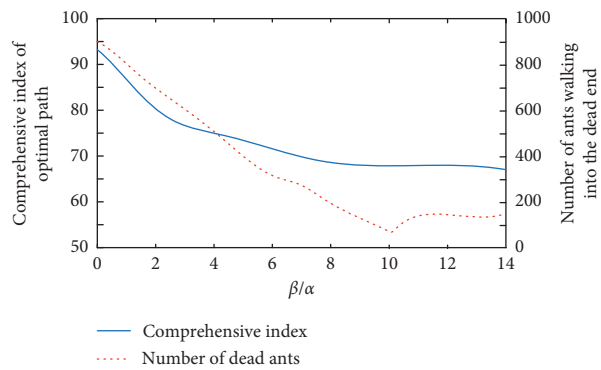


Figure 5: β/α variation curve.

Table 2: Comparison of 30 m × 30 m environmental experiment results.

Algorithm	Optimal path length (m)	Number of path nodes	Number of convergence iterations	Algorithm running time (s)	Robot running time (s)
Ant colony algorithm	44.623	21	40	22.368	150.876
Reference [26]	42.569	12	23	20.157	127.148
Reference [28]	42.357	11	20	20.845	125.758
Proposed algorithm	41.248	9	8	17.236	120.347

environment is Windows 10 64-bit operating system, Intel Core I5-3210m CPU @ 2.50 GHz processor, 8 GB memory.

At present, there is no perfect theoretical method for the parameter selection of the ant colony algorithm. The usual method is to take values according to experience and compare them through experiments to obtain better parameter values. In this article, the parameter selection experiment is carried out on a 30m × 30 m map. Because the main parameters of the ant colony algorithm are α and β, this article only shows the value selection process of α and β. The values of other parameters are shown in Table 1 according to experience and experimental simulation comparison.

Change α and β based on the above values. α reflects the importance of the inspiration of the previous ant colony to the current ant, and β reflects the importance of the inspiration of the current local environment to the ant. The values of the two are relative, so they must be combined to change. In order to facilitate the experiment, let $\alpha = 1$, so the value of β/α equals the value of β. The simulation results show that when α is large and β is small, the ant search path is chaotic, many ants are entering the dead end, and the number of ants moving towards the target grid is small, so the algorithm is easy to fall into local optimization. When the value of β/α is about 10, the algorithm has a wide search range and obtains better results, as shown in Figure 5. When the value of β/α continues to increase, the convergence speed of the algorithm slows down, and it is easy to fall into the local optimal situation only under the guidance of local heuristic information. Therefore, take $\alpha = 1$, $\beta = 10$.

4.2. Comparison of Simulation Results. At 30m × 30 m complex environment, the proposed algorithm, ant colony algorithm, the algorithm in literature [26], and the algorithm in literature [28] are compared and analyzed. Set the grid size to 1 m × 1 m; the average speed of the mobile robot is 0.5 m/s. Considering safety factors, set the deceleration of the mobile robot to 0.1 m/s when reaching each node. In addition, the turning time is set to 1 s at each node. The

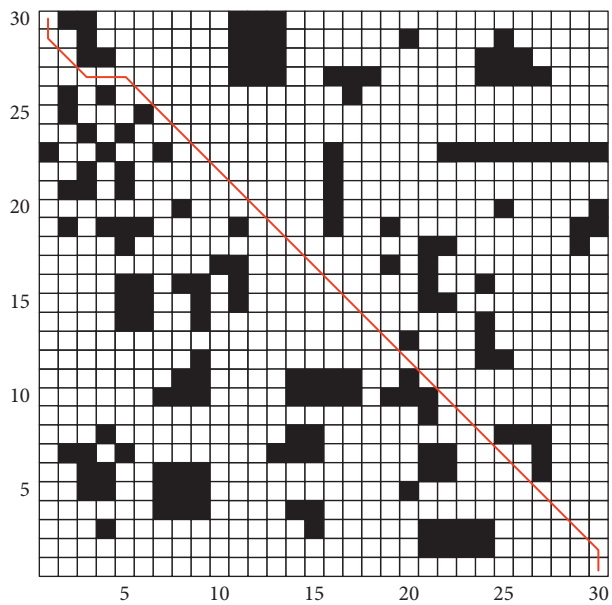


Figure 6: Improved algorithm planning path.

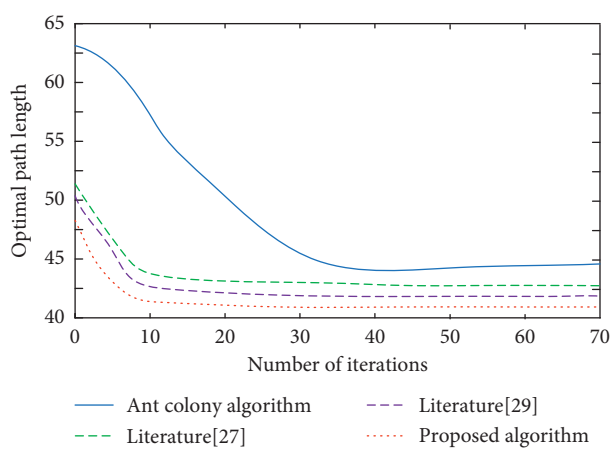


Figure 7: Comparison of convergence curves.

experimental data are shown in Table 2. The basic ant colony algorithm finds the global optimal value of 44.623 m in the 40th iteration. In [26], the algorithm converged to 42.569 m in the 23rd iteration, and in [28], the algorithm converged to 42.357 m in the 20th iteration. The proposed algorithm converged in the 8th iteration and found the global optimal value of 41.248 m. Compared with [26] and [28], the number of iterations of the algorithm is reduced by 65.22% and 60.00%, respectively, the number of turns is reduced by 25.00% and 18.18%, respectively, the path length is reduced by 3.10% and 2.62%, respectively, and the driving time of mobile robot is reduced by 5.35% and 4.30%, respectively [29]. Figure 6 shows the path planning diagram of the improved algorithm, Figure 7 shows the comparison of the convergence curves of the four algorithms, and Figure 8 shows the comparison of the number of high-quality solutions of the three algorithms.

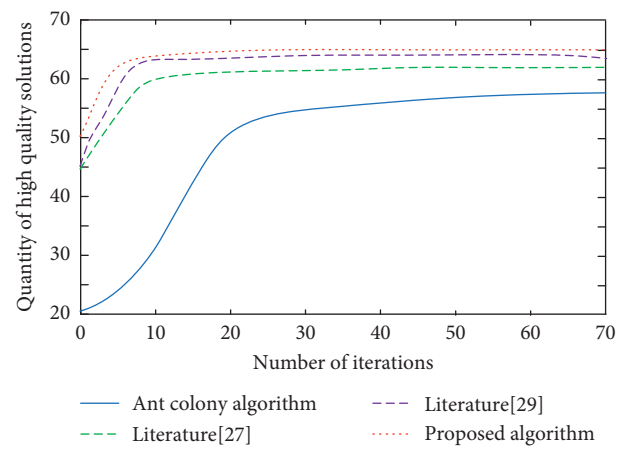


Figure 8: Comparison of the number of high-quality solutions of the four algorithms.

Through comparison, it can be seen that the improved algorithm in this article still achieves good results in improving deadlock ants in complex environment and greatly reduces the number of deadlock ants. This shows the superiority of this algorithm, which still has high global performance and fast convergence in complex environment. Due to the reduction of turning times and driving time, the energy consumption of mobile robot is most likely to be reduced. The larger the environmental scale is, the more complex the terrain is, and the more obvious the advantages are.

5. Conclusion

In order to solve the problems of slow convergence speed and easy to fall into local optimal value in path planning, a robot path planning method based on an improved ant colony algorithm is proposed. The main improvements are as follows: the grid method is used for modeling, and each grid is marked. Let the antecedent ants mark the basic advantage on a target value of the path with pheromone concentration so as to guide the subsequent pathfinding operation of the main ant group. The obstacle avoidance factor is added to improve the transition probability, and the amount of information on each path is dynamically adjusted according to the local path information to avoid the excessive concentration of pheromones. Experiments show that compared with the comparative references, the iteration times of the proposed algorithm are reduced by 65.22% and 60.00%, respectively, and the travel time of the mobile robot is reduced by 5.35% and 4.30%, respectively, which can effectively improve the work efficiency of the mobile robot in path planning.

In the future research, we will study how to complete efficient and high-quality path planning tasks in a dynamic environment. Moreover, a neural network can be used to dynamically calculate various parameters and weights to make the planning more intelligent.

Data Availability

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

Conflicts of Interest

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

Acknowledgments

This work was supported by the National Natural Science Fund Project (No. 42065004) and the project for improving the basic scientific research ability of young and middle-aged teachers in Guangxi Colleges and Universities (No. 2019KY0868).

References

[1] L. Xie, L. Yan, X. Zhang, and H. Zhang, "A security situation assessment model of information system for smart mobile devices," *Wireless Communications and Mobile Computing*, vol. 2020, Article ID 8886516, 11 pages, 2020.

[2] H. Choset, "Coverage for robotics–a survey of recent results," *Annals of Mathematics and Artificial Intelligence*, vol. 31, no. 1, pp. 113–126, 2001.

[3] S. Aggarwal and N. Kumar, "Path planning techniques for unmanned aerial vehicles: a review, solutions, and challenges," *Computer Communications*, vol. 149, no. 5, pp. 270–299, 2020.

[4] H. Jahanshahi, M. Jafarzadeh, N. Naeimeh, V.-T. Pham, V. H. Van, and Q. N. Xuan, "Robot motion planning in an unknown environment with danger space," *Electronics*, vol. 8, no. 2, pp. 201–212, 2019.

[5] B. Sahu, P. K. Das, M. R. Kabat, and R. Kumar, "Multi-robot cooperation and performance analysis with particle swarm optimization variants," *Multimedia Tools and Applications*, vol. 5, no. 9, pp. 1–24, 2021.

[6] R. Liu and X. Zhang, "A review of methodologies for natural-language-facilitated human–robot cooperation," *International Journal of Advanced Robotic Systems*, vol. 16, no. 3, pp. 172–179, 2019.

[7] G. S. Chirikjian, "Design and analysis of some non-anthropomorphic, biologically inspired robots: an overview," *Journal of Robotic Systems*, vol. 18, no. 12, pp. 701–713, 2001.

[8] B. C. Mohan and R. Baskaran, "A survey: ant colony optimization based recent research and implementation on several engineering domain," *Expert Systems with Applications*, vol. 39, no. 4, pp. 4618–4627, 2012.

[9] B. K. Patle, A. Pandey, D. R. K. Parhi, J. Anne, and B. L. Ganesh, "A review: on path planning strategies for navigation of mobile robot," *Defence Technology*, vol. 15, no. 4, pp. 582–606, 2019.

[10] B. Velusamy and S. C. Pushpan, "A review on swarm intelligence based routing approaches," *International Journal of Engineering and Technology*, vol. 9, no. 3, pp. 182–195, 2019.

[11] Z. Masoumi, J. Van Genderen, and A. Sadeghi Niaraki, "An improved ant colony optimization-based algorithm for user-centric multi-objective path planning for ubiquitous environments," *Geocarto International*, vol. 36, no. 2, pp. 137–154, 2021.

[12] V. Sangeetha, R. Krishankumar, K. S. Ravichandran, and S. Kar, "Energy-efficient green ant colony optimization for path planning in dynamic 3D environments," *Soft Computing*, vol. 25, no. 6, pp. 4749–4769, 2021.

[13] V. Sangeetha, K. S. Ravichandran, S. Shekhar, and M. T. Anand, "An intelligent gain-based ant colony optimisation method for path planning of unmanned ground vehicles," *Defence Science Journal*, vol. 69, no. 2, pp. 167–172, 2019.

[14] L. Wang, C. Luo, M. Li, and J. Kai, "Trajectory planning of an autonomous mobile robot by evolving ant colony system," *International Journal of Robotics and Automation*, vol. 32, no. 4, pp. 33–43, 2017.

[15] X.-Y. Zhang, M. Wu, J. Peng, and K.-C. Lin, "Target attraction based ant colony for dynamic path planning of rescue robot," *Journal of System Simulation*, vol. 9, no. 5, pp. 48–56, 2011.

[16] M. R. Jabbarpour, H. Zarrabi, J. J. Jung, and P. Kim, "A green ant-based method for path planning of unmanned ground vehicles," *IEEE access*, vol. 5, no. 11, pp. 1820–1832, 2017.

[17] C. Sahu, D. R. Parhi, and P. B. Kumar, "An approach to optimize the path of humanoids using adaptive ant colony optimization," *Journal of Bionics Engineering*, vol. 15, no. 4, pp. 623–635, 2018.

[18] M. J. Rostami, A. Zarandi, and S. M. Hoseininasab, "MSDP with ACO: a maximal SRLG disjoint routing algorithm based on ant colony optimization," *Journal of Network and Computer Applications*, vol. 35, no. 1, pp. 394–402, 2012.

[19] X. Jiang, Z. Lin, T. He, X. Ma, S. Ma, and S. Li, "Optimal path finding with beetle antennae search algorithm by using ant colony optimization initialization and different searching strategies," *IEEE Access*, vol. 8, no. 12, pp. 15459–15471, 2020.

[20] C. C. Yuan, Y. Wei, J. Shen et al., "Research on path planning based on new fusion algorithm for autonomous vehicle," *International Journal of Advanced Robotic Systems*, vol. 17, no. 3, pp. 172–181, 2020.

[21] J. Liu, J. Yang, H. Liu, and X. Tian, "An improved ant colony algorithm for robot path planning," *Soft Computing*, vol. 21, no. 19, pp. 5829–5839, 2017.

[22] M. R. Zeng, L. Xi, and A. M. Xiao, "The free step length ant colony algorithm in mobile robot path planning," *Advanced Robotics*, vol. 30, no. 23, pp. 1509–1514, 2016.

[23] J. Lee, "Heterogeneous-ants-based path planner for global path planning of mobile robot applications," *International Journal of Control, Automation and Systems*, vol. 15, no. 4, pp. 1754–1769, 2017.

[24] L. Wang, C. Luo, M. Li, and J. Cai, "Trajectory planning of an autonomous mobile robot by evolving ant colony system," *International Journal of Robotics and Automation*, vol. 32, no. 4, pp. 1500–1515, 2017.

[25] W. Deng, R. Chen, B. He, Y. Liu, L. Yin, and J. Guo, "A novel two-stage hybrid swarm intelligence optimization algorithm and application," *Soft Computing*, vol. 16, no. 10, pp. 1707–1722, 2012.

[26] H. Fa-Mei, X. Yi-na, W. Xu-ren, X. Meng-bo, and X. Zi-han, "An improved ant colony algorithm for solving time-dependent road network path planning problem," in *Proceedings of the 2019 6th International Conference on Information Science and Control Engineering (ICISCE)*, pp. 126–130, IEEE, Shanghai, China, December 2019.

[27] L. Wang, J. Kan, J. Guo, and C. Wang, "3D path planning for the ground robot with improved ant colony optimization," *Sensors*, vol. 19, no. 4, pp. 815–824, 2019.

[28] X. Dai, S. Long, Z. Zhang, and D. Gong, "Mobile robot path planning based on ant colony algorithm with A* heuristic method," *Frontiers in Neurorobotics*, vol. 13, no. 5, pp. 15–26, 2019.

[29] A. Majumder, "Contrast enhancement of multi-displays using human contrast sensitivity," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*, pp. 377–382, IEEE, Piscataway, NJ, USA, July 2005.

Hindawi
Journal of Robotics
Volume 2022, Article ID 9397783, 9 pages
https://doi.org/10.1155/2022/9397783

Research Article

An UAV-Assisted Edge Computing Resource Allocation Strategy for 5G Communication in IoT Environment

Hao Liu

Zhengzhou Health Vocational College, Xingyang, Henan 450100, China

Correspondence should be addressed to Hao Liu; liuhao@zzhvc.edu.cn

Received 29 January 2022; Accepted 1 March 2022; Published 31 March 2022

Academic Editor: Shan Zhong

As the computing capacity of existing mobile devices cannot fully meet the needs of users for communication quality, a computing resource allocation strategy for 5G communication in the Internet of Things (IoT) environment is proposed by applying UAV-assisted edge computing. First, a system model is constructed with the UAV deployed with mobile edge computing (MEC) servers to provide assisted computing services for multiple users on the ground. Based on the optimization of the UAV trajectory, communication scheduling, and the energy consumption model of the UAV, the problem of the total computational cost minimization is formulated. Then, the genetic algorithm is improved by introducing a penalty function to solve this problem, in which selection, crossover, and mutation operations are iterated to obtain the optimal allocation strategy for computational resources. Finally, a simulation platform is constructed to analyze the proposed method. The results show that the total cost and total time of the proposed strategy are better than other comparison strategies.

1. Introduction

With the increasing demands in applications and the rapid development of computing capacity, innovative applications are emerging quickly, making it difficult for the traditional 4G network architecture to serve exponential growth of data flows, diverse terminal devices, and various kinds of complex service scenarios [1]. The context of the Internet of Everything also places higher demands on wireless communications in terms of mobile performance, security, network delay, and communication reliability. At the same time, people have higher requirements for computational performance as artificial intelligence is developing quickly. However, the computing capacity of existing mobile devices cannot fully meet users' demands for communication quality [2, 3]. Therefore, mobile edge computing (MEC) technology is proposed to improve the system computing capacity so as to meet the users' demands for communication. In addition, 5G has become a key technology to achieve the best performance of MEC systems in the Internet of Things (IoT) environment, which can effectively solve the problems of delay, lack of storage, and poor computing capacity in wireless communication networks by combining reasonable resource optimization strategies for computing, storage, and communication [4]. It can also give full play to the characteristics of MEC networks, such as low response delay in computation task offloading and good network scalability [5].

By introducing UAVs into the MEC system, the flexible deployment of UAVs allows users to jointly collaborate on computing and communication, which improves the performance of the MEC system [6, 7]. However, the relationship between the trajectory control strategy and the capacity and energy consumption of traditional communication systems is unclear due to the fast mobility of UAVs. It is necessary to combine the classical air-to-ground line-of-sight channel model with a comprehensive multiobjective strategy optimization to improve the effectiveness of the algorithm based on the mobility of UAV during the communication [8].

There have been some research studies on traditional computation task offloading and resource allocation optimization at home and abroad. In Ref. [9], an exact linear search-based algorithm is proposed for finding the optimal

solution by fixing the status update rate. As a design alternative, a low complexity concave-convex procedure (CCP) is also formulated for finding a near-optimal solution relying on the original problem's transformation into a form represented by the difference of two convex problems. In Ref. [10], the optimization problem of computational resource allocation was transformed into the optimal matching for weighted bipartite graphs, which was solved by adding fake containers, thus obtaining the optimal strategy for resource matching of tasks in edge servers. But further improvement should be considered on the quality of service. Reference [11] proposed a dynamic resource allocation algorithm in the cloud edge environment and obtained the optimal scheme of resource scheduling based on tabu search algorithm, but the multiobjective optimization effect is not good. Reference [12] designed a multiobjective resource allocation method, in which Pareto archiving evolution strategy is used to optimize the time cost of IoT application, load balance, and energy consumption of MEC server, so as to find the optimal resource allocation scheme, but there is still room for further optimization for calculating delay. In Ref. [13], an enhanced resource allocation strategy in MEC using the reinforcement learning-based MOACO algorithm was proposed, in which MEC is used to predict the factors affecting delay and evaluate the performance of adjacent user equipment. A cognitive agent model was established to evaluate the validity of resource allocation. Zhang et al. proposed a MEC resource offloading method based on the UAV architecture and alternatively solved the problem of computational resource energy consumption and communication delay combined with an iterative algorithm, which improves the efficiency of optimal resource management [14]. However, the performance of overall optimization for multiple UAVs still needs to be improved. An iterative optimization algorithm with a double-loop structure for the optimal configuration of multi-UAV-assisted MEC systems was proposed by Zhang et al. [15]. Based on partial computing offload mode, jointly optimize user association and CPU cycle frequency for resource scheduling. In Ref. [16], a strategy for formulating and solving computational offloading problems using a model-free reinforcement learning framework was studied. And select an optimal overhead-aware computational offloading action based on its state, enabling the optimization of computational offloading policies in wireless cellular networks in MEC. However, the communication delay is still large due to the inadequacy of communication policies.

Based on the aforementioned analysis, an UAV-assisted edge computing resource allocation strategy for 5G communications in the IoT environment is proposed to address problems of energy consumption and delay in existing MEC networks. The main contributions are summarized as follows:

(1) In this article, a new mobile edge computing system is studied. In this system, the ground base station is used to serve the UAV for computing diversion, and some computing tasks are unloaded to the ground base station. Then, based on the energy consumption model of the rotor UAV, an optimization problem is formulated to minimize the total UAV energy consumption, At the same time, it meets the requirements of single target calculation and unloading of multiple ground base stations.

(2) In order to reduce the total system cost, the proposed method applies an improved genetic algorithm to solve the problem, in which the penalty function and combinatorial coding are used to improve the global search capability so as to obtain the optimal resource allocation scheme while improving the allocation efficiency.

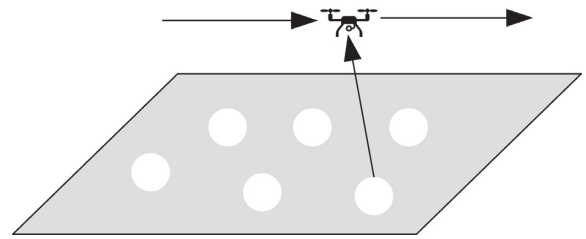


Figure 1: System model.

2. System Modeling

2.1. System Model. The scenario studied is the uplink communication in an UAV-assisted MEC system, where users on the ground perform computing tasks with the help of UAVs, as shown in Figure 1. In the system, there is a single-antenna rotary-wing UAV with a MEC server and N number of users on the ground, which can be denoted as $n \in \{1, 2, \cdots, N\}$. Considering the three-dimensional Cartesian coordinate system, the position of the nth ground user can be denoted as $l_n = (x_n, y_n)$. In addition, as the UAV flies in a defined region and provides assisted computing service to the ground users, the initial and final positions of the UAV can be denoted as $U_I = (x_I, y_I)$ and $U_F = (x_F, y_F)$, respectively.

Assume that each ground user offloads a portion of the computational task to the UAV and the remainder is executed locally due to the lack of computing capacity. It is also assumed that the total execution time T can be divided equally into K timeslots, where $T = K\delta$; here, δ refers to the length of the timeslot which is small enough to ensure that the position of the UAV does not change within each timeslot and $k = \{1, 2, \cdots, i \cdots, K\}$ denotes that there are K number of timeslots. In the ith timeslot, the position of the UAV can be represented as $(x[i], y[i])$ and the maximum flight speed of the UAV is $v_{\max}$, and then, the trajectory of the UAV can be constrained as

$$\begin{aligned} U[0] &= U_I, \\ U[K] &= U_F, \\ \frac{\|U[i] - U[i-1]\|}{\delta} &\leq v_{\max}, \end{aligned} \tag{1}$$

where the initial and final positions of the UAV during its flight are given and the flight speed of the UAV must not exceed the specified maximum flight speed. In addition, the energy consumption of the UAV during the flight in the ith timeslot can be calculated as

$$E^{\text{fly}}[i] = \alpha\|v[i]\|^2, \quad i \in k, \tag{2}$$

where $v[i] = [\|U[i+1] - U[i]\|/\delta]$ represents the flight speed of the UAV in the ith timeslot. $\alpha = 0.5J\delta$ and J is a constant related to the UAV payload.

2.2. Channel Model. Because the wireless channel between N ground users and UAV depends on the line-of-sight (LoS) channel. Therefore, the channel gain between the UAV and user n in the ith timeslot can be written as

$$h_n(i) = \frac{h_0}{d_n^2(i)} = \frac{h_0}{\sqrt{H^2 + \|U(i) - l_n\|}}, \tag{3}$$

where h_0 is the channel gain at a distance of 1 m, $d_n[i]$ is the distance between the user n and the UAV in the ith timeslot, and H denotes the fixed flight altitude of the UAV.

Additionally, assume that the transmitting power of the UAV is P_U and the offloading rate of the UAV data can be calculated as

$$\begin{aligned} R_n(i) &= B\log_2\left(1 + \frac{P_U h_n(i)}{\sigma^2}\right) \\ &= B\log_2\left(1 + \frac{\beta_0}{H^2 + \|U(i) - l_n\|}\right), \end{aligned} \tag{4}$$

where B represents the system bandwidth, σ^2 is the additive Gaussian white noise power, and β_0 represents the received signal to noise ratio (SNR) at a distance of 1 m.

2.3. Energy Consumption Model. From the above analysis, the energy consumption of the UAV consists of three main parts: the propulsion energy, computational energy, and communication energy. It is described in detail in the following section.

2.3.1. Propulsion Energy Consumption Model for UAV. The unit energy consumption (J/m) for the UAV during its flight at a speed of v can be calculated as

$$\begin{aligned} E_0(v) \triangleq \frac{P(v)}{v} &= P_0\left(\frac{1}{v} + \frac{3v}{V_{tip}^2}\right) \\ &+ P_u\left(\sqrt{v^{-4} + \frac{1}{4V_0^4}} - \frac{1}{2V_0^4}\right)^{1/2} \\ &+ \frac{1}{2}\tau_0\rho H A v^2, \end{aligned} \tag{5}$$

where V_{tip}^2 denotes the tip speed of the rotor blades and V_0 denotes the average rotor-induced hover speed. ρ is the air density. A denotes the area of rotor disc. τ_0 denotes the body drag ratio. P_0 and P_u denote the initial and process values of hovering power, respectively.

For any given UAV trajectory $\{U(i)\}$, the propulsion energy consumption can be expressed as

$$E_{fly}(T, \{U(i)\}) = \int_0^T P(\|v(i)\|)\mathrm{d}i, \tag{6}$$

where $v(i) \triangleq U(i)$ represents the speed of the UAV and $\|v(i)\|$ is the speed of the UAV at the moment i.

2.3.2. Communication Energy Consumption Model of UAV When Offloading Tasks. The time-division multiple access approach is considered here. Specifically, the UAV can establish contact with at most one ground user at any moment, and $\varphi_n(i) = 1$ indicates that the ground user n has established contact with the UAV. Otherwise, $\varphi_n(i) = 0$. $\{\varphi_n(i)\}$ is the user scheduling function, which should meet the constrains as

$$\sum_{n=1}^{N} \varphi_n(i) \le 1, \quad \forall i \in [0, T]. \tag{7}$$

The data volume $\overline{D}_n$ of the subtask offloaded from ground user n to the UAV is a function of the variable T, $U(i)$, and $\varphi_n(i)$, which can be expressed as

$$\begin{aligned} \overline{D}_n(T, \{U(i)\}, \{\varphi_n(i)\}) &= \int_0^T \varphi_n(i) R_n(i)\mathrm{d}i \\ &= B\int_0^T \varphi_n(i)\log_2 \\ &\left(1 + \frac{\beta_0}{H^2 + \|U(i) - l_n\|^2}\right)\mathrm{d}i. \end{aligned} \tag{8}$$

Hence, when a task packet with data size $(1 - \varpi_n)D_n$ is offloaded to the ground base station at the moment i, we can get

$$\overline{D}_n(T, \{U(i)\}, \{\varphi_n(i)\}) \ge (1 - \varpi_n)D_n, \quad \forall n \in N, \tag{9}$$

where D_n denotes users' total computation volume and ϖ_n denotes the bit allocation coefficient of the system.

Then, the communication energy consumption of the UAV when offloading tasks can then be obtained as

$$E_U = P_U\int_0^T\left(\sum_{n=1}^{N}\varphi_n(i)\right)\mathrm{d}i. \tag{10}$$

2.3.3. Computational Energy Consumption of UAV. The energy consumed in local execution is directly determined by the CPU workload F, which can be formulated as

$$F = D\zeta\ , \tag{11}$$

where D represents the total data volume of system tasks. ζ denotes the complexity of the application which obeys a Gamma distribution. In order to minimize CPU energy consumption, it is necessary to dynamically adjust the clock frequency in each CPU cycle to achieve the best working state. As a result, the computational energy consumption required to execute the computation task whose data size is $\varpi_n D_n$ within time T can be calculated as

$$E_{comp} = \frac{G(\varpi_n D_n)^3}{T^2}, \tag{12}$$

where the coefficient $G = 10^{-11}$ is a constant that takes the effective switching capacity and the probability of completion of the application execution into account.

To sum up, the total energy consumption of UAV is composed of propulsion energy consumption, calculated energy consumption, and communication energy consumption, which is calculated as follows:

$$\begin{aligned} E_{total} = {} & E_{fly}(T, \{U(i)\}) + P_U \int_0^T \left(\sum_{n=1}^{N} \varphi_n(i) \right) \\ & + \sum_{n=1}^{N} \frac{G(\varpi_n D_n)^3}{T^2}. \end{aligned} \tag{13}$$

2.4. Optimization Objectives. The optimization objective is to minimize the energy consumption of the UAV, which needs to take the user scheduling function $\varphi_n(i)$, the bit allocation coefficient of the system ϖ_n and the UAV trajectory $\{U(i)\}$ into consideration. Then, the problem can be formulated as

$$\begin{aligned} & \min_{T, \{U(i)\}, \{\varphi_n(i)\}, \varpi_n} E_{total} \\ \text{s.t. } & \overline{D}_n(T, \{U(i)\}, \{\varphi_n(i)\}) \geq (1 - \varpi_n) D_n, \quad \forall n \in N, \\ & \|\dot{U}(i)\| \leq v_{\max}, \quad \forall i \in [0, T], \\ & U(0) = U_I, \\ & U(T) = U_F, \\ & \varphi_n(i) \in \{0, 1\}, \quad \forall n \in N, i \in [0.T], \\ & 0 \leq \varpi_n \leq 1. \end{aligned} \tag{14}$$

It should be noted that the problem requires to jointly optimize the UAV trajectory $U(i)$, the communication scheduling $\varphi_n(i)$, and the bit allocation coefficient ϖ_n. And the trajectory $U(i)$ and the communication schedule $\varphi_n(i)$ are both continuous functions with respect to time i. Furthermore, the problem constraints are nonconvex and contain binary constraints, making it difficult to obtain optimal results.

3. Computational Allocation Strategy Based on an Improved Genetic Algorithm

Minimizing energy consumption is a nonlinear 0-1 programming problem, so the binary coded genetic algorithm is selected to solve the problem. Genetic algorithm is an efficient global search algorithm, which has certain advantages in solving complex problems such as nonlinear, multipeaked, large space, and globalization. It is famous for simulating the law of "superiority and inferiority" in the evolution of organisms in nature. The main idea is to encode the solution of the objective optimization problem into chromosomes and then simulate the genetic evolution of the chromosomes to gradually approach the optimal solution.

3.1. Encoding and Initial Populations. When the genetic algorithm is applied to globally find the optimal offloading strategy of a task, each offloading strategy needs to be encoded into a corresponding chromosome to facilitate later genetic operations such as crossover and mutation [17]. In fact, encoding is the process of converting the feasible solutions of the actual problem from its solution space to the search space of the genetic algorithm. And choosing the appropriate encoding method is the fundamental task in solving the genetic algorithm.

There are a variety of encoding methods. Since in the energy consumption minimization model for task offloading, the scheduling decision variable φ_n is a binary vector, binary encoding is most suitable and conforms to the encoding principle of minimum symbol set. The population size is set to be Q and task offloading decisions are randomly generate, which should be encoded based on the binary chromosome encoding method. Then, the initial chromosome population $S(0)$ can be obtained.

Since an offloading decision variable φ_n contains n number of decisions and it is known that those subtasks that can only be executed locally do not require a decision, they always have 0 decision, that is, $\varphi_n = 0$. Suppose that a mobile application contains n number of subtasks, where ϑ_0 number of tasks must be executed locally and ϑ_1 number of tasks can be offloaded, and $\vartheta_0 + \vartheta_1 = n$. Therefore, it is sufficient to make offloading decisions for the ϑ_1 number of subtasks. Namely, a chromosome can be constructed by using ϑ_1 number of binary bits (genes) in the genetic algorithm, and after arriving at the optimal solution, the completed offloading decision φ_n can be formed by combining them with another ϑ_0 number of binary bits whose values are 0. An example of the combined encoding solution is shown in Figure 2.

As shown in Figure 2, the mobile application is divided into 10 subtasks, with 3 locally executed tasks and 7 offloading tasks. Only the decision variables of 7 offloading subtasks need to be encoded and solved. When the optimal solution is achieved, it is combined with the decision variables of the other 3 local subtasks to arrive at the final offloading decision.

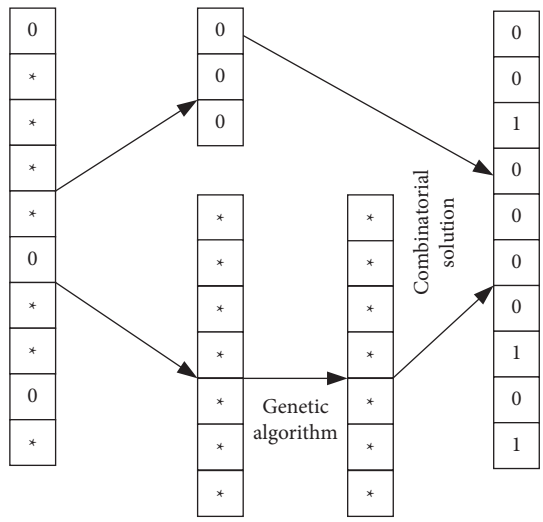


Figure 2: Combined encoding process.

3.2. Fitness Function. The optimization problem with constraints can be converted to an unconstrained problem by introducing a penalty function [18, 19]. The penalty function can combine the constraints with the objective function by introducing a multiplicative factor, and then, the optimization problem can be converted, while the optimal solution to the original optimization problem remains unchanged. The specific process can be written as

$$\phi = E_{fly}(T, \{U(i)\}) + P_U \int_0^T \left(\sum_{n=1}^{N} \varphi_n(i) \right) + \sum_{n=1}^{N} \frac{G(\varpi_n D_n)^3}{T^2} + \sum_{i=1}^{T} \gamma_1 \times \overline{D}_n(T, \{U(i)\}, \{\varphi_n(i)\}) + \gamma_2 \times \|\dot{U}(i)\| + \gamma_3 \times \varphi_n(i), \tag{15}$$

where γ_1, γ_2, and γ_3 are penalty factors, which are the weights of the inequality constraint functions in the optimization problem. The fitness function of the genetic algorithm can be combined with the original optimization problem based on the method of the penalty function. Since the proposed optimization problem is to minimize the total energy consumption of the task allocation under the constraints, the fitness function here will be designed to be inversely proportional to equation (15), which can be calculated as

$$\text{Fitness} = \frac{1}{\phi}. \tag{16}$$

3.3. Genetic Operations

3.3.1. Selection. The main function of selection is to simulate the operation of survival of the fittest in nature, select the chromosomes corresponding to some individuals with the best adaptation to produce the next generation, and eliminate the parents with poor adaptability [20, 21]. In this process, the number of individuals in parent and offspring populations should be the same.

The most common strategy used in selection is the "proportional selection" strategy. For example, the probability of selecting the bth element with priority a in a j th population can be expressed as

$$p_j^{a,b} = \frac{Fitness_j^{a,b}(x_b)}{\sum_{b=1}^{n} Fitnes_j^{a,b}(x_b)}, \quad x_b = \{0, 1\}. \tag{17}$$

Proportional selection is usually achieved using the Roulette Wheel Selection (RWS) algorithm, which yields probabilities $p_j^{a,b}$ that are always distributed in the interval [0,1] and satisfy

$$\sum_{b=1}^{n} p_j^{a,b} = 1. \tag{18}$$

Thus, for a population with a number of priority tasks in each generation, assuming that the total number of ath priority tasks is W and the RWS algorithm first generates W random numbers in the interval [0,1], [0, 1] and divides the wheel into W parts according to the individual probabilities $p_j^{a,b}$. For each random number, if it is within $[\sum_{r=1}^{b-1} p_j^{a,r}, \sum_{r=1}^{b} p_j^{a,r}]$, the individual corresponded to the task will be selected to generate the next population with W number of individuals. Thus, the RWS algorithm can be regarded as a scheme that decides whether to select an individual based on the probability corresponding to its fitness value. Therefore, this selection method has no safeguard mechanism and may allow some well-adapted individuals to be lost in genetics. According to the proposed model, elitist strategy selection is used to ensure that the better-adapted individuals will not be lost in the next generation. The elitist strategy selection is an improvement of the basic genetic algorithm. In order to prevent the optimal solutions generated during evolution from being destroyed by crossover and mutation operations, the optimal one in each generation will be copied directly into the next generation [22, 23]. The genetic algorithm with the

introduction of elitist strategy tends to have faster convergence and better stability [24]. Since the execution time of tasks with higher priority affects the waiting time for the subsequent tasks with lower priority, the fitness value and the maximum time for the execution of tasks with the same priority should be considered simultaneously when performing elitist strategy selection.

3.3.2. Crossover and Mutation. Individuals are selected for the crossover and mutation operations based on the probabilities of crossover and mutation. The crossover operation for two 6×4 matrices is shown below, with the positions of the lines chosen at random:

$$\text{Before crossover: } \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}$$

$$\text{After crossover: } \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}$$

3.4. Genetic Update. The new population is formed by combining the population obtained through the above operations with the initial population, and the energy consumption and the corresponding fitness value of each individual in the new population need to be calculated [25, 26]. The minimum value ϕ is taken as the result of this iteration, and the individual with the greatest fitness value is selected to update the population, that is, $i = i + 1$.

4. Experiments and Analysis

In the simulation experiments, numerical results will be provided to validate the proposed algorithm. The altitude of the UAV is set to $H = 100$ m, and the total communication bandwidth is $B = 3$ MHz. The SNR is $h_0 = 60$ dB, and the communication power consumption of the UAV is fixed at $P_U = 75$ W. The maximum flight speed is $v_{\max} = 60$ m/s, and the initial and final positions of the UAV are set to $U_I = [0, 0]^T$ and $U_F = [1200, 1200]^T$. The specific simulation parameter settings for the scenario when the number of users is $N = 50$ are shown in Table 1.

The parameters of the genetic algorithm are set as follows. The number of populations is 50. The maximum number of iterations is 500. The crossover probability is 0.7, and the initial mutation probability is 0.01. The penalty factors γ_1, γ_2, and γ_3 are set to be 107.

4.1. UAV Trajectory. The UAV trajectory based on the proposed method under different task execution time T is shown in Figure 3, where the red circle indicates the positions of the ground users.

Table 1: Simulation parameter setting.

Simulation parameters	Value
Flight altitude H	100 m
Channel bandwidth B	3 MHz
SNR received at reference distance 1 m h_0	60 dB
Communication power of UAV P_U	75 W
Maximum flight speed $v_{\max}$	60 m/s
Initial position U_I	[0, 0]
Final position U_F	[1200, 1200]
Number of users N	50

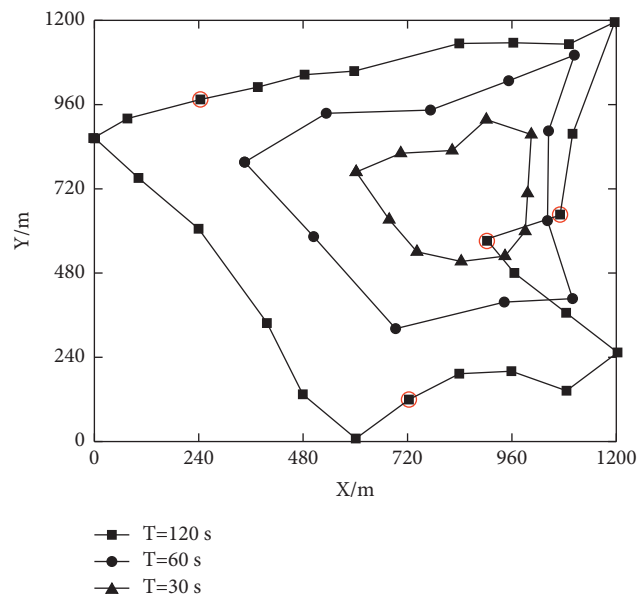


Figure 3: The map of UAV trajectory.

As can be seen in Figure 3, when $T = 30$ s, the UAV flies along a closed-loop trajectory and has a small flight range. The UAV is farther away from the user when offloading data and the communication quality is poor. As T increases, that is, $T = 60$ s and $T = 120$ s, the distance between the UAV and users decreases and the flight range expands, improving the communication quality. When T is large enough, the UAV continues to fly along a closed-loop trajectory and serves one user for a certain number of timeslots. Thus, it can be concluded that the larger T leads to the greater flight range of the UAV, making it is closer to the users, and then brings better communication quality.

4.2. Comparison with Other Algorithms. To demonstrate the performance of the proposed method, it is analyzed in comparison with algorithms proposed in Ref. [10], Reference [12], and Ref. [16].

4.2.1. Effect of the Number of Tasks on Total Cost. When the weight of time in the total cost is 0.5, there are 10 levels of task priorities, and the effect of changes in the maximum number of tasks at each level on the total cost is shown in Figure 4. The horizontal coordinate indicates the maximum

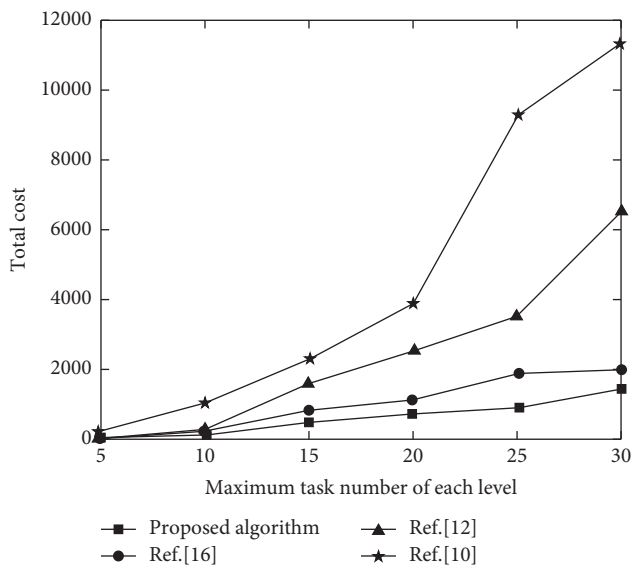


Figure 4: Effect of the number of tasks on total cost.

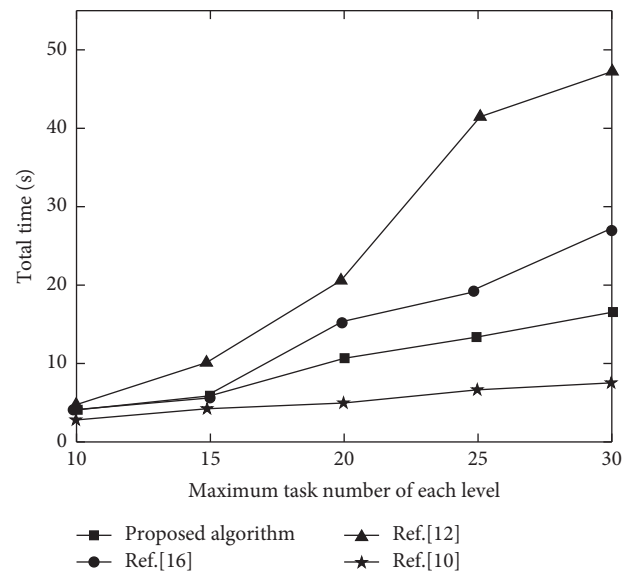


Figure 5: Effect of the number of tasks on total time.

number of tasks at each level. In the experiment, for each level, a random number from 1 to the maximum is randomly generated and the tasks are numbered, and then, the dependencies on the tasks at the previous level are randomly generated according to the mutation of different tasks. The vertical coordinate shows the total cost corresponding to the allocation results obtained in different algorithms.

It can be seen in Figure 4 that the total cost grows gradually as the number of tasks at each level increases. Meanwhile, the lowest total cost is obtained in the proposed method as the UAV is used for computational task offloading and the improved genetic algorithm is introduced for problem-solving. When the task size is 30, the total cost reaches 1800. Both Ref. [10] and Ref. [12] use a traditional optimization algorithm, and the total cost grows most rapidly with the increase of tasks in each level. Reference [16] realizes the resource allocation based on a reinforcement learning network. However, it lacks a reasonable offloading target, so the total computational cost is higher than the proposed method. In addition, when there are more tasks per level, greater reduction in the total cost of executing all tasks can be achieved using the proposed method and the algorithm proposed in Ref. [16], while the proposed method has a more significant advantage.

4.2.2. Effect of the Number of Tasks on Total Time. When there are 10 levels in task priority, the effect of the change in the number of tasks on the total time to execute the same task is illustrated in Figure 5. The horizontal coordinate indicates the maximum number of tasks at each level, and the vertical coordinate indicates the total time to execute all tasks.

As depicted in Figure 5, the total computation time is the shortest due to the simplicity of the algorithm proposed in Ref. [10], which creates a weighted complete bipartite graph by adding false containers to achieve the optimal matching of task resources in the edge server. Reference [12] uses a Pareto archiving evolutionary strategy to optimize time cost, load balancing of MEC servers, and energy consumption. The overall time consumed rises rapidly with the increase of tasks due to too many optimization targets and the slow search speed of the optimization algorithm. The reinforcement learning network in Ref. [16] is more complex, and its search time is longer compared to the proposed method, which uses an improved genetic algorithm. And the proposed method introduces the UAV to dynamically adjust the offloading strategies and reduce the data transmission time; thus, its total time cost is 17 s when the task size is 30, which is about 39% less than Ref. [16].

4.2.3. Effect of the Number of Mobile Terminals on Total Cost. The effect of the change in the number of mobile terminals on the total system cost is shown in Figure 6.

As illustrated in Figure 6, when the number of mobile terminals is less than 30, the computing resources are less and the MEC server resources are relatively sufficient. Therefore, most tasks are unloaded to the MEC server for execution, and the overhead obtained by collaborative scheduling of different algorithms is not different. With the increase of mobile terminals, resources in the MEC server gradually decrease as the collaborative scheduling uses the computational resources of the mobile terminals. The total cost of Ref. [10] and Ref. [12] grows rapidly to over 4×104 due to their poor performance in finding the optimal solution. The proposed method allocates resources based on an improved genetic algorithm and the UAV platform, and its system cost is lower than that of Ref. [16] which applies the reinforcement learning network. Specifically, the cost of the proposed method is only half of that of Ref. [16] when there are 100 mobile terminals in the system, demonstrating that the resource allocation strategy in the proposed method is ideal.

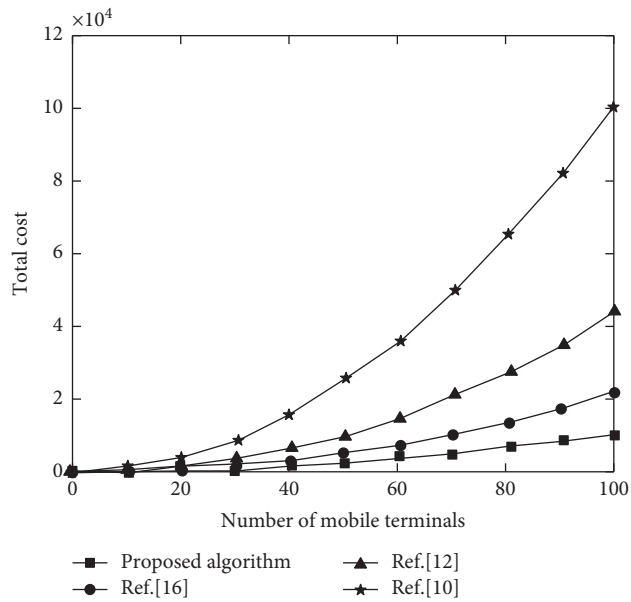


Figure 6: Effect of the number of mobile terminals on the total cost.

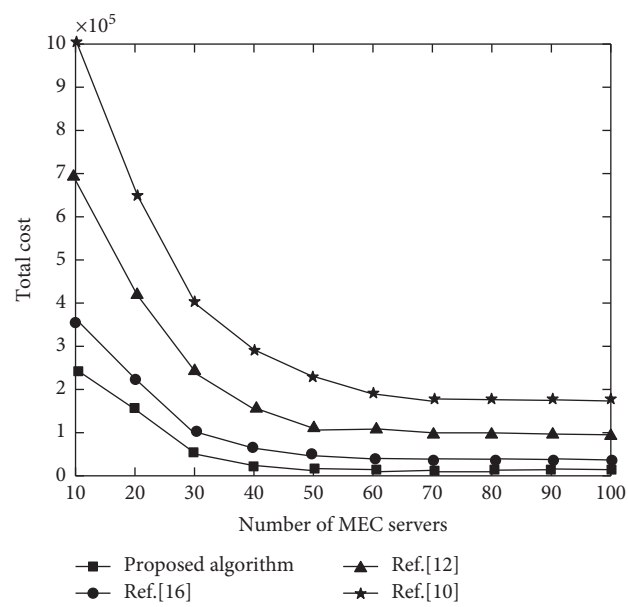


Figure 7: The effect of the number of MEC servers on the total cost.

4.2.4. Effect of the Number of MEC Servers on Total Cost. The effect of the change in the number of MEC servers on the total system cost is shown in Figure 7.

It is depicted in Figure 7 that the queuing time at the MEC server side decreases as the total computational cost gradually decreases. The proposed method has the smallest total computational cost, which is less than 0.2×105 when the number of MEC servers is 100. This is because the use of UAV allows dynamic adjust on the positions of MEC servers, causing less transmission loss. Furthermore, as the number of MEC servers continues to increase, all tasks are offloaded to the MEC server for execution when resources in the MEC server are sufficient. When the number of servers exceeds 70, the overhead of the whole system is almost unchanged, which shows that only increasing the number of MEC servers cannot increase the total overhead of the system. With the increase of MEC servers, the queuing latency decreases and the cost when all tasks are offloaded to MEC servers decreases at first, but it remains constant when there are enough MEC servers.

5. Conclusion

As a key and fundamental technology in 5G, the MEC system needs to meet the requirements of low latency and low energy consumption. Applying the high flexibility of UAVs, an UAV-assisted edge computing resource allocation strategy for 5G communication in the IoT environment is proposed. A system model of an UAV deployed with a MEC server to provide assisted computing services for multiple users on the ground is constructed. And the minimization of total system cost, which is composed of channel model and energy consumption model, is solved by the improved genetic algorithm, obtaining the best resource allocation strategy. The simulation results show that

(1) The system model introduces the UAV-assisted edge computing with high flexibility, which can effectively reduce the total system time to be less than 17 s when the number of tasks increases.

(2) Improving the genetic algorithm by using a penalty function and combination coding can enhance its ability of global optimization so as to obtain a resource allocation scheme that minimizes the total system cost. The total cost of the proposed method is 1800 when the number of tasks per level is 30 and it is less than 0.2×105 when the number of MEC servers is 100.

The main scenario studied in this article is a single UAV providing assisted computing services to users, while multi-UAV scenarios are more important in practical applications. Therefore, in future research studies, the focus will be attached on the scenarios of multiple UAVs serving ground users.

Data Availability

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

Conflicts of Interest

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

References

[1] Z. Fang, J. Wang, J. Du, X. Hou, Y. Ren, and Z. Han, "Stochastic optimization-aided energy-efficient information collection in Internet of underwater things networks," *IEEE Internet of Things Journal*, vol. 9, no. 3, pp. 1775–1789, 2022.

[2] Q. Wang, S. Guo, J. Liu, and Y. Yang, "Energy-efficient computation offloading and resource allocation for delay-

sensitive mobile edge computing," *Sustainable Computing: Informatics and Systems*, vol. 21, no. 3, pp. 154–164, 2019.

[3] D. Zhu, C. Xiang, and S. Bing, "Biologically inspired self-organizing map applied to task assignment and path planning of an AUV system," *IEEE Transactions on Cognitive and Developmental Systems*, vol. 10, no. 99, pp. 304–313, 2018.

[4] Y. Chen, Z. Li, B. Yang, K. Nai, and K. Li, "A Stackelberg game approach to multiple resources allocation and pricing in mobile edge computing," *Future Generation Computer Systems*, vol. 108, no. 3, pp. 273–287, 2020.

[5] M. Mukherjee, S. Kumar, C. X. Mavromoustakis et al., "Latency-driven parallel task data offloading in fog computing networks for industrial applications," *IEEE Transactions on Industrial Informatics*, vol. 16, no. 9, pp. 6050–6058, 2020.

[6] H. Wang, "Collaborative task offloading strategy of UAV cluster using improved genetic algorithm in mobile edge computing," *Journal of Robotics*, vol. 2021, Article ID 3965689, 9 pages, 2021.

[7] Y. Zhang, M. Zhang, and C. Fan, "Computing resource allocation scheme of IOV using deep reinforcement learning in edge computing environment," *EURASIP Journal on Applied Signal Processing*, vol. 2, no. 1, pp. 1–19, 2021.

[8] D. Wang, H. Qin, B. Song, X. Du, and M. Guizani, "Resource allocation in information-centric wireless networking with d2d-enabled MEC: a deep reinforcement learning approach," *IEEE Access*, vol. 7, no. 10, pp. 114935–114944, 2019.

[9] Z. Fang, J. Wang, Y. Ren, Z. Han, H. V. Poor, and L. Hanzo, "Age of information in energy harvesting aided massive multiple access networks," *IEEE Journal on Selected Areas in Communications*, p. 1, 2022.

[10] Wei, Feng, and Zou, "Distance control strategy for deploying content replica servers in an edge cloud," *The Journal of China Universities of Posts and Telecommunications*, vol. 26, no. 2, pp. 95–102, 2019.

[11] H. Tang, C. Li, J. Bai, J. Tang, and Y. Luo, "Dynamic resource allocation strategy for latency-critical and computation-intensive applications in cloud-edge environment," *Computer Communications*, vol. 134, no. 6, pp. 70–82, 2019.

[12] Q. Liu, R. Mo, and X. Xu, "Multi-objective resource allocation in mobile edge computing using PAES for Internet of Things," *Wireless Networks*, vol. 4, no. 3, pp. 1–13, 2020.

[13] V. Shanmuganthan, M. Khari, and N. Dey, "Enhanced resource allocation in mobile edge computing using reinforcement learning based MOACO algorithm for IIOT," *Computer Communications*, vol. 151, no. 4, pp. 355–364, 2020.

[14] X. Zhang, J. Zhang, J. Xiong, L. Zhou, and J. Wei, "Energy-efficient multi-UAV-enabled multiaccess edge computing incorporating NOMA," *IEEE Internet of Things Journal*, vol. 7, no. 6, pp. 5613–5627, 2020.

[15] J. Zhang, L. Zhou, F. Zhou et al., "Computation-efficient offloading and trajectory scheduling for multi-UAV assisted mobile edge computing," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 2, pp. 2114–2125, 2020.

[16] Y. Wei, Z. Wang, D. Guo, and F. Richard Yu, "Deep Q-learning based computation offloading strategy for mobile edge computing," *Computers, Materials & Continua*, vol. 59, no. 1, pp. 89–104, 2019.

[17] X. Li, Y. Dang, M. Aazam, X. Peng, T. Chen, and C. Chen, "Energy-efficient computation offloading in vehicular edge cloud computing," *IEEE Access*, vol. 8, no. 5, pp. 37632–37644, 2020.

[18] M. Hua, Y. Wang, C. Li, Y. Huang, and L. Yang, "UAV-aided mobile edge computing systems with one by one access scheme," *IEEE Transactions on Green Communications and Networking*, vol. 3, no. 3, pp. 664–678, 2019.

[19] M. Sheng, Y. Wang, X. Wang, and J. Li, "Energy-efficient multiuser partial computation offloading with collaboration of terminals, radio access network, and edge server," *IEEE Transactions on Communications*, vol. 68, no. 3, pp. 1524–1537, 2019.

[20] P. A. Apostolopoulos, E. E. Tsiropoulou, and S. Papavassiliou, "Risk-aware data offloading in multi-server multi-access edge computing environment," *IEEE/ACM Transactions on Networking*, vol. 28, no. 3, pp. 1405–1418, 2020.

[21] C. Gza, Z. Hao, and B. Yla, "5G network-oriented hierarchical distributed cloud computing system resource optimization scheduling and allocation," *Computer Communications*, vol. 164, no. 5, pp. 88–99, 2020.

[22] S. F. Abedin, M. G. R. Alam, S. M. A. Kazmi, N. H. Tran, D. Niyato, and C. S. Hong, "Resource allocation for ultra-reliable and enhanced mobile broadband IoT applications in fog network," *IEEE Transactions on Communications*, vol. 67, no. 1, pp. 489–502, 2019.

[23] R. Siddavaatam, I. Woungang, and A. Anpalagan, "Joint optimisation of radio and infrastructure resources for energy-efficient massive data storage in the mobile cloud over 5G HetNet," *IET Wireless Sensor Systems*, vol. 9, no. 5, pp. 323–332, 2019.

[24] Y. Wei, F. R. Yu, M. Song, and Z. Han, "Joint optimization of caching, computing, and radio resources for fog-enabled IoT using natural actor-critic deep reinforcement learning," *IEEE Internet of Things Journal*, vol. 6, no. 2, pp. 2061–2073, 2019.

[25] S. Zhang, Y. Liang, J. Ge, M. Xiao, and J. Wu, "Provably efficient resource allocation for edge service entities using hermes," *IEEE/ACM Transactions on Networking*, vol. 28, no. 4, pp. 1684–1697, 2020.

[26] C. Tang, C. Zhu, X. Wei, H. Wu, Q. Li, and J. J. P. C. Rodrigues, "Intelligent resource allocation for utility optimization in RSU-empowered vehicular network," *IEEE Access*, vol. 8, no. 1, pp. 94453–94462, 2020.

Hindawi
Journal of Robotics
Volume 2022, Article ID 8954060, 9 pages
https://doi.org/10.1155/2022/8954060

Research Article

Intelligent Obstacle Avoidance Algorithm for Mobile Robots in Uncertain Environment

Liwei Guan,[1,2,3] **Yu Lu**,[1,2,3] **Zhijie He**,[1,2,3] **and Xi Chen**[1,2,3]

[1]*College of Physics and Energy, Fujian Normal University, Fujian Provincial Key Laboratory of Quantum Manipulation and New Energy Materials, Fuzhou, Fujian 350117, China*
[2]*Fujian Provincial Collaborative Innovation Center for Optoelectronic Semiconductors and Efficient Devices, Xiamen, Fujian 361005, China*
[3]*Fujian Provincial Engineering Technology Research Center of Solar Energy Conversion and Energy Storage, Fuzhou, Fujian 350117, China*

Correspondence should be addressed to Liwei Guan; guanlw@fjnu.edu.cn

Received 30 January 2022; Revised 7 March 2022; Accepted 9 March 2022; Published 30 March 2022

Academic Editor: Shan Zhong

The application of mobile robots and artificial intelligence technology has shown great application prospects in many fields. The ability of intelligent obstacle avoidance is the basis for the deep application of mobile robots. However, there are often more or less uncertain factors in the actual operating environment of the robot, such as people or objects that are not updated in time or temporarily appear. Therefore, it is an important step to complete the automatic learning of obstacle avoidance for mobile robots. In a nondeterministic environment, a mobile robot intelligent obstacle avoidance algorithm based on an improved fuzzy neural network with self-learning is firstly proposed. The mobile robot intelligent obstacle avoidance system is constructed through the reaction layer, the deliberation layer, and the supervision layer. Through the analysis of sensor performance, model accuracy, path obstacle avoidance optimization, and obstacle avoidance simulation, the following conclusions are drawn. First, through network training, the accuracy rate of the test set is stable at 98%, and the loss of the function value has also been reduced from the original 0.79 to 0.08, which is 10 times smaller. Second, the traditional single sensor cannot meet the obstacle avoidance requirements of robots, and mobile robots must combine multipurpose technology. Third, the algorithm in this paper encounters the following. When there are obstacles, the path is dominated by straight lines, obstacle avoidance planning is optimal, and the distance is shorter. Fourth, the larger $N:M$, the larger the solution space, indicating that this algorithm gradually improves the search efficiency to the greatest extent and can handle any form of medium and large scale task allocation problem.

1. Introduction

Robot technology is an important manifestation and symbol of the degree of industrial automation and the national high-tech level. Robot technology is a high-tech formed on the basis of interdisciplinary. It is one of the hot spots of current scientific research [1]. Mobile robots are widely used in various fields, especially in harsh and dangerous environments such as military reconnaissance, antinuclear pollution, mine clearance, and material handling in civil use. Therefore, mobile robots greatly facilitate people's production and life. The research on robots has received widespread attention.

On the basis of the continuous improvement of artificial intelligence, the development of robot technology is becoming more and more intelligent, and the ability of mobile robots to intelligently avoid obstacles is an important indicator of their intelligence [2]. They not only reflect the efficiency, feasibility and energy consumption of mobile robot motion but also reflect the way the robot detects obstacles, processes obstacle information, and avoids obstacles. Today's mobile robots have moved from structured work spaces to uncertain environments [3]. In an unknown and uncertain environment, it is an important step for mobile robots to learn to avoid obstacles autonomously. It enables mobile robots to have human-like behaviour

strategies and avoid obstacles, thereby realizing intelligent navigation of mobile robots [4]. How to intelligently avoid these obstacles unknown to the environment map in an uncertain environment is a key link of mobile robots and one of the research hot spots of robotics.

At present, great progress has been made in the research of autonomous obstacle avoidance algorithms for mobile robots in the world, and many algorithms are applied in the autonomous obstacle avoidance systems of mobile robots. These intelligent obstacle avoidance algorithms mainly include fuzzy control, grid map, artificial potential field, and neural network [5]. The fuzzy control algorithm has the advantages of simple algorithm, easy to understand, and strong robustness, but there are also problems such as the need for experience in design, low control accuracy, and no learning ability; neural network algorithms can use a large amount of data to train models, which can automatically learn parameters to obtain the final end-to-end network model, but there are problems of high network complexity and practicability. This paper fully considers the practicability and operability of intelligent obstacle avoidance of mobile robots, combines fuzzy control algorithm with neural network algorithm, and takes the lead in proposing an intelligent obstacle avoidance algorithm for mobile robots based on improved fuzzy neural network with self-learning.

2. Related Work

A mobile robot is a robot system that completes certain work functions. It can sense the external environment and its own situation through sensors and realize its autonomous movement in an environment with obstacles [6]. Mobile robots have become a research hot spot in robotics because they have shown more and more extensive application prospects in various aspects of agriculture, industry, aerospace, medicine, and human life [7].

The earliest research on mobile robots was from 1966 to 1972, when Nils Nilssen and Charles Rosen of Stanford Research Institute developed an autonomous mobile robot [8]. In the 1980s, a number of universities such as Stanford and MIT established a special scientific research team for ALV research [9]. In the 1990s, Japan developed working robots that can operate in extreme environments and successfully developed humanoid robots [10]. In 2003, Tsinghua University in China independently developed the robot automatic navigation system [11]. All over the world, Japan has always been at the forefront of the world in the field of humanoid robots and wheeled robots. The ASIMO biped walking robot developed by Honda represents the highest level in the world. ASIMO uses a variety of sensors in the realization of its functions, including cameras on the head, ultrasonic sensors around, and ground pressure sensors on the bottom of the feet. In terms of functional realization, ASIMO not only realizes the functions of walking, positioning, and navigation but also adds functions such as face recognition and communication by voice or gesture [12].

In recent years, breakthroughs have been made in the research on the theory and algorithm of navigation control of mobile robots in unknown environments. Yan Z. pointed out that based on global path planning, navigation and obstacle avoidance are performed according to the planned path method [13]. Jin J. combined fuzzy logic and grid graph in the target navigation of mobile robots, using the minimum risk criterion as the evaluation function to improve the effect of path planning [14]. Zafar M. proposed a learning algorithm based on neural network error backpropagation to adjust the membership function parameters of the fuzzy logic system to improve the trajectory smoothness of the mobile robot [15]. Cui Min applied the improved neural network algorithm to the path planning of mobile robots, which improved the operation efficiency of the algorithm [16]. Chen D. proposed a reinforcement learning mechanism based on fuzzy neural network, which utilizes the residual in the learning algorithm and obtains a better algorithm convergence speed in the navigation process of the mobile robot [17].

In the previously mentioned research, different algorithms are used in mobile robot path planning and navigation and obstacle avoidance, and the performance of the system has been improved to a certain extent. However, in the nondeterministic environment, the previously mentioned obstacle avoidance and planning algorithms have some deficiencies in flexibility, real-time, humanization, and intelligent performance. This paper focuses on the problems existing in the previously mentioned algorithms. In a nondeterministic environment, an intelligent obstacle avoidance algorithm for mobile robots based on an improved fuzzy neural network with self-learning is firstly proposed, which improves the flexibility, autonomy, and stability of mobile robots.

3. Implementation of Intelligent Obstacle Avoidance Algorithm

This paper fully considers the importance of obstacle avoidance for mobile robots in nondeterministic environments and proposes an intelligent obstacle avoidance algorithm for mobile robots based on improved fuzzy neural network autonomous learning. First, in view of the high complexity and poor real-time performance of the fuzzy algorithm, the necessary simplification of the fuzzy neural network is carried out to perform image preprocessing and rough positioning of obstacles. Second, the parameters of the fuzzy membership function are automatically and dynamically adjusted by the neural network with self-learning ability, so that the fuzzy control rules have stronger object adaptability. Furthermore, for dynamic obstacles in autonomous obstacle avoidance, edges are introduced. The detection operator implements the dynamic obstacle avoidance strategy of the mobile robot.

3.1. Improved Fuzzy Neural Network Algorithm. This paper applies the improved fuzzy neural network algorithm to the field of image processing, analyzes the image to detect obstacles by counting the neurons in the visible range, and uses the knowledge of distance images to convert the detection results to finally determine the location of the obstacles.

3.1.1. Image Preprocessing and Coarse Positioning of Obstacles. Image preprocessing includes image greyscale, image denouncing, and effective image region selection. Image greyscale is to complete the conversion from colour to greyscale. In the system, the video image format conversion is completed through the acquisition of the camera. The value of the three colour components of *R*, G, and B in the RGB space of the image is stored in the image, and greyscale can be achieved by using

$$\text{gray} = 0.5 \times \text{Red} + 0.23 \times \text{Green} + 0.27 \times \text{Blue}. \quad (1)$$

The size of the redundancy is related to the probability of occurrence of each basic element in the information. Its expression is as follows:

$$M(X) = E[\log q(ai) \times 10] = \sum_{i=0}^{n} qi(ai) \log qi(ai). \quad (2)$$

Among them, ai is the event in the sample population X. The previously mentioned formula indicates that the more average the number of occurrences of each sample event, the greater the number of different events in the sample population *X*, and the greater the amount of information. If there are N groups of samples, $X_1, X_2 \ldots X_n$, the corresponding sample numbers are $k_1, k_2 \ldots k_n$, and the information entropy is $M_1, M_2 \ldots M_n$. The unit information entropy is the average value of the information contained in the unit sample in the sample population, which is embodied as the ratio of the population neuron to the number of spatial samples, namely,

$$D(X) = M(X) * n = \sum_{i=0}^{n} qi(ai) \log qi(ai) \times \frac{1}{n}. \quad (3)$$

Among them, D(X) represents the unit neuron of the sample space X. Various colour components of an image can be used as the basic measure of the amount of information, and grayscale representation is also a method of expressing information. Each grayscale level is the basic element of the composition. The probability of occurrence is used to characterize the amount of information in the area.

Use the knowledge of neurons to analyze the image and locate obstacles. The unit information entropy expression formula of the image is applied to a certain frame of greyscale image and can be written as

$$M = + \sum_{i=0}^{255} kqi + \log qi. \quad (4)$$

Among them, qi represents the probability of the occurrence of a pixel whose gray level is *i*; k is the number of gray levels whose qi is not zero in all gray levels. In the actual experiment, the horizontal image analysis process is as follows: select a row of pixels as the data source for statistics qi; that is, count the gray level distribution of the pixels in this row, and obtain the corresponding pixel points in each gray level *i*. The number of pixels, qi, is the ratio of the number of pixels contained in the gray level to the total number of pixels in the row; then, the unit information entropy is calculated for all gray levels whose qi is not 0.

3.1.2. Accurate Positioning of Obstacles in the Image. Since the experimental environment is relatively simple, the relatively simple Roberts edge detection operator can be used for detection to obtain edge point information, which can effectively improve the real-time performance of the system. The Roberts operator in an image pixel array looks like this

$$k(a,b) = \sqrt{[t(a,b) + t(a+1,b+1)]^2 - [t(a,b+1) + t(a+1,b)]^2}. \quad (5)$$

The Roberts operator is a diagonal derivative computed in a 2×2 neighbourhood. In practical applications, the previously mentioned formula can be replaced by a simplified calculation form, which is expressed as follows; that is, the absolute value of Roberts is replaced.

$$k(a,b) = |t(a,b) + t(a+1,b+1)| + |t(a,b+1) + t(a+1,b)|. \quad (6)$$

3.1.3. Distance Image Obstacle Location. The range image stores the depth information of the ray associated with each pixel and the first focus of the scene observed by the camera. According to the idea of interpolation, using the geometric principle of camera imaging, using Manlius's theorem, establish a three-dimensional distance function, calculate the distance from any position in the monitoring area to the camera, and establish a distance image about the distance between any point in the monitoring area and the camera:

$$\alpha = \arctan h \times y1, \quad (7)$$

$$\beta = \arctan h \times (y1 - y2), \quad (8)$$

$$\lambda = \arctan(x1 \times y1), \quad (9)$$

where *h*, y1, y2, and x1 are data that can be measured. After obtaining α, β, and γ, *x* and *y* can be obtained from the trigonometric function relationship, and the derivation formula is as follows:

$$\begin{aligned} y &= h \times \tan\left[(45+\alpha) - \left(s_y - u\right)s_y \times (\alpha+\beta)\right], \\ x &= y \times \tan\left[s_x + v\right] \times s_x + \gamma, \\ B &= \sqrt{x^2 - y^2} \times \sqrt{x^2 + y^2}. \end{aligned} \quad (10)$$

Among them, B is the distance between the vertical projection points of the camera coordinates; S_x and S_y are the total number of pixels in the *x* and y directions in the image

plane; u and v represent the horizontal and vertical imaging planes, respectively.

The position of the target point in ABCD can be obtained by formulas (7)–(9), that is, the coordinates of the target point in the world coordinate system. By analyzing the image through image neurons and edge information in the image, the characteristic edge points of obstacles can be obtained. Through the knowledge in this section, these correspondences with the world coordinate system are established to obtain the position of the obstacle in the actual situation.

3.2. Intelligent Obstacle Avoidance System. A very important task area for mobile robots is navigation, which is to find a route so that it can move safely without collisions around all obstacles. This planning technology that can autonomously avoid obstacles and complete tasks is the frontier technology for intelligent mobile robots to achieve autonomous behaviour control. Therefore, it is the most basic problem to apply the improved fuzzy neural network algorithm to the intelligent obstacle avoidance system of mobile robots. The mobile robot intelligent obstacle avoidance system designed in this paper integrates different functions in the system, as shown in Figure 1.

Much of the structure of the intelligent obstacle avoidance system involves planning, using a reactive planner called PRS-L. PRS-L can accept human natural language instructions, and then, start to run navigation tasks and perceptual recognition routines [18]. Both planning and execution rely on a locally aware spatial-centrist model of the environment. The reactive components of an intelligent architecture consist of behavior. These behavior extract virtual sensor input and output fuzzy rules from the local perception space of the central environment model and then synthesize control instructions through fuzzy logic.

From the intelligent obstacle avoidance system, it can be seen that a hybrid architecture should have the following modules and objects.

Sequencer agent is used to generate the set of tasks required to complete the subtasks and to determine all timing and activation conditions. Timing is usually represented as a correlation network or finite state machine, and sequencers can generate these structures and modify them dynamically [19].

Resource manager is used to allocate resources for behavior including selection of schema libraries.

Cartographer is used to generate, store, and maintain map or spatial information and provide a means of accessing data.

Mission planner interacts with the user, passes instructions to the robot, and generates mission plans.

Performance supervision and problem solving agents are used to let the robot notice if it is making progress.

3.3. Intelligent System Hierarchy. Because the mobile robot obstacle avoidance system adopts the modular design method based on mufti-agent, the functions are independent of each other. Therefore, it is very easy to realize the function expansion and can be used as the design and experiment platform of the ideal hybrid architecture. In this chapter, we are based on the mobile robot intelligent obstacle avoidance architecture, follow the general design principles of intelligent system structure, and add supervision management and learning functions on the basis of the original hybrid structure. Through the improved fuzzy neural network algorithm, the robot can learn behavior autonomously in the dynamic unknown environment so as to adapt to the changes of the environment, as shown in Figure 2.

Reactive behaviour control layer contains reflective behaviour and reactive behaviour, as well as a structure for coordinating reactive behaviour. Reflective behaviour is used to react to emergencies in the environment, or it can be a single control rule. For the navigation tasks of mobile robots, reactive behaviors are designed for obstacle avoidance, goal orientation, and roaming. These behaviors can be pare-designed by expert experience or acquired through learning and evolution. All behavior can directly respond to the information felt from the outside world and can also receive control signals from the deliberation layer and the supervision layer to perform actions.

Regarding the deliberate behaviour control layer, the task for the navigation class includes five basic modules, and other functions can be added on this basis. The modules are environment model knowledge base, task planning module, positioning module, navigation module, and sequencer. The environmental model knowledge base can directly receive the environmental model transmitted from the outside world and store it in the knowledge base through human-computer interaction and can also generate a map based on the collected sensor information, maintain the map information in time, and provide orientation data to other modules. The task planning module receives the instructions input by the user and transmits it to the robot for task planning. The positioning module uses the shaft angle encoder combined with the feature quantity extracted from the external sensor data to determine the position of the robot in the environment at each moment. Navigation accepts the task from the task planning module and calculates the path and decomposes it into subtasks. The sequencer is used to generate the task set required to complete the subtasks and determine all the timing and activation conditions. When the purposeful planning is required, the reaction layer obtains the ordered actions through the sequencer.

Supervise and manage the behaviour control layer: used to monitor the execution of the deliberation layer and the behaviour layer, so that the robot can notice whether it is making progress. At present, five modules are set up, namely, a cross module for coordinating and reflecting deliberation behavior, a fault supervision module, a supervision module for the execution of behavior at the reaction layer, a supervision module for behaviour planning at the deliberation layer, and a learning evolution unit. The learning evolution unit has two functions. One is to learn the local optimal path planning behaviour in a static environment of deliberation behaviour, and obtain the design of reactive behaviour autonomously. The second is to use the collected sample data for training and learning in a dynamic

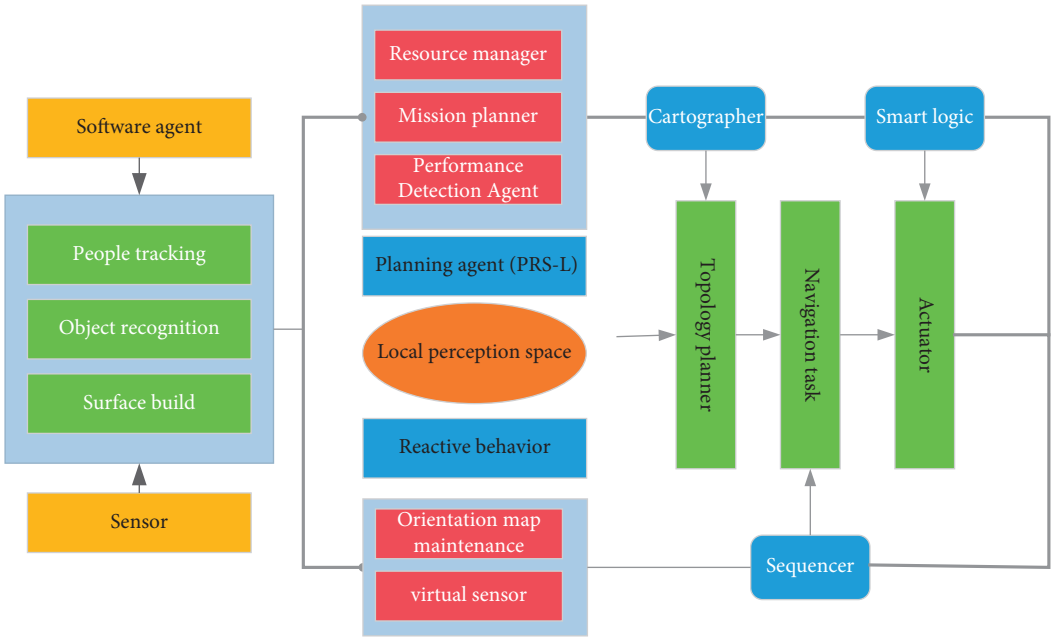


Figure 1: Intelligent obstacle avoidance system diagram.

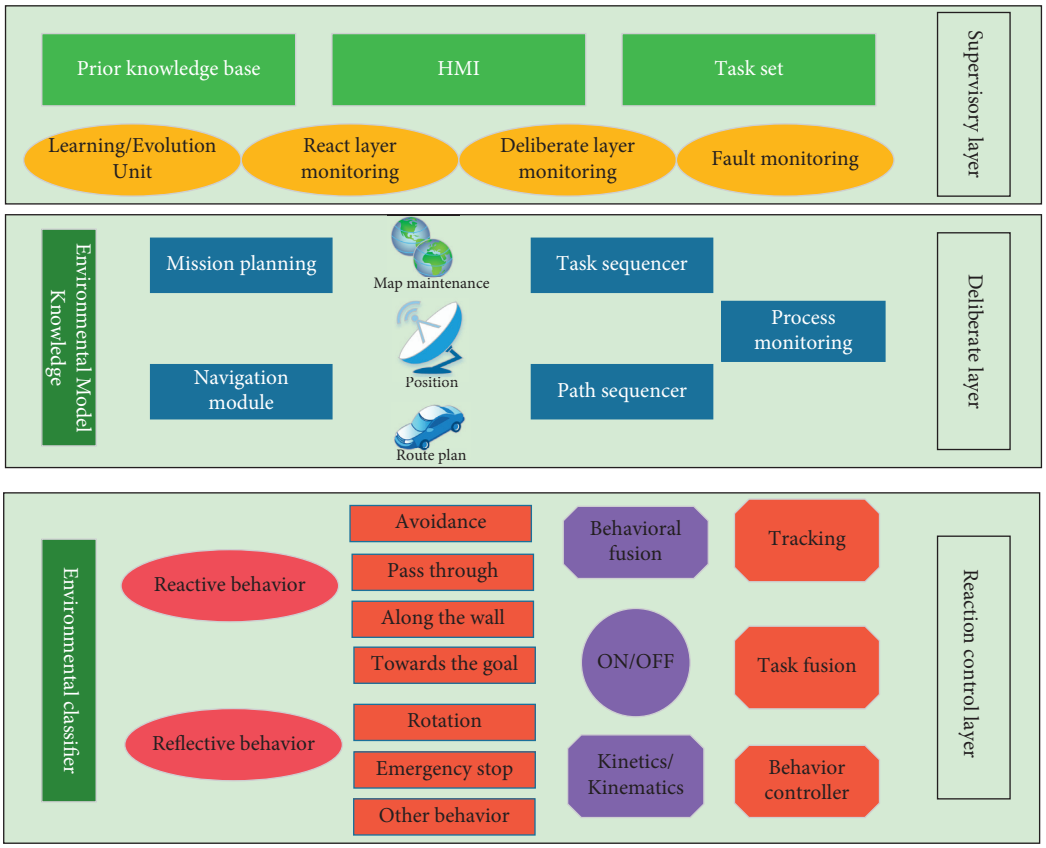


Figure 2: Mobile robot hierarchy diagram.

environment to establish a prediction model for dynamic obstacle avoidance. Learning the functions of evolution could enable robots to autonomously adapt to environmental changes and enhance their intelligence.

3.4. Obstacle Avoidance Control Process. In robot route planning, various information obtained from the robot itself and its environment is synthesized, enabling the robot to understand its environment and make decisions through

controller processing. So as to avoid obstacles, find the optimal path and move autonomously [20]. Figure 3 shows the flowchart of the mobile robot path planning.

The mobile robot is equipped with a positioning system that can detect the global position and heading of the mobile robot. Six ultrasonic sensors are used to detect local obstacle information. The mobile robot performs detection every 1s, and the acquired sensor data is fused as the input of the controller. By verifying the effectiveness of the previously mentioned improved fuzzy neural network algorithm, a system simulation model is established in Mat lab with Simulation, the control rules are simulated, and the fuzzy logic toolbox software is used to simulate the neural network algorithm. In this simulation system, the steps of path planning simulation are as follows.

(1) Establish environmental information: environmental information is to establish the coordinates, dimensions of obstacles, and the starting point and target point of the robot.

(2) Establish a simulated robot, including some parameters such as robot size and moving speed. According to the kinematic model of the robot in the actual system, it is assumed that the travelling speed of the mobile robot is 0.6 m/s, and it can realize 360° in situ.

(3) Establish a simulated sensor: used to perceive the simulated environment information, that is, to obtain the value of the obstacle distance information d and the target direction angle ?.

(4) The robot controller is designed by an improved fuzzy neural network algorithm, and the driving instructions of the robot are obtained by analyzing the obstacle and target data obtained by the sensor.

(5) Transmission of control instructions: the driving instructions are transmitted to the simulated robot, the mobile robot moves according to the instructions, and then the map coordinates of the robot after executing the driving instructions are calculated; repeat steps (3) to (5) until the robot reaches the predetermined target point.

4. Intelligent Obstacle Avoidance Algorithm Results and Analysis

4.1. Algorithm Accuracy Analysis. In this experiment, the two network models before and after the improvement were used for training, respectively, and the training results were compared after 300 times of effective training. Before training, you need to import the saved network model data in advance, and use the training data set to continue training the network model before the improvement. At the same time, the loss value and the accuracy curve are obtained, and the accuracy curve of the test set is obtained, as shown in Figure 4.

It can be seen from Figure 4 that since the training, the loss value has been continuously reduced and the accuracy has been continuously improved in the process of reverse gradient learning. This is a supervised learning process. Before adding the $1 * 1$ constitutional layer, after multiple straining, the final test set accuracy rate was stable at about 95%. After adding the $1 * 1$ constitutional layer, the network is retrained, the test set accuracy rate is stable at 98%, and the loss function value is also reduced from the original 0.79 to the current 0.08, which is 10 times smaller. Since the added $1 * 1$ constitutional layer contains nonlinear units, the nonlinear expression ability of the model is improved. After adding $1 * 1$ to two different curves, the fluctuation is obviously reduced compared with the previous curve, and it is relatively stable.

4.2. Sensor Property Index Analysis. The sensor control system and the information processing system constitute the robot's obstacle avoidance system. Among them, the sensor control system is composed of sensors and microprocessors that obtain environmental information. It mainly collects information from unknown environments, which is the only way for robots to understand environmental information. Therefore, the reasonable configuration of the sensor determines the accuracy of the system's acquisition of the external environment.

As can be seen from Figure 5, the radar has the best performance index and high cost. Ultrasonic and infrared have the lowest cost. In view of the need for obstacle avoidance, combined with the cost of the sensor, the performance index of the collected information, the hardware implementation circuit, volume, and other comprehensive factors, the sensor control system in this paper mainly uses ultrasonic and infrared sensors to collect the obstacle distance of the external environment and electronic compass to obtain the target object. Considering the unknowns and complexity of the external environment, the multipurpose technology is applied to the mobile robot obstacle avoidance system, and the hardware and algorithms are processed accordingly to enhance the intelligence of the robot.

4.3. Path Obstacle Avoidance Optimization Analysis. In a complex environment with obstacles, the ultimate goal of path planning is to solve an optimal route, so that the robot can move smoothly and avoid all obstacles without collision [21]. In practical applications, the environment is unknown to the mobile robot. The unknown environment, the detection accuracy of the obstacle detector and the difference of the path planning algorithm have a great influence on whether the path planning can be successfully implemented. This paper compares and analyzes the experimental paths of the BUG algorithm and the improved fuzzy neural network algorithm.

It can be seen from Figure 6 that when the BUG algorithm is used, the robot moves along the edge contour of the obstacle. When it can move to the target point, it will directly leave the edge of the obstacle and move beyond the target point. It is not limited to the distance judgment between the obstacle and the target point. H1 and H2 are the arrival points, and L1 and L2 are the separation points. When using

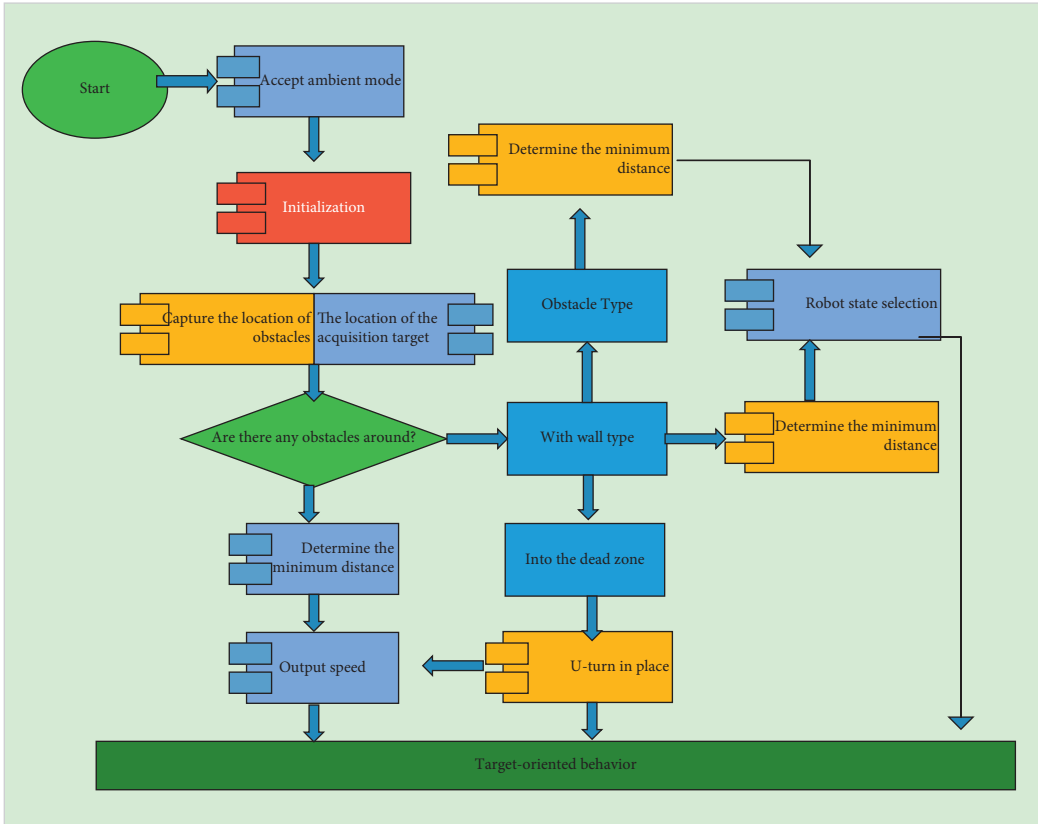


Figure 3: Flowchart of intelligent obstacle avoidance path.

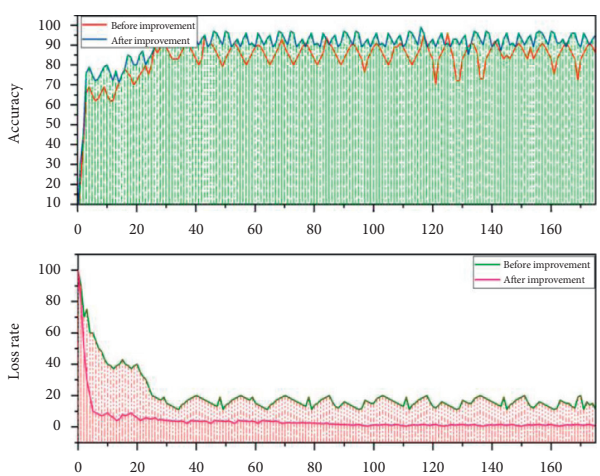


Figure 4: Model loss value and accuracy curve.

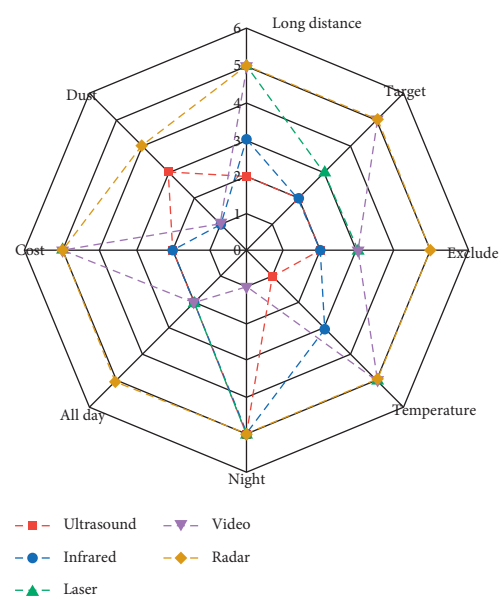


Figure 5: Sensor performance graph.

the intelligent algorithm, the robot initially moves along the line connecting the starting point and the target point. When encountering an obstacle, it moves along the tangential direction of the obstacle until no obstacle is detected and then updates the current point to the target point. It extends the updated main line and repeat this process until the target point is reached. After comparison, it is found that the improved fuzzy neural network algorithm in this paper, the path distance is dominated by straight lines, the distance is shorter, and the distance is shortened by 5–10m.

4.4. Obstacle Avoidance Algorithm Simulation. In order to verify the performance of the algorithm, a lot of simulation research has been done in this paper. Let the values of N and M be 4, 10, 30, and 60, respectively, the selection probability is 0.1, the cross-mutation probability is 0.08, the a and β values are both 1, and the population evolution stops at 200

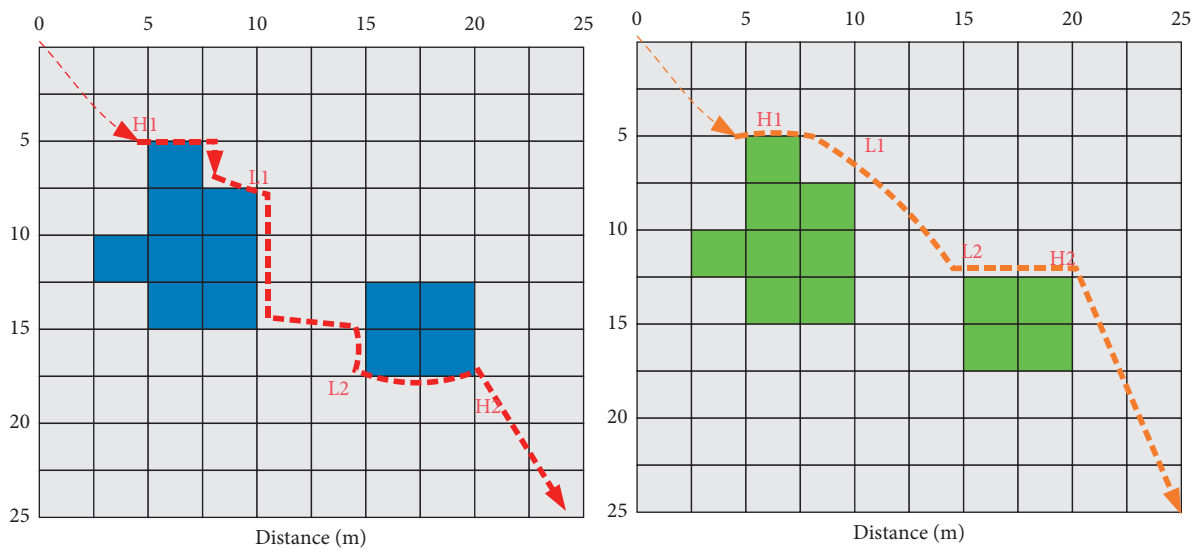


Figure 6: Comparison and analysis diagram of path obstacle avoidance optimization.

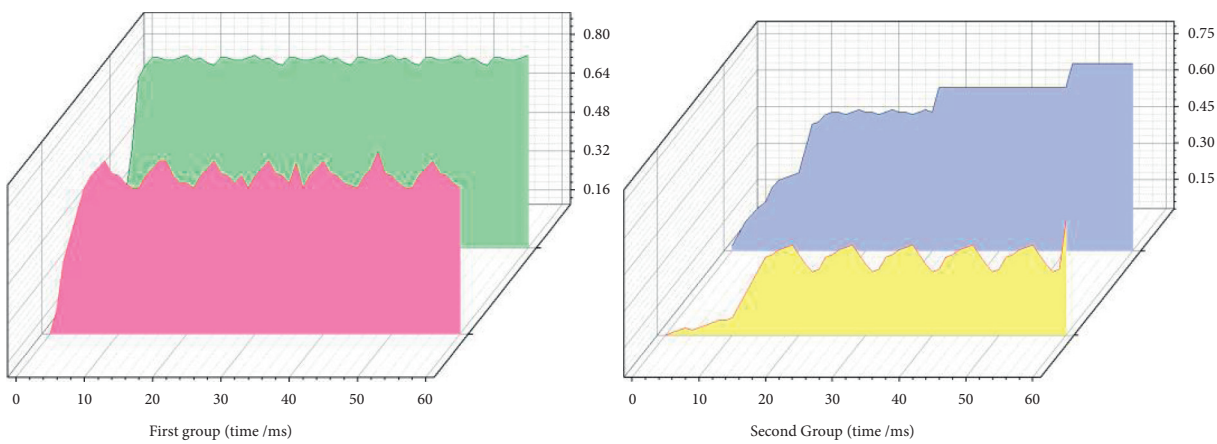


Figure 7: Obstacle avoidance algorithm simulation curve.

or 1000 generations. Research the convergence performance of the algorithm evolution under the same parameters: the evolution algebra when the algorithm converges and the execution time is consumed by the evolution calculation. The changes of the maximum fitness value during the evolution of each group of experiments are shown in Figure 7.

It can be seen from the figure that, with the increase of N : M, the method will gradually improve the search efficiency. The larger N : M is, the larger the solution space is, but the convergence speed does not change, which shows that the search efficiency of this method is the largest as the value increases. When the solution space of the problem is large, the speed of convergence is obviously slowed down, the algebra of convergence is improved, and the convergence time is prolonged. This further verifies that the method is suitable for medium and large scale problem spaces. In summary, the algorithm can handle any form of medium-to-large-scale task assignment problem.

5. Conclusion

The ability to autonomously avoid obstacles is the main indicator to measure the intelligence of mobile robots, and it is also an important condition for intelligent robots to drive safely [22]. This paper is the first to propose an intelligent obstacle avoidance algorithm for mobile robots based on improved fuzzy neural network, which achieves precise positioning through adaptive learning. Following the general principles of intelligent control system design, on the basis of response and deliberation, a monitoring layer is added. Supervise and coordinate the implementation of deliberative layer behaviors to learn adaptive behaviors in unknown environments. By collecting sample data for training, it learns to build predictive models to avoid dynamic obstacles. Through the obstacle avoidance experiment simulation, the accuracy and real-time performance of the system are further verified, and finally, a good obstacle avoidance effect is achieved.

Data Availability

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

Conflicts of Interest

The authors declare that there are no conflicts of interest or ethics in this article.

Acknowledgments

This study was supported by the Key Project of Industry Guidance of Science and Technology Department of Fujian Province"R and Dand application of large space automatic tracking and positioningJet Fire Control System for Intelligent Security"(no.2020Y0021).

References

[1] S. M. H. Rostami, A. K. Sangaiah, and J. Wang, "Obstacle avoidance of mobile robots using modified artificial potential field algorithm," *EURASIP Journal on Wireless Communications and Networking*, vol. 22, no. 4, pp. 1–19, 2019.

[2] B. K. Patle, A. Pandey, A. Jagadeesh, and D. R. Parhi, "Path planning in uncertain environment by using firefly algorithm," *Defence technology*, vol. 14, no. 6, pp. 691–701, 2018.

[3] M. Radmanesh, M. Kumar, and P. H. Guentert, "Overview of path-planning and obstacle avoidance algorithms for UAVs: a comparative study," *Unmanned Systems*, vol. 6, no. 12, pp. 95–118, 2018.

[4] A. Nasrinahar and J. H. Chuah, "Intelligent motion planning of a mobile robot with dynamic obstacle avoidance," *Journal on Vehicle Routing Algorithms*, vol. 1, no. 2, pp. 89–104, 2018.

[5] A. Adib and B. Masoumi, "Mobile robots navigation in unknown environments by using fuzzy logic and learning automata," *Artificial Intelligence and Robotics*, vol. 56, no. 15, pp. 111–123, 2017.

[6] Q. Wu, H. Lin, Y. Jin, Z. Chen, S. Li, and D. Chen, "A new fallback beetle antennae search algorithm for path planning of mobile robots with collision-free capability," *Soft Computing*, vol. 24, no. 3, pp. 2369–2380, 2020.

[7] L. Zhang, J. Wang, and Z. Lin, "Distributed cooperative obstacle avoidance for mobile robots using independent virtual centre points," *Journal of Intelligent and Robotic Systems*, vol. 98, no. 9, pp. 791–805, 2020.

[8] J. Y. Jhang, C. J. Lin, and C. T. Lin, "Navigation control of mobile robots using an interval type-2 fuzzy controller based on dynamic-group particle swarm optimization," *International Journal of Control, Automation and Systems*, vol. 16, no. 10, pp. 2446–2457, 2018.

[9] F. Shamsfakhr and B. Sadeghibigham, "A neural network approach to navigation of a mobile robot and obstacle avoidance in dynamic and unknown environments," *Turkish Journal of Electrical Engineering and Computer Sciences*, vol. 25, no. 23, pp. 1629–1642, 2017.

[10] S. Muthukumaran and R. Sivaramakrishnan, "Optimal path planning for an autonomous mobile robot using dragonfly algorithm," *International Journal of Simulation Modelling*, vol. 18, no. 4, pp. 397–407, 2019.

[11] R. Zhao and H. K. Lee, "Fuzzy-based path planning for multiple mobile robots in unknown dynamic environment," *Journal of Electrical Engineering and Technology*, vol. 12, no. 11, pp. 918–925, 2017.

[12] G. Lee and D. Chwa, "Decentralized behaviour-based formation control of multiple robots considering obstacle avoidance," *Intelligent Service Robotics*, vol. 11, no. 9, pp. 127–138, 2018.

[13] Z. Yan, J. Li, and G. Zhang, "A real-time reaction obstacle avoidance algorithm for autonomous underwater vehicles in unknown environments," *Sensors*, vol. 18, no. 12, pp. 438–448, 2018.

[14] J. Jin and W. Chung, "Obstacle avoidance of two-wheel differential robots considering the uncertainty of robot motion on the basis of encoder odometer information," *Sensors*, vol. 19, no. 17, pp. 289–376, 2019.

[15] M. N. Zafar and J. C. Mohanta, "Methodology for path planning and optimization of mobile robots: a review," *Procedia Computer Science*, vol. 133, no. 32, pp. 141–152, 2018.

[16] M. Cui, H. Liu, and W. Liu, "An adaptive unscented kalman filter-based controller for simultaneous obstacle avoidance and tracking of wheeled mobile robots with unknown slipping parameters," *Journal of Intelligent and Robotic Systems*, vol. 92, no. 46, pp. 489–504, 2018.

[17] D. Chen, "Fuzzy obstacle avoidance optimization of soccer robot based on an improved genetic algorithm," *Journal of Ambient Intelligence and Humanized Computing*, vol. 11, no. 12, pp. 6187–6198, 2020.

[18] Z. Yi, G. Li, and S. Chen, "A navigation method for mobile robots using interval type-2 fuzzy neural network fitting Q-learning in unknown environments," *Journal of Intelligent and Fuzzy Systems*, vol. 37, no. 35, pp. 1113–1121, 2019.

[19] W. Zheng, H. B. Wang, and Z. M. Zhang, "Multi-layer feedforward neural network deep learning control with hybrid position and virtual-force algorithm for mobile robot obstacle avoidance[J]," *International Journal of Control, Automation and Systems*, vol. 17, no. 8, pp. 1007–1018, 2019.

[20] D. Q. Bao and I. Zelinka, "Obstacle avoidance for swarm robot based on self-organizing migrating algorithm," *Procedia Computer Science*, vol. 150, no. 22, pp. 425–432, 2019.

[21] A. K. Rath, D. R. Parhi, H. C. Das, M. K. Muni, and P. B. Kumar, "Analysis and use of fuzzy intelligent technique for navigation of humanoid robot in obstacle prone zone," *Defence Technology*, vol. 14, no. 6, pp. 677–682, 2018.

[22] D. C. Rao, M. R. Kabat, and P. K. Das, "Hybrid IWD-DE: a novel approach to model cooperative navigation planning for multi-robot in unknown dynamic environment," *Journal of Bionics Engineering*, vol. 16, no. 10, pp. 235–252, 2019.

Hindawi
Journal of Robotics
Volume 2022, Article ID 8514235, 9 pages
https://doi.org/10.1155/2022/8514235

Research Article

Network Resource Allocation Strategy Based on UAV Cooperative Edge Computing

Shuo Wang and Ning Kong

School of Energy and Intelligence Engineering, Henan University of Animal Husbandry and Economy, Zhengzhou, Henan 450044, China

Correspondence should be addressed to Shuo Wang; 81239@hnuahe.edu.cn

Received 9 February 2022; Revised 9 March 2022; Accepted 10 March 2022; Published 29 March 2022

Academic Editor: Shan Zhong

Aiming at the problem that fixed mobile edge computing (MEC) server is difficult to meet the needs of mobile users and temporary computing services, this study proposes a network resource allocation strategy based on unmanned aerial vehicle (UAV) cooperative edge computing. First, a UAV-aided MEC scene is designed, and a single UAV with an MEC server is used to provide auxiliary computing services for ground multiusers. Then, an optimization model aiming at total system delay is constructed by considering the system communication model and calculation model. Finally, Deep Q-Network is used to solve the optimization problem to obtain the best resource allocation scheme. Based on the experimental platform, the proposed strategy is demonstrated and analyzed. The results show that when the number of user equipment is 40, the total delay is about 33s, which is 35.29%, 31.25%, and 15.38% lower than other comparison strategies and effectively reduces the computing delay of users.

1. Introduction

In recent years, with the development of 5G mobile communication technology and the Internet of Things (IoT), smart mobile equipment have shown an explosive growth [1–3]. In the context of cloud computing, MEC is considered to be a technical key to improving the computing efficiency of mobile edge equipment [4]. In MEC systems, task offloading is the key for mobile equipment to support resource-intensive applications by executing on edge cloud resources [5, 6]. However, computing servers are usually deployed in fixed base stations. In practical scenarios, fixed base stations cannot meet dynamic services such as user mobility, base station damage, and temporary hotspot areas. It is particularly significant to meet the dynamic communication and multiequipment access requirements.

According to the existing research, computing servers can be deployed on an unmanned aerial vehicle (UAV) to meet the communication requirements. UAV themselves have the advantages of low cost and high mobility. In UAV-assisted MEC networks, mobile equipment can offload tasks to UAV with high computing power and flexible connectivity at the network edge [7, 8]. This method utilizes the flexibility of UAV and better channel gain to offload computing tasks for users, which can not only save computing delay and user energy but also reduce the traffic load on fixed cloud servers [9]. Reference [10] proposed a two-stage joint hovering altitude and power control solution for the resource allocation problem in UAV networks considering the inevitable cross-tier interference from space-air-ground heterogeneous networks.

At present, significant progress has been made on the research of computing task offloading. For example, reference [11] proposed a collaborative service deployment and application allocation algorithm to achieve the final edge service policy deployment. The minimum energy consumption was obtained through the minimum resource ratio increasing algorithm, and computing tasks are redistributed in combination with the load balancing algorithm to balance the computing load. However, its overall computational efficiency needs to be improved. Reference [12] proposed a computing framework for coordinating terminals, edge

nodes, and cloud centers based on the pipeline offloading scheme. According to the computing and communication capabilities of the entire network, they reasonably allocated computing-intensive tasks to specific terminals or clouds, effectively improving computing efficiency. Reference [13] proposed a vehicle-assisted computational offloading architecture for UAVs. The proposed framework used vehicle-assisted computing offload for UAV computing tasks and network resource optimization, which has commonality with the research topic. Reference [14] introduced agents into computing task offloading and proposed a UAV-MEC (UMEC) agent-enabled computing task offloading framework to help users, UAV, and edge clouds perform computing task offloading. Reference [15] proposed a computational offloading scheme to minimize the time and energy consumption of computational task costs. The scheme reduced the execution cost by coordinating the allocation of computing resources between mobile equipment and edge servers. However, many iterations were needed to find the optimal solution. Reference [16] designed a multiround iterative auction algorithm based on auction theory. But the overall performance of the algorithm needs to be further optimized. Reference [17] proposed an optimized auction-based incentive mechanism. The mechanism can optimize long-term system welfare by operating in an online fashion. However, it does not have good scalability for resources that cannot be covered by network hardware.

With the continuous development of computer technology, intelligent algorithms such as machine learning are continuously applied to edge computing. As in reference [18], a heuristic-based algorithm with a low time cost is proposed. Compared with the existing centralized resource allocation and decision-making algorithms, this scheme acquired a higher number of successfully offloaded tasks in different scenarios, but the offloading of computing tasks in complex situations cannot achieve collaborative optimization. Reference [19] proposed a blockchain-driven collaborative framework for MEC. Reference [20] proposed a caching mechanism based on Q-learning to reduce the backhaul traffic load and transmission delay from the cloud. Reference [21] proposed a collaborative computing framework based on deep neural networks. The experimental results indicate that the proposed method has batter effectiveness than the traditional methods. The above methods based on machine learning not only consider the efficiency and timeliness of MEC network resource allocation to a certain extent but also do not have good scalability. Therefore, the offloading strategy based on UAV has been further researched and developed. Reference [22] proposed a deep reinforcement learning-based MEC UAV-assisted computing offloading. The total cost minimization was taken as the objective function. Although this method expands the wireless network to a certain extent, it needs to be further optimized for efficient offloading of computing tasks and reasonable allocation of network resources in complex IoT.

Aiming at the problem that it is difficult for fixed MEC servers to meet the computing needs of mobile users and computing efficiency of existing UMEC, a network resource allocation strategy based on UAV collaborative edge computing was proposed. Compared with the traditional network resource optimization allocation strategies, the innovations of this study are as follows:

(1) For the low-latency service guarantee for users, the UAV carrying MEC server is designed to improve the auxiliary edge computing system model for users. Taking the total system delay as the optimization goal, the transmission delay of user calculation is greatly shortened.

(2) Due to the huge amount of data and high real-time requirements of the 5G system, the proposed strategy uses Deep Q-Network (DQN) to solve the optimal resource allocation scheme, which improves the analysis efficiency of the network state.

2. System Model and Optimization Goal

2.1. System Model. The scenario of human-machine-assisted MEC system uplink communication is studied, as shown in Figure 1, that is, UAVs provide auxiliary computing services for multiple ground user equipment. Only the scenario where a single rotor UAV provides services to N user equipment on the ground is considered.

For the convenience of theoretical analysis, a three-dimensional Cartesian coordinate system is considered, where $L_n = (x_n, y_n)$ represents the position of n user equipment on the ground. The drone flies in a circle within a specified range and provides auxiliary computing services to ground user equipment. Among them, UAV is at a fixed height and the initial position of UAV that is denoted as $d[1]$, and the final position is denoted as $d[K]$. Each ground user equipment offloads some of computing tasks to UAV and the rest is locally computed.

It is assumed that the total execution time is denoted T, and it can be equally divided into K slots and denoted $t = \{1, 2, \ldots, K\}$, where $\tau = T/K$, where τ refers to the length of each slot. In the t slot, the position of drone is $d[t] = (x[t], y[t])$. Assuming that the maximum flight speed of UAV is $V_{\max}$, the trajectory constraints of UAV are as follows:

$$d[1] = d[K],$$
$$\frac{\|d[t+1] - d[t]\|}{\tau} \leq V_{\max}, \tag{1}$$

where formula (1) indicates that the flight speed of the UAV cannot exceed its maximum flight speed. Since UAV flight energy consumption is related to UAV flight speed, UAV flight energy consumption expression is as follows:

$$E^{\text{fly}}[t] = \eta \|v[t]\|^2, \tag{2}$$

where $v[t] = [\|d[t+1] - d[t]\|/\tau]$ represents the flight speed of UAV in the t time slot; η represents the proportional factor of energy consumption and flight speed.

2.2. Communication Model. Assuming that the wireless channel between UAV and each user equipment is a line-of-

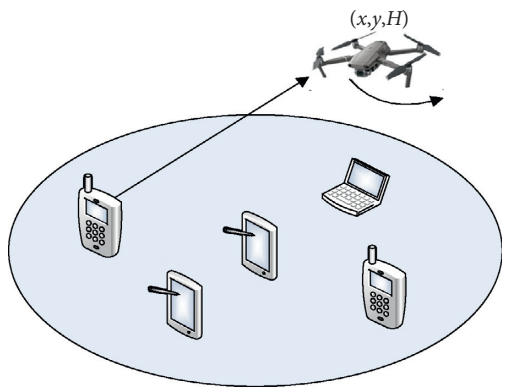


Figure 1: System model.

sight (LOS) channel, the channel power gain between UAV and user equipment n can be expressed as follows:

$$h_n[t] = \beta_0 d_n^{-2}[t] = \frac{\beta_0}{H^2 + \left\| d[t] - L_n \right\|^2}, \tag{3}$$

where β_0 represents the channel power gain per unit distance; $d_n^{-2}[t]$ represents the horizontal distance between UAV and user equipment n in the t gap; and $\| \cdot \|$ represents Euclidean norm.

2.3. Computational Model

2.3.1. Local Computing. When user equipment n selects the local mode, all computing tasks will be locally performed, and the computing delay can be expressed by the following formula:

$$T_n^{\text{loc}} = \frac{C_n D_n}{f_n}, \quad n \in N, \tag{4}$$

where f_n is the computing capability of user equipment n (number of cycles per second); C_n represents the number of cycles required to calculate 1 bit; and D_n is the size of input data.

2.3.2. Offload to UAV for Calculation. The UAV carrying MEC server periodically flies at a fixed altitude above the ground user equipment. The user equipment that selects MEC mode can offload tasks to UAV through time-division multiple access (TDMA) [23, 24]. The set of ground users is set to be offloaded to MEC as N', and UAV can choose to relay some tasks to access point (AP) for calculation. Therefore, it can be assumed that the ratio of user equipment n that selects the relay to the ground base station in the t time slot is $\zeta_n[t]$, $n \in N'$, $t \in K$; then, the task that needs to be calculated by UAV in the t time slot is $(1 - \zeta_n[t])D_n[t]$.

(1) The transmission speed of user equipment n in time slot t:

$$v_n^u[t] = B_1 \log_2\left(1 + \frac{P_n h_n[t]}{\sigma^2}\right), \quad n \in N', \tag{5}$$

where B_1 is the bandwidth between UAV and user n; P_n is the maximum transmission power of the user n, and the user equipment uses the maximum power transmission; and $h_n[t]$ is t time slot channel gain between the user n and UAV communication.

(2) Transmission delay of user equipment n: assuming that the bits offloaded by the user n in time slot t are $D_n[t]$, so the user n needs to offload all bits in K time slots, and the constraints are expressed as follows:

$$\sum_{n=1}^{N} D_n[t] = D_n. \tag{6}$$

The total transmission delay of the user equipment is expressed as follows:

$$T_n^{\text{mec}} = \sum_{t=1}^{K} \frac{D_n[t]}{v_n[t]}. \tag{7}$$

(3) The calculation delay of UAV: according to the relay ratio, the task calculated by UAV for user equipment n in the t time slot and then the calculation delay of user equipment in one time slot in MEC mode can be obtained, and finally, the calculation task of UAV for a single user equipment can be obtained. The total delay is as follows:

$$T_k^u = \sum_{t=1}^{K} \frac{(1 - \zeta_n[t])D_n[t]C_n}{f_c}, \tag{8}$$

where f_c represents the computing power of MEC server.

2.3.3. Offload to Base Station on the Ground for Calculation. UAVs can choose to relay part of computing tasks to AP computing. The proportion of UAVs choosing relays in time slot t is $\zeta_n[t] \in [0, 1], n \in N'$. Since the ground base station can have multiple MEC servers built in, the calculation delay of ground base station is considered to be ignored. At the same time, the result of the calculation task is usually very small, and it is considered to ignore the transmission delay of calculation result. Thus, the delay of UAV relaying tasks to the computing part of the ground base station only includes the transmission delay of the UAV.

The flying height of the UAV is relatively high from the ground base station and has a good line-of-sight link. It is considered that the communication between the UAV and ground base station adopts LOS channel. The transmission rate of UAV relaying the task to AP in the time slot is calculated as follows:

$$v_n^{U2E}[t] = B_2 \log_2\left(1 + \frac{P_u h_u[t]}{\sigma^2}\right), \tag{9}$$

where B_2 is the bandwidth between UAV and AP; P_u is the transmission power of UAV; and $h_u[t]$ is the communication channel between UAV and AP.

The transmission delay from UAV relay to ground AP can be expressed as follows:

$$T_n^{U2E} = \sum_{t=1}^{K} \frac{\zeta_n[t] D_n[t]}{v_n^{U2E}[t]}. \tag{10}$$

Assuming that UAV can calculate and transmit part of computing tasks to the ground AP at the same time, the total calculation delay offloaded to UAV is as follows:

$$T_n^{\max} = \max\left\{T_n^{U2E}, T_n^{u}\right\}. \tag{11}$$

When the user equipment selects MEC mode, the total delay of offloading computing tasks to UAV is as follows:

$$T_n^{\Omega} = T_n^{\text{mec}} + T_n^{\max}. \tag{12}$$

2.4. Model Constraint Description. The system goal is to minimize the total delay of all requesting user equipment, so the system goal can be defined as follows:

$$T_{\sum} = \sum_{n=1}^{N} T_n, \tag{13}$$

where T_n represents the calculation delay of the user n.

Considering different computation and offloading modes, T_n is defined as follows:

$$T_n = (1 - \alpha_n) T_n^{\text{loc}} + \alpha_n T_n^{\Omega}, \tag{14}$$

where α_k represents the mode selected by the user equipment. When the variable is 1, it means that the offloading to the drone is selected, and when the variable is 0, it means that the local calculation is selected. T_n^{loc} represents the total delay locally calculated by the user equipment; T_n^{Ω} represents the total delay calculated by the user to select all offload to the drone.

Optimization of user equipment offloading strategy, UAV relay ratio, UAV trajectory, and user bit allocation are combined to minimize the total delay of user equipment. The mathematical expression is as follows:

$$\begin{aligned}
&\text{P1: } \min \sum_{n=1}^{N} \left((1-\alpha_n) T_n^{\text{loc}} + \alpha_n T_n^{\Omega}\right),\\
&\text{C1: } \alpha_n \in \{0,1\}, \quad n \in N,\\
&\text{C2: } \zeta_n[t] \in [0,1], \quad n \in N',\\
&\text{C3: } T_n \le T_n^{\max}, \quad n \in N,\\
&\text{C4: } \sum_{t=1}^{K} D_n[t] = D_n, \quad n \in N',\\
&\text{C5: } d[1] = d[K],\\
&\text{C6: } \|d[t+1] - d[t]\|^2 \le (v_{\max}\Delta)^2, \quad t \in \tau,\\
&\text{C7: } \textstyle\sum_{n=1}^{\dot{N}} D_n[t] \le C_u, \quad t \in K,
\end{aligned} \tag{15}$$

where C1 represents that the user equipment n can only choose one mode; C2 represents that the relay ratio of UAV is a variable between 0 and 1. C3 represents that the calculation delay for each user equipment n must be less than the maximum tolerated delay. C4 represents that the user equipment n that selects the MEC mode needs to complete the offloading of tasks within K time slots. C5 means that the drone flies in cycles; C6 means that the maximum horizontal distance of the drone in one time slot cannot exceed the threshold. C7 represents that the bits offloaded by all user equipment to UAV in a time slot cannot exceed the computational threshold of UAV.

3. Solutions Based on Deep Reinforcement Learning

3.1. Reinforcement Learning Modeling. The optimization problem in equation (15) is solved by a single-agent Q-learning algorithm based on reinforcement learning. The algorithm model consists of four parts, namely, agent, state, action, and reward [25].

(1) Agent: Using UAV as the agent of the algorithm, UAV will be responsible for collecting the information of each equipment in the system and making scheduling decisions.

(2) Action: $a = [\lambda, \varpi_n]$ is defined to represent each action decision variable of UAV. Among them, $\lambda = n$ indicates that the drone will provide on-demand services for the n user equipment in the current state, and $\lambda \in N$ is satisfied. ϖ_n represents the computing mode of required computing task of n user equipment.

(3) State: The state variable $s = [\omega, z, t^{\text{serve}}, t^{\text{fly}}]$ is defined to represent the service state in the system. The state quantity consists of the following four parts: the user equipment serves state quantity $\omega = [\omega_1, \ldots, \omega_N]$ and satisfies $\omega_n \in \{0, 1\}$. ω_n represents whether the n user equipment in the system has been serviced by drone and the service status has been completed, $\omega_n = 1$ represents that the drone has completed the on-demand service for the first user equipment, and otherwise, $\omega_n = 0$. The maximum tolerant delay amount $z = [z_1, \ldots, z_N]$ and satisfies $z_n = T_n^{\max}$, that is, z_n represents the maximum tolerant delay of n user equipment in the system. Waiting service delay amount $t^{\text{serve}} = [t_1^{\text{serve}}, \ldots, t_N^{\text{serve}}]$; flight delay amount $t^{\text{fly}} = [t_1^{\text{fly}}, \ldots, t_N^{\text{fly}}]$.

(4) Rewards: The proposed optimization problem aims at maximizing the number of network user equipment services under UAV-MEC system architecture, while the proposed Q-learning algorithm based on reinforcement learning aims to obtain maximum reward feedback. Combining the above two points, and based on the current system state variable, $s = [\omega, z, t^{\text{serve}}, t^{\text{fly}}]$ and the selected action variable $a = [\lambda, \varpi_n]$. The system reward feedback function is defined as $r(s, a) = \sum_{n=1}^{N} \omega_n$, that is, when UAV selects the action decision variable as a in the state s, the total number of network equipment in the system can meet the service requirements [26, 27].

The training iterative process of the Q-learning algorithm based on reinforcement learning satisfies the Bellman equation. The selection principle of action decision performed by UAV under different state variables is based on ε-greedy mechanism.

3.2. DQN-Based Offloading Strategy Optimization Algorithm. The pseudocode of the DQN-based offloading strategy optimization algorithm is shown in Algorithm 1.

First, the network parameters are initialized, including the capacity C_1 of replay experience pool, the parameter θ of action value function $Q(s, a; \theta)$, the parameter $\widehat{\theta} = \theta$ of target action value function $\widehat{Q}(s, a; \widehat{\theta})$, and the initial state s. Then, at each episode t, each user equipment randomly picks an action from the feasible action space according to ε-greedy policy. A random action with probability ε or an action satisfying the formula $a = \arg\max Q(s, a; \theta)$.

Then, resource allocation is performed according to the action selected by user equipment. If the local processing task is selected, let the local CPU cycle frequency be the maximum computing power of equipment, namely, $F_n = f_n^{\text{loc}}$. If the user equipment chooses to upload tasks to UAV for processing, let its uplink transmission power be the maximum available power, i.e., $p_n = P_n$.

After the resource allocation is completed, reward r can be calculated according to the designed reward function, a new state s' can be obtained, and a new sample (s, a, r, s') can be stored in the experience pool. Finally, a batch of samples is randomly sampled from the experience pool, and these samples are used to update the parameters θ of Q-Network, synchronizing the parameters $\widehat{\theta} = \theta$ of target Q-Network every e steps. Note that although each user equipment independently selects actions, since the resources of UAVs are shared, the results of their action selections influence each other [28, 29]. Assuming that all users have selected the same drone for association, the drone may be overloaded, resulting in the user's equipment mission not being able to complete within the specified latency limit, or the drone running out of energy. In order to balance the load as much as possible to avoid this extreme situation, an invalid action pool is added. After all user equipment complete the action selection, if the action set composed of all the user equipment cannot meet the energy consumption constraints of the base station or the load is too concentrated so that most of the user equipment cannot meet the delay limit $T_n^{\max}$, the action set is added to invalid action pool [30]. In this way, when the user equipment performs action selection again, if the current action combination has already appeared in an invalid action pool, it is ignored and the selection is made again. This avoids invalid calculations to improve efficiency.

4. Experiments and Analysis

Simulation results are used to evaluate the performance and efficiency of the proposed strategy. It is assumed that the user equipment is randomly distributed in a two-dimensional area of 200 m × 200 m, and the AP is located in the upper right corner of the two-dimensional area. Each user equipment has different computing tasks, time delay tolerance, and the number of cycles required for computing a 1 bit task. The specific parameters are shown in Table 1.

4.1. Average Reward of Algorithm Training Process. The average reward value of multiple episodes in the training phase of the DQN-based offloading strategy optimization algorithm is shown in Figure 2. In order to increase the exploratory nature, a certain amount of noise is added to the action during the experiment, which makes the original image have many glitches and become unsmooth. Figure 2 shows the reward curve after taking the moving average.

It can be seen from Figure 2 that the proposed algorithm has converged at 2,000 time slots, and the average reward value fluctuates around 220. It can be seen that the algorithm has a fast convergence speed and an ideal average reward value.

4.2. Relationship between the Number of Convergence Iterations and the Number of IoT Equipment. As the number of equipment in the system increases, the number of times that the proposed algorithm training reaches convergence also increases, as shown in Figure 3.

As can be seen from Figure 3, when the number of system equipment increases, more iterations are required to achieve convergence. But when the number of equipment exceeds 35, the increase in the number of iterations increases. This is mainly because the Q-table size of the algorithm is closely related to the number of equipment present in the system. When the number of equipment exceeds the threshold, the processing pressure of the algorithm sharply increases, resulting in a rapid increase in the number of iterations required for convergence. However, the algorithm can achieve convergence regardless of the number of equipment, thus demonstrating the effectiveness of the proposed algorithm.

4.3. Relationship between the Calculated Energy Efficiency and Maximum Energy Consumption of Different Loading Models. The relationship between the computing energy efficiency of user equipment and the maximum energy consumption under different loading models is shown in Figure 4.

As can be seen from Figure 4, the local computing model is that the user equipment only performs local computing, while the global offloading model is that the user equipment completely offloads computing tasks to UAV for computing. Both schemes optimize the flight trajectory of UAV. Using the DQN-based offloading optimization algorithm for partial offloading can achieve higher computational energy efficiency. When the maximum energy consumption is 4J, the computing energy efficiency of its user equipment exceeds 200 bits/J. This is because the user equipment can flexibly allocate resources according to the quality of channel state information, so as to select offload computation or local computation under the partial offload model. Furthermore, the global offload model outperforms the local computing

Initialization Experience pool capacity C_1, Invalid action pool capacity C_2, Parameters θ of action value function Q, Parameters $\widehat{\theta} = \theta$ of target action value function $\widehat{Q}$, Initial state s.
Begin
(1) While $t \leq t_{\max}$
(2) For $n = 1$: N
If rand$(0, 1) < \varepsilon$
User n randomly selects action a from action space A;
Else
User n selection action $a = \arg\max_{a \in A} Q(s, a; \theta)$
End if
End for
(3) Allocate resources according to user actions
(4) For $n = 1$: N
(5) Obtain a reward r and a new status s', and store the new sample (s, a, r, s') in the experience pool.
(6) Update status $s = s'$
(7) Random sampling (s_i, a_i, r_i, s_{i+1}) is conducted from the experience pool to form small batch samples.
(8) Perform gradient descent on $(y_i - Q(s_i, \alpha_i; \theta)^2$ with respect to parameter θ.
(9) Reset the target action value function $\widehat{Q} = Q$ every e steps.
(10) End for
(11) $t = t + 1$
(12) End While
(13) Determine the final user association matrix according to all action sets.
End

Algorithm 1: Pseudocode of offloading strategy optimization algorithm based on DQN.

Table 1: System simulation parameters.

System parameters	Value
Bandwidth between UAV and user/MHz	10
Bandwidth between UAV and AP/GHz	1
UAV computing power/MHz	1200
User task size/Mbits	[20, 30, 50]
User transmission power/W	3
Maximum user delay/s	[10, 15, 25]
Flight cycle time slot of UAV	32
Fixed altitude of UAV/m	60
Maximum available speed of UAV/m·s^{-1}	45
Gauss white noise/W	10^{-10}

model, and the computing energy efficiency of user equipment increases with the increase in maximum consumed energy. The reason is that as the energy of user equipment increases, the user equipment has more energy to perform local computations or offload computations. In addition, in the local computing mode, the computing energy efficiency does not change as the energy increases. It can be considered that when the maximum energy consumption value is 1J, the computing energy efficiency of user equipment has reached the maximum value. Subsequently, the user equipment does not need to consume more energy to improve the computing energy efficiency.

4.4. Relationship between Total Delay and the Number of IoT Equipment. As the number of equipment in the system changes, the total latency of the IoT system will also accordingly change. In order to demonstrate the performance of the proposed strategy, it is compared with reference [13], reference [14], and reference [22], and the results are shown in Figure 5.

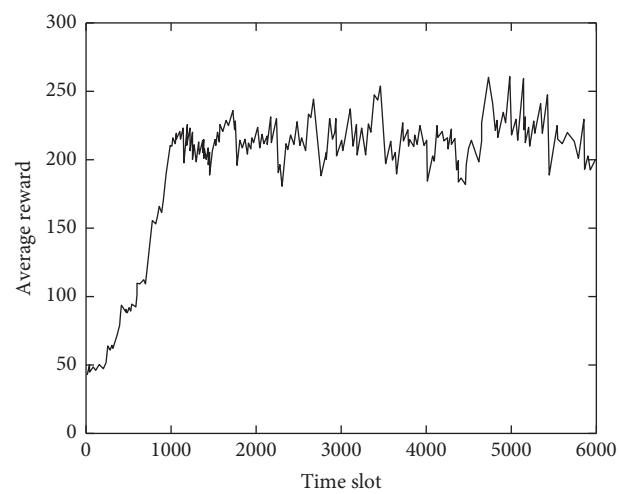


Figure 2: The average reward in the training process of the proposed algorithm.

As can be seen from Figure 5, as the number of equipment increases, the total system delay also increases as expected. The main reason is that when the number of equipment in the system increases, UAV flight delay, UAV edge computing delay, equipment upload delay, and equipment local computing delay corresponding to the newly added equipment will be added to the system. will be added to the total delay of system. Furthermore, the proposed strategy significantly outperforms other comparative strategies in reducing the total system latency. When the

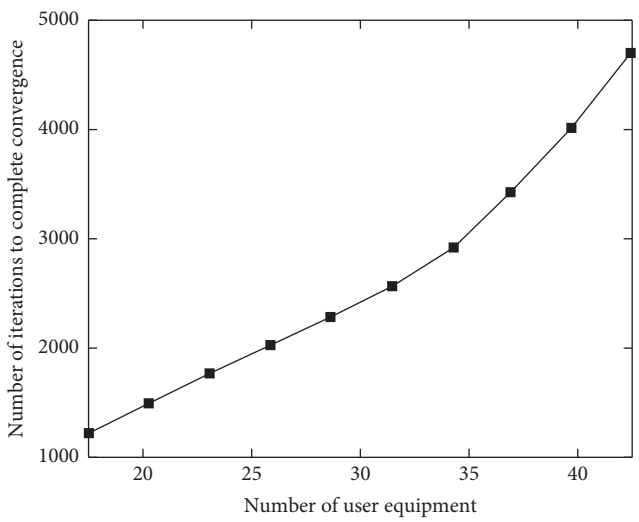


Figure 3: Relationship between the number of convergence iterations and the number of system equipment.

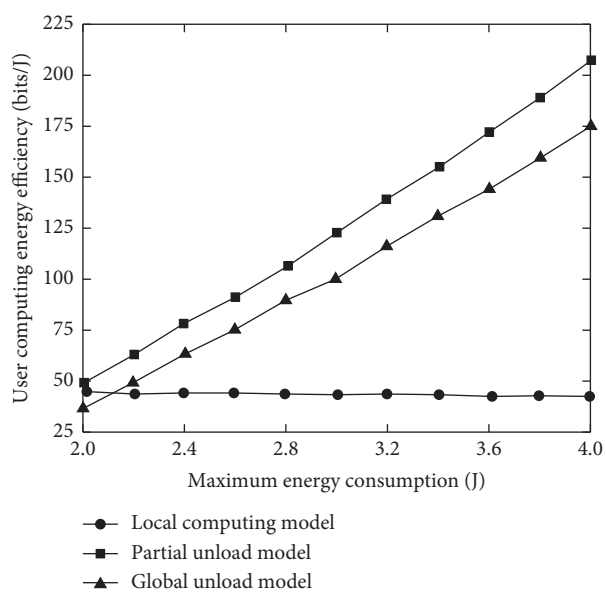


Figure 4: Relationship between calculated energy efficiency and maximum energy consumption of different loading models.

number of user equipment is 40, the total delay is about 33s, which is 35.29%, 31.25%, and 15.38% lower than reference [13], reference [14], and reference [22], respectively. Since the proposed strategy utilizes mobile UAVs for computational offloading and optimizes resource allocation with DQN, the delay can be minimized. Reference [13] proposed a vehicle-assisted computing offloading architecture for UAVs. Since the optimization algorithm has a weak ability to seek optimization, it takes a long time, more than 50s. Reference [14] proposed a UAV-assisted agent-enabled computing task offloading framework to help users, UAVs, and edge clouds perform computing task offloading. However, this algorithm requires a large amount of computation and takes a long time, resulting in increased delay.

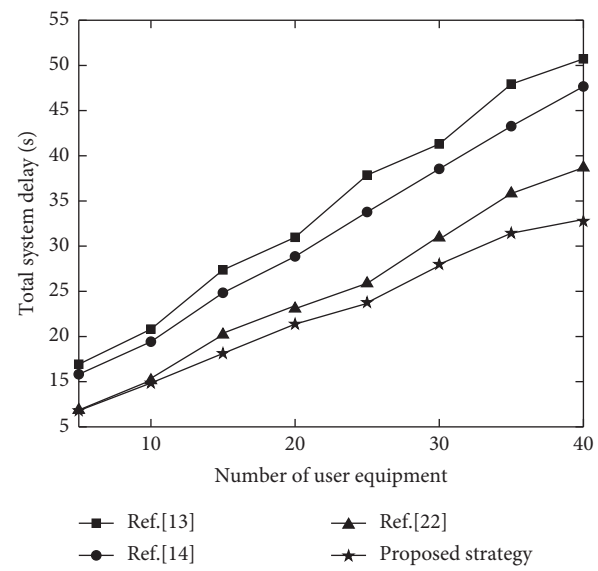


Figure 5: Relationship between time delay and the number of user equipment under different strategies.

Reference [22] proposed a deep reinforcement learning-based MEC UAV-assisted computing offloading scheme. When the number of user equipment is small, the total delay is close to the proposed strategy. However, when the number of user equipment is large, the processing timeliness is not strong, and the delay is relatively large.

5. Conclusions

In order to give full play to the advantages of a UAV-assisted MEC network, it is necessary to further formulate a reasonable user offloading strategy and UAV flight trajectory. To this end, this study proposes a network resource allocation strategy based on UAV collaborative edge computing. Based on the system scenario of collaborative computing between UAVs and ground users, an optimization model is constructed to minimize the total system delay and use DQN to optimize, so as to obtain the best resource allocation scheme. The experimental results show the following:

(1) With the fast mobility of UAVs, users in the system can offload computing tasks in real time, which greatly reduces the transmission delay. Especially when the computing power of UAV reaches 2000 MHz, the total system delay is about 20s.

(2) The proposed strategy uses DQN to optimize the offloading strategy, which not only has a fast optimization speed but also has an efficient optimization result, and realizes a reasonable allocation of resources while reducing the delay. Especially when the amount of user data is large, the optimization effect is more obvious. When the number of user equipment is 40, the total delay is about 33s, which is more than 10% lower than the total delay of other comparison strategies.

As the network scale becomes larger, the network scenarios also become more complex. In order to be suitable for large-scale complex network scenarios and meet the real-time requirements of MEC systems, we will focus on researching and designing a simpler and more effective online scheduling strategy in the future. This is crucial for the large-scale application of MEC.

Data Availability

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

Conflicts of Interest

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

Acknowledgments

This work was supported by the Provincial Education Reform Project (no. 2020JGLX081) and the Department Level Education Reform Project (no. KCSZ-202022).

References

[1] W. Wu, F. Zhou, R. Q. Hu, and B. Wang, "Energy-efficient resource allocation for secure noma-enabled mobile edge computing networks," *IEEE Transactions on Communications*, vol. 68, no. 1, pp. 493–505, 2020.

[2] M. Sirajuddin, C. Rupa, C. Iwendi, and C. Biamba, "TBSMR: a trust-based secure multipath routing protocol for enhancing the qos of the mobile ad hoc network," *Security and Communication Networks*, vol. 2021, pp. 1–9, 2021.

[3] X. Chen, Z. Liu, Y. Chen, and Z. Li, "Mobile edge computing based task offloading and resource allocation in 5g ultra-dense networks," *IEEE Access*, vol. 7, no. 7, pp. 184172–184182, 2019.

[4] Z. Song, Y. Liu, and X. Sun, "Joint radio and computational resource allocation for noma-based mobile edge computing in heterogeneous networks," *IEEE Communications Letters*, vol. 22, no. 12, pp. 2559–2562, 2018.

[5] W. Fang, S. Ding, Y. Li, W. Zhou, and N. Xiong, "OKRA: optimal task and resource allocation for energy minimization in mobile edge computing systems," *Wireless Networks*, vol. 25, no. 5, pp. 2851–2867, 2019.

[6] Y. Gao, Y. Cui, X. Wang, and Z. Liu, "Optimal resource allocation for scalable mobile edge computing," *IEEE Communications Letters*, vol. 23, no. 7, pp. 1211–1214, 2019.

[7] R. Sridharan and S. Domnic, "Placement strategy for intercommunicating tasks of an elastic request in fog-cloud environment," *Scalable Computing: Practice and Experience*, vol. 20, no. 2, pp. 335–348, 2019.

[8] D. Zhu, C. Xiang, and S. Bing, "Biologically inspired self-organizing map applied to task assignment and path planning of an auv system," *IEEE Transactions on Cognitive and Developmental Systems*, vol. 10, no. 99, pp. 304–313, 2018.

[9] Z. Fang, J. Wang, and J. Du, "Stochastic optimization aided energy-efficient information collection in internet of underwater things networks," *IEEE Internet of Things Journal*, vol. 9, no. 3, pp. 1775–1789, 2021.

[10] J. Wang, C. Jiang, and Z. Wei, "Joint UAV hovering altitude and power control for space-air-ground iot networks," *IEEE Internet of Things Journal*, vol. 6, no. 2, pp. 1741–1753, 2018.

[11] B. Yu, X. Zhang, and I. You, "Collaborative cache allocation and transmission scheduling for multi-user in edge computing," *IEEE Access*, vol. 8, no. 4, pp. 163953–163961, 2020.

[12] C. Kai, H. Zhou, Y. Yi, and W. Huang, "Collaborative cloud-edge-end task offloading in mobile-edge computing networks with limited communication capability," *IEEE Transactions on Cognitive Communications and Networking*, vol. 7, no. 2, pp. 624–634, 2021.

[13] M. Dai, Z. Su, Q. Xu, and N. Zhang, "Vehicle assisted computing offloading for unmanned aerial vehicles in smart city," *IEEE Transactions on Intelligent Transportation Systems*, vol. 22, no. 3, pp. 1932–1944, 2021.

[14] W. A. Rui, C. A. Yong, and B. An, "Agent-enabled task offloading in UAV-aided mobile edge computing," *Computer Communications*, vol. 149, no. 7, pp. 324–331, 2020.

[15] Z. Qin, X. Qiu, J. Ye, and L. Wang, "User-edge collaborative resource allocation and offloading strategy in edge computing," *Wireless Communications and Mobile Computing*, vol. 2020, no. 11, pp. 1–12, 2020.

[16] B. Liu and D. Xu, "Auction-based resource allocation for mobile edge computing networks," *IEICE - Transactions on Fundamentals of Electronics, Communications and Computer Sciences*, vol. 103, pp. 718–722, 2020.

[17] J. He, D. Zhang, Y. Zhou, and Y. Zhang, "A truthful online mechanism for collaborative computation offloading in mobile edge computing," *IEEE Transactions on Industrial Informatics*, vol. 16, no. 7, pp. 4832–4841, 2020.

[18] Y. Liao, L. Shou, Q. Yu, Q. Ai, and Q. Liu, "Joint offloading decision and resource allocation for mobile edge computing enabled networks," *Computer Communications*, vol. 154, no. 2, pp. 361–369, 2020.

[19] B. Wu, K. Xu, Q. Li, S. Ren, Z. Liu, and Z. Zhang, "Toward blockchain-powered trusted collaborative services for edge-centric networks," *IEEE Network*, vol. 34, no. 2, pp. 30–36, 2020.

[20] W.-C. Chien, H.-Y. Weng, and C.-F. Lai, "Q-learning based collaborative cache allocation in mobile edge computing," *Future Generation Computer Systems*, vol. 102, no. 6, pp. 603–610, 2020.

[21] E. Li, L. Zeng, Z. Zhou, and X. Chen, "Edge ai: on-demand accelerating deep neural network inference via edge computing," *IEEE Transactions on Wireless Communications*, vol. 19, no. 1, pp. 447–457, 2020.

[22] H. Wang, H. Ke, and W. Sun, "Unmanned-aerial-vehicle-assisted computation offloading for mobile edge computing based on deep reinforcement learning," *IEEE Access*, vol. 8, no. 7, pp. 180784–180798, 2020.

[23] C. Gza, Z. Hao, and B. Yla, "5G network-oriented hierarchical distributed cloud computing system resource optimization scheduling and allocation," *Computer Communications*, vol. 164, no. 4, pp. 88–99, 2020.

[24] M. Mukherjee, S. Kumar, C. X. Mavromoustakis et al., "Latency-driven parallel task data offloading in fog computing networks for industrial applications," *IEEE Transactions on Industrial Informatics*, vol. 16, no. 9, pp. 6050–6058, 2020.

[25] A. U. Rehman, Z. Ahmad, A. I. Jehangiri et al., "Dynamic energy efficient resource allocation strategy for load balancing in fog environment," *IEEE Access*, vol. 8, no. 2, pp. 199829–199839, 2020.

[26] C. Jiang, Y. Li, and J. Su, "Research on new edge computing network architecture and task offloading strategy for Internet of Things," *Wireless Networks*, vol. 6, no. 2, pp. 1–13, 2021.

[27] L. Ruan, Z. Liu, X. Qiu, Z. Wang, S. Guo, and F. Qi, "Resource allocation and distributed uplink offloading mechanism in fog

environment," *Journal of Communications and Networks*, vol. 20, no. 3, pp. 247–256, 2018.

[28] B. Al-Manthari, N. Nasser, and H. Hassanein, "Congestion pricing in wireless cellular networks," *IEEE Communications Surveys & Tutorials*, vol. 13, no. 3, pp. 358–371, 2011.

[29] S. Min, "Energy-efficient multiuser partial computation offloading with collaboration of terminals, radio access network, and edge server," *IEEE Transactions on Communications*, vol. 68, no. 3, pp. 1524–1537, 2019.

[30] T.-M. Pham and T.-T.-L. Nguyen, "Optimization of resource management for nfv-enabled iot systems in edge cloud computing," *IEEE Access*, vol. 8, no. 7, pp. 178217–178229, 2020.

Hindawi
Journal of Robotics
Volume 2022, Article ID 2365006, 13 pages
https://doi.org/10.1155/2022/2365006

Research Article

Unmanned Aerial Vehicle Surveying and Mapping Trajectory Scheduling and Autonomous Control for Landslide Monitoring

Shifang Liao,[1] Manzhu Ye,[2] Rongcai Yuan,[3] and Wanzhi Ma[4]

[1]*School of Resource & Environment and Historical Culture, Xianyang Normal University, Xianyang, Shaanxi 712000, China*
[2]*School of Surveying & Testing, Shaanxi Railway Institute, Weinan, Shaanxi 714000, China*
[3]*Xi'an Zhongke Xingyun Space Information Research Institute Co., Ltd.,, Xi'an, Shannxi 710000, China*
[4]*School of Educational Science, Ningxia Normal University, Guyuan, Ningxia 756000, China*

Correspondence should be addressed to Wanzhi Ma; mawanzhi79@nxnu.edu.cn

Received 29 December 2021; Revised 8 February 2022; Accepted 9 February 2022; Published 24 March 2022

Academic Editor: Shan Zhong

Real-time and efficient monitoring of geological disasters has received extensive attention in the application of UAV surveying and mapping control technology. The application of traditional landslide monitoring methods lacks the accuracy of control algorithms, which has become a hot issue currently facing. Based on the landslide surface subsidence monitoring method, this article designs the UAV trajectory scheduling subsidence monitoring software, which can monitor the UAV's flight status and navigation information, and draw the flight trajectory in real time. At the same time, the model solves the problem of storage and management of landslide inspection results by the landslide inspection management system, and realizes the functions of entering and querying landslide information, viewing inspection results, landslide safety judgment, generating reports, and autonomous control. The simulation results show that the global accuracy reaches 0.975, and the algorithm recognition degree reaches 99.8%, which promotes the reliability of the landslide monitoring data for the identification of the surveying and mapping trajectory, and provides a decision-making basis for landslide treatment.

1. Introduction

Due to the subsidence and deformation of the ground surface caused by landslide activities, the methods of subsidence monitoring have undergone a long-term process, ranging from traditional leveling traverse measurement and GPS monitoring to the more commonly used three-dimensional laser scanning monitoring and D-InSAR monitoring. These monitoring methods range from full-field to few-field measurements, and the observation efficiency is greatly improved, but their monitoring range, accuracy, and efficiency still cannot meet the needs of comprehensive monitoring of regional subsidence [1]. The emergence of UAV low-altitude photogrammetry at this stage has greatly improved the efficiency of monitoring with its high mobility and real-time performance. Its high-precision digital results are applied to various monitoring fields, including monitoring of mountain landslides, mudslides, and collapses. The terrain conditions of highway construction in mountainous areas are complex and changeable. Due to the consideration of many aspects, highway construction cannot avoid the use of high filling and deep excavation methods, which will form a high landslide structure, because such high landslides are all reconstructed from the original landform. Breaking the original balance relationship and balance conditions of the landform, the stress field will change accordingly, and the open-air slope will have cracks and landslide settlement and sliding due to the local lithology and structure, surface water, ground runoff, etc. [2–5].

At present, with the continuous development and progress of drone technology, autonomous drone navigation technology has gradually been applied to all walks of life in society, providing it with high-performance technology services such as safety, convenience, intelligence, and environmental protection, making people's life and social production have become more and more convenient and efficient [6–8]. The research and development of UAV autonomous navigation technology will continue to promote the innovation of social science and technology, which is of great significance to

the future development. With the vigorous development of national highway traffic, highway mileage has increased year by year. In the road operation stage, road maintenance and management are particularly important, and landslide inspection is a basic but important task in road maintenance and management. Daily regular inspections are conducive to real-time grasp of the stability and safety of landslides. Timely early warning and treatment of diseases are essential. At present, landslide inspections use manual inspections, generally using visual inspection, tapping, and touching. The terrain conditions of landslides are complex and changeable. When inspectors face high landslides and mountain landslides, manual inspections are very dangerous or even to reach certain areas of the landslide, resulting in inspection results that cannot fully reflect the true situation of the landslide. Therefore, conducting drone inspections for landslide disease has important practical significance for disaster reduction and prevention and traffic safety. In traditional manual inspection, visual inspection is the main method. Therefore, the operation mode of drones is used to replace or assist traditional inspection personnel's short-distance visual inspection, which makes landslide safety inspection possible [9–11].

UAV low-altitude photogrammetry as a fast, convenient, and safe image acquisition technology has been studied in the field of geological disaster investigation and other fields. The purpose of this article is to study the technical methods of using drones for landslide inspection to replace or assist manual road inspections, so as to ensure the safety of inspectors. According to the relevant requirements of the landslide inspection, starting from the actual inspection work, two technical methods of overall inspection and fine inspection of landslides based on unmanned low-altitude photogrammetry are proposed. In order to realize the early deformation recognition of the landslide in the area, this article uses drone photogrammetry technology to take multiple phases of the landslide to study its accuracy and deformation recognition methods. First, the image of the survey area is obtained by flying the drone, and then, the image is processed to obtain the point cloud data. Through continuous improvement of the number and location of ground control points, the accuracy of UAV aerial photography technology has been improved, and the accuracy has been increased from the decimeter level to the centimeter level. Before entering the test in the subsidence area, the UAV aerial survey plan was designed. Combining two more classic sparse and dense deployment schemes, the accuracy of the UAV digital results under the two schemes is verified to obtain a set of reasonable image control point schemes. The two sets of data, respectively cite the post difference technology to verify the digital results and verify the accuracy and superiority of the post difference technology. Under the same scheme, the DEM accuracy of tilt and vertical photography was compared, and the best tilt aerial photography plan was obtained.

2. Related Work

Compared with traditional methods of monitoring the regional surface, the rapid development of drone aerial photogrammetry technology in recent years has become more and more widely used. The drone technology is highly maneuverable, flexible, and lightweight in the process of acquiring images. The acquired images have high resolution and play an important role in surveying and mapping geographic information, earthquake mitigation and disaster relief, agricultural production estimation, water conservancy and hydropower, and transportation and construction. The most widely used technology in UAV photogrammetry technology is tilt photogrammetry technology. This technology has become one of the current research hotspots. With its high precision, high mobility, and high efficiency, it is suitable for operation in a variety of complex environments. [12–15].

At present, there are three methods for differential interferometry, namely, the two-track method of two InSAR images, the three-track method of three InSAR images, and the four-track method of four InSAR images. The deformation result measured by InSAR technology is the estimated value of the spatial average change in a certain area. This method makes up for the shortcomings of traditional monitoring methods for the sparse measurement points. It has the characteristics and advantages of large area, real time, and high precision, which effectively complements traditional monitoring. Zanol et al. [16] used the X-band and L-band of D-InSAR to monitor the subsidence of the central area of Utah. The results show that the L-band can accurately monitor the subsidence value of this area, and it is less affected by slope changes. Although it is difficult to achieve a high-precision level for subsidence monitoring using X-band, it can also monitor the survey area with a reasonable and accurate subsidence range. Figueiredo et al. [17] used high-resolution images to extract and monitor tunnel cracks to provide a solid guarantee for the safe operation of the subway. However, the uncertainty in the selection of the segmentation threshold directly affects the fracture extraction results. Julge et al. [18] applied it to the UAV image classification experiment and achieved 83% global accuracy and 77.36% Kappa accuracy. The classification results are highly consistent with the reference data type; in order to improve the extraction accuracy of ground fissures, the hit-to-hit transform algorithm connects the broken ground fissures, and at the same time, the small pattern removal algorithm combining shape and area proposed in this article is used to realize the removal of ground fissures and combined with ground fissure extraction results to conduct experiments to verify the effect of ground fissure fine treatment.

Zhang et al. [19] studied a fault-tolerant path planning algorithm, using multisensor fusion to study fault-tolerant path planning and using wireless communication systems for positioning. Basir et al. [20] studied the path planning algorithm based on a firefly under an uncertain environment, using the attraction law between fireflies to reduce the number of iterations and helping the robot to search for the optimal path in the constantly changing dynamic environment according to the local static environment. At the same time, a linear computational complexity graph search algorithm is proposed, which uses the priority queue, graph search, and the combination of a cattle to study a two-way sublist path planning algorithm that is faster than the

algorithm under global information. Domestic researchers have also made a lot of contributions in the field of path planning. Scholars have studied the PRM path planning algorithm based on distance transformation and constructed a narrow channel undirected graph based on image processing and random icon technology to study the density of obstacles [21–23]. The researcher proposed a multithreaded SA∗ path planning algorithm, adding multithreaded parallel computing to the A-Xing algorithm and using heuristic search to improve the SA∗ algorithm to improve the speed and accuracy of path calculation. Some scholars try to use the genetic iterative algorithm to extract road cracks, but the selection of the best parameters in the algorithm is still a problem worthy of in-depth study. In order to extract the cracks of concrete bricks, multiple threshold segmentation algorithms are used for experiments, although the final result is optimized to the subpixel level, but the threshold segmentation algorithm cannot distinguish targets with similar gray values, resulting in a large amount of background noise; in order to realize the automatic detection of cracks in concrete pipes, the researchers proposed a method to segment pipe images based on morphological methods [24–26].

3. Analysis of Landslide Monitoring System

3.1. Structural Diseases of Landslides. In the landslide structure, the broken rock mass of the slope top and the slope surface is the broken rock mass formed by the fracture and disintegration of the rock mass located on the top of the slope and the slope surface. At the initial stage of the formation of a landslide, the redistribution of stress resulted in the concentration of tensile stress on the top of the slope, pressure cracking on the slope, the development of tension and cracks in the rock mass, and the fragmentation of the rock mass. Collapse is a phenomenon in which the rock and soil on a high and steep landslide suddenly completely separates from the parent body under the action of gravity and external force, and then rolls, jumps, dumps or falls, and accumulates to the foot of the slope or the road:

$$\lim_{x\longrightarrow\infty} u_q(t,x) - \lim_{x\longrightarrow\infty} u_q(wxt+q)U = 0. \tag{1}$$

When the quaternion is updated, the Euler angle can be used to obtain the attitude angle in the dispatch, but the yaw angle is not used as the final yaw angle of the attitude, and it needs to go through the first-order complementary calculation. In the calculation, the complementary gyroscope calculated output specific gravity is increased, while the magnetometer output specific gravity is decreased, and the yaw angle is obtained by the fusion calculation:

$$\lim_{x\longrightarrow\infty} u_d(x,t) = \lim_{x\longrightarrow\infty} u_d(wxt-d)U. \tag{2}$$

The protective structure is a force-bearing structure set on the landslide slope surface to support the rock and soil pressure. The protective structure is mainly affected by water erosion and rock and soil pressure. The coagulation structure has a large amount of cement. The internal cracks in the concrete are the primary breakthrough points for water erosion. In these locations, the concrete is easy to peel off and the steel bars are exposed, causing structural damage and loss of function:

$$\lim_{x\longrightarrow\infty} u_d(\cos x, \sin x) = \lim_{x\longrightarrow\infty} \tan(xt-d)U. \tag{3}$$

The protective structure is mainly subjected to rock and soil pressure. When the local structural strength is lower than the structural stress, the structure will bend and fail. The magnetometer is susceptible to external interference when the drone is autonomously navigating. Therefore, the magnetometer data are not used during complementary fusion, but the value is the magnetometer output when the system initializes the quaternion:

$$\begin{bmatrix} u_d \\ u_q \end{bmatrix} \times \begin{bmatrix} u_d & u_q \end{bmatrix} = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \times \begin{bmatrix} -u_d \\ u_q \end{bmatrix}. \tag{4}$$

In the safety judgment module, according to the drone landslide inspection results, the four aspects of landslide weathering degree, drainage system, protective structure, and other landslide phenomena are evaluated, and the corresponding options are selected in the inspection management system, such as weathering less intense, less intense, more intense, and intense. The inspectors make choices based on the actual inspection results. The system integrates these four options to determine the safety of the road landslide. The judgment results are as follows: stable, basically stable, understable, and unstable. After the judgment, the corresponding recommended treatment measures are displayed.

3.2. Classification of Landslide Safety Index. The landslide safety index information needs to be properly preserved. In the information management module, the information has been stored in the database. Users can export landslide information and inspection results to files by generating reports, which is convenient for users to save and can effectively ensure the safety of landslide data. In the safety prewarning module, users can also send the results of landslide inspection and treatment suggestions to managers' mailboxes to realize real-time communication and feedback functions:

$$F(s,t,\theta) = \begin{bmatrix} \sin\theta' & \cot\theta \\ \cot\theta' & \sin\theta \end{bmatrix} \times \begin{bmatrix} 1-u_s \\ 1-u_t \end{bmatrix}. \tag{5}$$

Finding the extreme point (gradient position) in the first-order derivative function of the gray value is a troublesome task, while finding the position where the second-order derivative function value is 0 is relatively easy. The second-order edge detection algorithm is to detect the pixel position in the image where the second-order derivative function value of the gray value is 0, which is regarded as an edge. Similar to the first-order derivative function, the second-order derivative function can be replaced by the second-order difference form of pixel gray value.

Under the premise of ensuring accuracy, the use of drone camera measurement technology can be used to monitor the deformation of the landslide group. For areas with a very

large number of landslides in Table 1, the deformation area has been monitored, and the identification of potential landslide areas has also been realized:

$$\prod f(x,t) \times U_i(s) - \prod f(x-1,t-1) \times \frac{s(s+w)}{w^2 - s^2} U = 0. \tag{6}$$

The above system errors have destroyed the central projection relationship between the objects, and the influence of the curvature of the Earth belongs to the difference caused by different projection transformations. The geoid is an ellipsoid. The ground coordinate system used in map projection uses a plane as a horizontal plane. When the aerial photography range is large, this difference will affect the accuracy of the aerial three encryption results, so it needs to be corrected. In the length of the buffer number of NS, the historical value of the appropriate interval is taken as the estimated value of the current moment according to the delay time of the auxiliary sensor:

$$\sum s(s+w) \times \sum w^2 - s^2 - \sum f(s,t) \times \frac{s(s+w)}{w^2 - s^2} = 0. \tag{7}$$

If it is under high-speed operation, the delay correction needs to use the least square fitting; that is, the delay correction buffer (effective broadband autonomous control algorithm value) is polynomial fitting, and the curve is fitted according to the buffer area. Then, use the median value average recursive filtering to calculate the optimal estimated value at the next moment, and refresh the buffer area.

In the same way, through the working principle of each module, the data manual can calculate that the maximum sampling frequency of ultrasound in this study is 75 ms, the maximum sampling frequency of GPS M8N is 100 ms, the maximum sampling frequency of the barometer is 20 ms, and the maximum sampling frequency of the optical flowmeter is 8 ms. The previous surveying and mapping scheduling, serial port receiving buffer, digital filtering, etc. will also have a period of time.

3.3. UAV Fine Inspection. The development of UAV fine remote sensing and photogrammetry science has enabled humans to use low-altitude, high-altitude, and even outer space sensors to obtain various image data reflecting the characteristics of the surface. By extracting the physical characteristics and information of various target objects, the spatial shape, location, nature, change, and the relationship with the environment are studied. Feature points are pixels that describe the geographic location of the same information in two or more images and are used for image geometric transformation, image splicing, and three-dimensional model generation:

$$\lim_{\theta \longrightarrow \infty} \Delta\theta - \left|\theta - 2\theta'\right| - \left|wt - 2\theta'\right| = 0. \tag{8}$$

The extraction of feature points is an important step in the processing of oblique photogrammetry. The relationship in Figure 1 between multiview images is constructed by conjugate points. At present, the photogrammetry community has proposed many methods for automatic extraction of connection points. Commonly used automatic connection point extraction algorithms include scale-invariant feature transform (SIFT) and fast robust feature (SURF), Harris, MSER (the most stable extreme value region), and FAST (features from accelerated segment test) algorithms.

Table 1: Monitoring instructions for deformed areas.

Monitoring test	Partial deformed areas		
	Phase A	Phase B	Phase C
1	2.30	2.44	2.58
2	1.77	1.87	1.97
3	1.24	1.30	1.36
4	0.71	0.73	0.75
5	0.18	0.16	0.14
6	−0.35	−0.41	−0.47
7	−0.88	−0.98	−1.08
8	−1.41	−1.55	−1.69
9	−1.94	−2.12	−2.30

In the actual field operation of tilt photogrammetry, the image will inevitably be affected by uncertain external factors, such as radiation distortion and geometric distortion in the image. For external influence factors, these operators can accurately extract feature points according to their respective characteristics and appropriate measures to extract feature points, and make the feature points have good uniqueness, antirotation, antiscaling, and antilight variation. For the matched connection points, the quality of these connection points has a greater impact on the later feature matching, which is directly related to its accuracy and accuracy:

$$\begin{cases} \lim\limits_{\theta \longrightarrow \infty} \sin\theta \times \cos\theta = \left|\theta - 2x'\right|, \\ \lim\limits_{\theta \longrightarrow \infty} \cos\theta \times \cos\theta = \left|\theta + 2x'\right|. \end{cases} \tag{9}$$

The monitoring personnel set the route and UAV parameters in the ground control system, and the aerial part of the UAV then performs the flight mission according to the instructions of the ground control system. UAV flight data are transmitted to the ground control system in real time using wireless transmission channels, and ground monitoring personnel can change the flight plan or let the UAV continue to execute instructions based on the received data. When you take supplementary shots of poorly-photographed areas, emergency situations, and landing, you can switch the flight status from automatic to manual control:

$$\left\langle \begin{matrix} u_\alpha(t) - u_\alpha(t-1) \\ u_\alpha(t) + u_\alpha(t-1) \end{matrix} \middle| \begin{matrix} \sin(wt) \\ \sin(w(t-1)) \end{matrix} \right\rangle = 1. \tag{10}$$

On the basis of network socket communication, the TCP or UDP protocol needs to be added. This study takes network communication based on the TCP protocol as an example. This protocol mainly uses the TcpClient and TcpListener classes built on socket. The attributes of the TcpClient class mainly include Available, ReceiveBufferSize, Client, Connected, LingerState, NoDelay, SendTimeout, etc. The TcpListener class is mainly used to monitor and receive incoming connection requests. It mainly includes five stages:

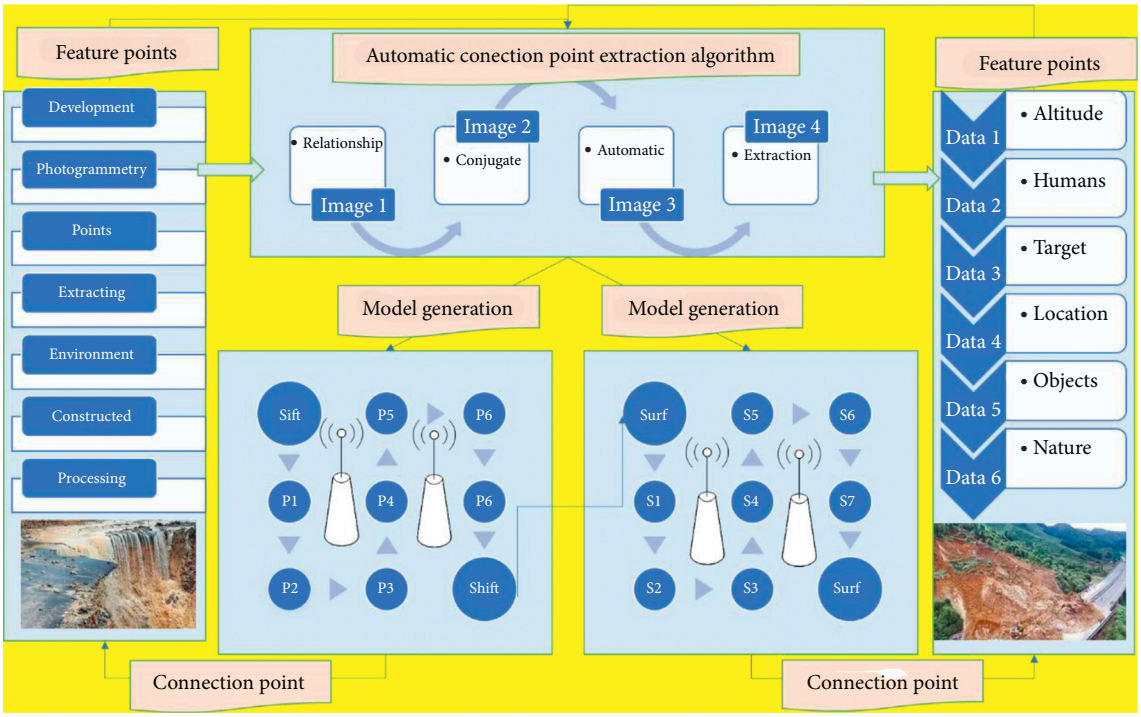


Figure 1: UAV fine inspection topology.

creating TCP instances, monitoring connections, receiving connection requests, sending and receiving data, and stopping services.

3.4. Image Processing Technology. The low-altitude UAV photogrammetry technology is mainly used for small surveys or emergency survey missions. It is characterized by unmanned aircraft as a flying platform, various types of sensors are used to obtain ground information, and computer graphics are used. Image technology processes the acquired images and provides basic data for various types of photogrammetry applications.

The data processing part includes a data preprocessing system and a data postprocessing system. The data preprocessing system includes photogrammetric data download, flight quality and data quality inspection, camera calibration, and distortion correction; data postprocessing includes air three processing, drawing production, interpretation, and result evaluation:

$$\int \left[u_{\alpha}(t) - \sin(wt) + \sin(w(t+1))\right] \mathrm{d}t = 0. \tag{11}$$

The UAV low-altitude photogrammetry system in Table 2 consists of two parts: data acquisition and data processing. Data acquisition refers to the route planning function of ground personnel through the ground monitoring system or the remote control card control system of the UAV remote control. We realize the take-off and landing of the UAV, and achieve the precise control of the drone dagger platform. Ground fissures are a form of existence of fissures on the ground, and their spatial form is characterized by dark tones and strip-like linear features.

The edge detection algorithm in computer vision is one of the earliest methods used for linear target extraction, which is mainly divided into the first-order edge detection algorithm and the second-order edge detection algorithm. The area where the local gray value changes obviously in the image can be regarded as the edge, and the change in the local gray value can be expressed by the first-order derivative function. If the value of the first derivative function is larger, the probability of becoming an edge is higher.

From a qualitative point of view, the deformation shown by the monitoring data and the orthographic image is basically the same as the difference result; from a quantitative point of view, there is a certain difference between the amount of deformation displayed by the difference model and the actual amount of deformation:

$$\left\langle \begin{array}{l} \sin(wt) + \sin[w(t+1)] = 0 \\ \cos(wt) - \cos[w(t+1)] = 0 \end{array} \right\rangle. \tag{12}$$

The amount of data in the same area has greatly increased, and the increase in the number of lenses has increased the amount of data, which brings a burden to subsequent image processing, and puts forward higher requirements on computer performance. The clustered multi-CPU and multi-GPU architecture is now used to handle tilt massive data obtained from photography. When the flight control board is rotated, the 3D modeling icon of the anonymous host computer will also rotate and the data will also change. It can be judged whether the IMU module can communicate normally according to the consistency of the rotation direction and the beating amplitude of the data change:

Table 2: UAV low-altitude photometric system.

Features step algorithm		Connection point 1	Connection point 2	Connection point 3	Connection point 4
Algorithm accuracy	SIFT 1	2.37	0.47	0.57	2.67
	SIFT 2	1.60	0.66	0.72	1.78
	SURF 1	0.83	0.85	0.87	0.89
	SURF 2	0.06	0.04	0.02	0.00
	MSER 1	−0.71	−0.77	−0.83	−0.89
	MSER 2	−1.48	−1.58	−1.68	−1.78

$$\text{if}\,(\sin t > \cos t), \text{c.t.out}\; f\left(x, t|x^2 + t^2 = 0\right) = \frac{\partial \sin(wt)}{\partial w \partial t}. \tag{13}$$

The low-altitude indoor autonomous control module, which is a combination of ultrasonic and optical flow meter, is connected to the reserved UART port of the flight control system. You can use the serial debugging assistant to check whether the data output is normal or the onboard OLED screen on the flight control system to observe the data.

4. UAV Surveying and Mapping Trajectory Scheduling and Autonomous Control Model Construction for Landslide Monitoring

4.1. UAV Mapping Trajectory Scheduling. The cameras carried by UAV photogrammetry are all nonmeasurement cameras, and the images acquired by them have a certain degree of distortion due to the time difference. The image must be removed before data processing. Through the adjustment of the position of the principal point of the image and the correction of the distortion parameters, the processing of the distortion effect on the image is realized.

In order to improve the matching speed of the image points with the same name and increase the density and accuracy of the matching points, it is necessary to create a pyramid for the image and perform image matching based on the image classification matching strategy that is gradually refined by the image pyramid, so that good reliability and high accuracy can be obtained to extract similar feature points from two adjacent images with a certain degree of overlap in Figure 2, and then use gray-scale or feature-based matching algorithms to match points with the same name, and obtain a point set after relative orientation.

The data transmission system consists of a modem and a radio station. The reference station modem encodes and modulates the relevant data, and then, the radio station transmits it. The radio station on the rover station receives and demodulates the relevant signal through the radio station. Finally, the three-dimensional coordinates of the point are calculated and processed by the real-time dynamic measurement software system.

In order to make up for the lack of edge extraction accuracy caused by the randomness of the threshold selection in the first-order edge detection algorithm, the second-order edge detection algorithm was proposed to search for the maximum point of the local gradient value of the first-order derivative function (i.e., the value of the first-order derivative function):

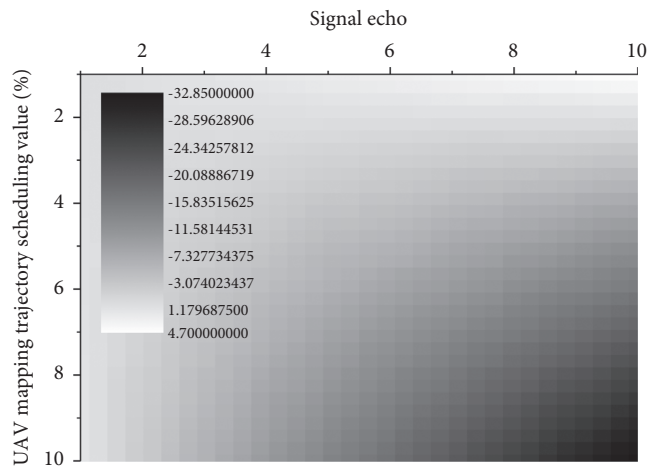


Figure 2: UAV surveying and mapping trajectory scheduling distribution.

$$\frac{\partial \sin(wt)}{\partial (wt)} + \frac{\partial \sin(w(t+1))}{\partial \cos(wt)} = 1, \quad \text{if}\,\{\cos(wt) - 1 < 0\}. \tag{14}$$

However, the analog image taken by a traditional film camera (i.e., an image with continuous image point and gray distribution) or a digital camera equipped with a charge-coupled device (CCD) is limited by the storage of digital devices such as computers. The captured digital image (i.e., the image in which the image point and gray-scale distribution are expressed in the form of discrete data), to be processed in the computer, must be stored in the form of discrete data.

4.2. Composition of Landslide Inspection Management System. The landslide inspection management system is divided into two parts: air and ground. The air part contains electronic equipment for sending data, external antennas, and ports for receiving data, while the ground part contains electronic equipment for sending data and various antenna interfaces. These are used for data transmission between air cruise control equipment and ground observation stations.

This part of the system contains equipment that enables the aircraft to take off and land. The launch system provides guarantee for the accelerated take-off of the UAV, and the recovery part provides guarantee for the safe landing of the UAV. Ground support equipment for air flight includes support equipment during the transportation phase and equipment during filming.

The function of the transportation safety assurance equipment is to protect the shooting equipment and mission equipment from damage, while the shooting safety assurance equipment refers to the equipment that allows the drone to take pictures safely in the field. In actual research, the raw data of the accelerometer and gyroscope in Figure 3 can be sent to MATLAB for FFT change, and the spectrum graph near the throttle can be analyzed, so that the frequency range of noise can be observed, and the MATLAB filter design toolbox can be used to study the filter parameter design to determine various parameters of Butterworth low-pass filtering. While determining the order of the filter, consider that although the higher the order, the faster the stopband attenuation, the system phase delay will also increase:

$$[\![\frac{-b \pm \sqrt{b^2 - 4ac}}{2a} \cos t]\!] \underset{|\theta - 2\theta'|}{\rightleftharpoons} \tan(ab) \cos t. \tag{15}$$

After comparing the theory of tilt and vertical photography, and based on the high-precision aerial photography plan of vertical photography, that is, postdifferential technology and reasonable image control point layout, tilt photography uses its multiview data acquisition and multiview joint adjustment to compare the postdifferential.

In the application of technology and image control point program, the elevation accuracy of tilt photography data results was compared with that of vertical photography. It was found that the elevation accuracy of tilt photography was improved to a certain extent compared with that of vertical photography, which provided high-precision program support for subsidence monitoring of tilt photography.

Under the condition of the same degree of overlap, the basis can only be increased by increasing the heading CCD width or shortening the focal length. However, the shorter the photographic baseline B, the more automatic the matching points, the easier the image matching, and the higher the relative orientation accuracy. SPP extracts fixed-dimensional features from the input object image. First, the maximum pooling divides the image into 4 feature maps, each of which extracts features of the same dimension and then pools the image into 16 feature maps with consistent dimensions.

After receiving the survey task, it is necessary to first analyze the survey task and objectives in detail to determine the flight platform and photogrammetric sensor to be used, and then choose to carry out drone aerial photogrammetry work based on the geological and meteorological conditions of the survey area in Table 3. The best season and time are finally used as the basis for field reconnaissance and comprehensive analysis based on the previous geological work degree and other data.

Optimizing the flight plan can effectively reduce the number of invalid flights and time. Based on a comprehensive understanding of the location and area of the target area, full consideration should be given to flight efficiency, and the minimum number of routes to the target should be used under the condition of ensuring full coverage of the target area.

4.3. Autonomous Control of Safety Early Warning. Safety early warning needs to introduce the tilt aerial photography program into the subsidence area, obtain the two phases of ground drone tilt data corresponding to underground mining activities, and use the level to measure the two phases of ground observation station data; carry out the two phases of drone tilt data.

The empirical parameters of the mine are used to predict the sinking basins, and the sinking basins are predicted to form. The sinking basins obtained by tilt photography are quantitatively analyzed using the measured observation line data, and Table 4 is used to predict the sinking basins. Qualitative analysis was performed on the acquired subsidence basin.

Compared with the subsidence difference between the two phases of ground observations, the two subsidence values are basically controlled within 150 mm, and the median error is 102 mm. Although there is a certain gap in the monitoring of the 10 mm terrain subsidence with slight changes, the mining parameters of this type, the aerial survey monitoring method has high reliability.

In order to facilitate reading and identification, it is necessary to renumber the images acquired in all directions and modify the image names according to certain rules. Since the original image is being shot, there are effects such as uneven illumination and changes in the sun's altitude angle, and the oblique image has forward and back light conditions. A uniform color and uniform light processing can be performed during preprocessing to ensure the natural color transition of the image:

$$\begin{aligned} &\oiint |wt - 2\theta'||wt + 2\theta'| \mathrm{d}w\mathrm{d}t \\ &\quad - \oiint |wt - \theta'||wt + \theta'| \mathrm{d}w\mathrm{d}t = 0. \end{aligned} \tag{16}$$

After superimposing the predicted subsidence basin with the aerial photography subsidence basin, we can intuitively see that the areas with larger subsidence values in the aerial photography subsidence basin tend to be on the west side of the aerial photography subsidence basin and are basically the same as the subsidence center of the predicted subsidence basin. The operations such as cropping and scaling, which are easy to lose information on the original image, are removed and replaced by a spatial pyramid pooling layer on the convolutional features.

The coincidence reflects the accuracy of aerial photography of the sinking trend of the subsidence basin; in addition, it can be seen that the subsidence range obtained by drone aerial photography is more comprehensive and broad to reflect the area affected by the subsidence of the mining area, breaking the traditional monitoring methods that can only be from one line. Observational bottlenecks reflect the comprehensiveness of using aerial photography to monitor subsidence trends:

$$\sqrt{\lim_{\theta \longrightarrow \infty} \Delta\theta \lim_{\theta \longrightarrow \infty} \Delta 2\theta} + \sqrt{\cos a^2 + \cos b^2} = 0. \tag{17}$$

Therefore, this study adopts the second-order Butterworth low-pass filter. The original data of the accelerometer and the data changed by FFT are simulated in MATLAB, and

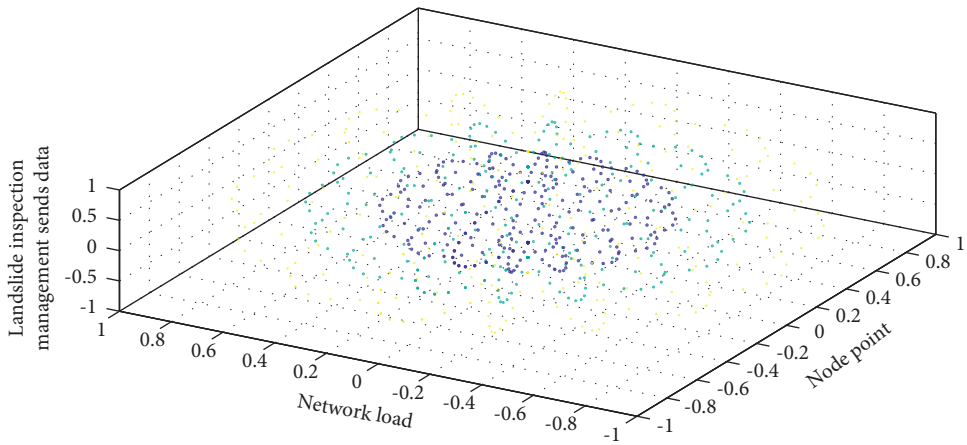


Figure 3: Distribution of data sent by landslide inspection management.

Table 3: Photogrammetric control algorithm of flight platform.

Steps number	Algorithm description	Code explanation
1	Autonomous navigation path $U_i(s)$	Int Index(char ∗ S, char ∗ T)
2	Vertical and oblique images $(n!/r!(n-r)!)$	B=(int ∗)malloc(sizeof(A));
3	The two phases of field $t-1$	Int $i=1, j=1$;
4	Its accuracy results $x-1$	While(i<=s[0]&&j<=T[0])
5	During the application process $\lim_{x\longrightarrow\infty}(n!/r!n!)$	//Merge(A,0,3,7);
6	Based on the Smart3D $\partial s\partial t$	If($S[i]==T[j]$){
7	Contextcapture platform $f(x,t)$	++i; ++j; }else
8	The point cloud of the survey area $\partial f(x,t)$	If($j==0 \| S[i]==T[j]$){
9	UAV can directly carry out stable	{mergesort(A,0,7);
10	Drone tilt photogrammetry $U_i(s,t,x)$	I = i − j + 2; $j=1$;
11	Mages obtained by technology $i, j, k \in R$	Printf("%d ",$A[i]$); }
12	For the overall modeling of the target $F(a,b)$	If($j>$ T[0])return i-T[0];
13	UAV can directly carry out stable $\sum w^2 - s^2$	Else int next[], int pos
14	The host computer software $-b+\sqrt{b^2-4ac}$	Void get_nextone(char T[], int next[]){
15	Geometric center of the flight platform $\cos a$	While(i<=T[0])
16	The tilt angles of the pictures obtained $\cos b$	Int kmpone(char S[], char T[])
17	Object from multiple perspectives $R(\sin x, \cos s)$	While(i < S[0]&&j < T[0])

Table 4: Analysis of autonomous control of safety early warning.

Parameter name	Channel 1 subsidence value	Channel 2 subsidence value	Parameter requirements
Control method	19.31	47.37	External connection
Autonomous communication	19.62	47.74	Wireless communication
Warning bandwidth	19.65	47.77	500~1500 MHz
Control subsidence supply	19.25	47.30	Reserved for external interface
Control subsidence life	18.27	46.11	Continuous work conditions
Warning range	16.56	44.03	−40~−20 dbm
Analysis of channels	14.14	41.11	10

then, the accelerometer data after the second-order Butterworth filter with a cut-off frequency of 30HZ are simulated, and FFT simulation is performed. The processing results are as follows. We regard GCP coordinates as true values, image point coordinates (x, y) as observations, and image internal and external azimuth elements, various distortion parameters, and spatial coordinates of pending points as unknowns.

5. Application and Analysis of UAV Surveying and Mapping Trajectory Scheduling and Autonomous Control Model for Landslide Monitoring

5.1. Preprocessing of Landslide Monitoring Data. For the dangerous situation of landslide monitoring data, in addition to human operation errors, other situations can be avoided as much as possible. The DJI Phantom 4 Pro drone used in this article is equipped with front and rear obstacle avoidance functions, which can avoid hitting obstacles. In the case of weak GPS information, the drone will automatically hover and wait for a signal or return home automatically, which can avoid flight accidents.

On the basis of theoretical research, this article has done systematic software design and development of Beidou-based high-precision autonomous navigation and monitoring technology for UAVs. On the basis of self-made monitoring upper computer, self-made remote navigation communication system, and anonymous flight control system in Figure 4 UAV comprehensive test platform was built, and the software design of UAV high-precision positioning, UAV autonomous navigation inertial navigation algorithm, and UAV autonomous navigation path planning algorithm was completed. The bar is the b-box regression. As of now, the two network modes 1 and 2 do not share parameters, but are trained separately.

After the drone completes the aerial photography of the overall inspection, due to the drone's own image transmission function, the aerial image can be checked at the ground control station, and the landslide can be judged according to the method of visual interpretation. The image can be judged that the target landslide is within the image range. Then, check the quality of the image to see whether it is blurred or deformed. If such problems occur, it is generally caused by camera settings, lens contamination, high wind speed, etc. After the problem is resolved, we retake aerial photography or only re-shoot a part of the image:

$$\iiint \partial \frac{U_i(s,t,x)}{\partial s \partial t} + \frac{\partial f(x,t,s)}{\partial s \partial t} + \frac{\partial f(x,t)}{\partial s} = 1. \tag{18}$$

Aerial triangulation calculation is a critical step in UAV image processing. First, we use photogrammetry software to correct the distortion of the image acquired by the UAV, then perform air triple encryption, and use the results of air triple encryption to perform high-precision matching editing and acquisition. The DSM filter edit the DSM to get the DEM that filters the buildings, etc., and then, use the DEM to digitally differentiate and correct the image to obtain multiple digital orthoimages in units of frames, namely, the DOM, and finally stitch them into the final image data.

5.2. Realization of UAV Surveying and Mapping Trajectory Simulation. In order to realize the autonomous navigation and monitoring of UAVs, it is necessary to realize the technology of remote monitoring of UAVs by the host computer and remote binding tasks for UAVs online, so it needs to have the function of wireless communication. At the same time, a WIFI chip is added. If there is a WIFI signal in the surrounding environment, the host computer software of the monitoring system and the UAV can directly carry out stable and efficient two-way communication in the local area network:

$$\begin{aligned} &\lim_{x \longrightarrow \infty} \frac{n!}{r!(n-r)!} - \lim_{x \longrightarrow \infty} \frac{n!}{r!n!} \\ &= 0, \quad \text{for}\{i, j, k \in R(\sin x, \cos s)\}. \end{aligned} \tag{19}$$

Oblique photogrammetry carries cameras evenly distributed at different angles around the geometric center of the flight platform. There are two-lens oblique photography and five-lens oblique photography. The tilt angles of the pictures obtained by oblique photography are all greater than 3°, and the pictures obtained are called oblique pictures. Since oblique photography captures the target object from multiple perspectives, it is particularly advantageous for the interpretation of ground objects and targets, and has a higher quality assurance for the overall modeling of the target area:

$$\prod_{s.t.}^{a,c,t} F(a,b) * G(\cos a, \cos b) = \overbrace{\frac{-b+\sqrt{b^2-4ac}}{2a} * \cdots * \frac{-b-\sqrt{b^2-4ac}}{2a}}^{\text{for}\{a,b,c \in C\}}. \tag{20}$$

Aerial triangulation is the air three encryption. This is the most important link in the aerial photogrammetric image processing industry, and it is also a difficult point in the entire processing process. Its accuracy results directly affect the later digital elevation model DEM, digital ground model DSM, and digital positive. On the basis of datasets 1 and 2, it has been improved to form a new algorithm body model, which introduces a network structure instead of the selective search algorithm, and integrates the candidate frame generation into the deep neural network, which greatly simplifies the data set. The computational complexity and model complexity of 3, 4 improve the computational speed and the accuracy of the final result.

The dense point cloud is generated using the result of air three encryption, and the digital surface model data of the entire survey area are obtained. The digital surface model is filtered to remove some surface objects such as vegetation and then manually edited to obtain the digital elevation data that meets the requirements. Finally, it can be used to make 3D terrain of landslide.

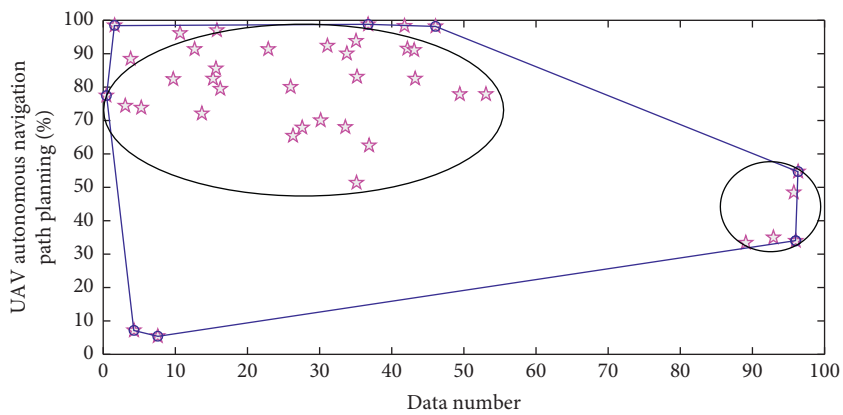


Figure 4: UAV autonomous navigation path planning.

The database is used to store the basic information of the landslide in Figure 5, such as the name of the landslide, lithology, topography, and rainfall, and it can also store the UAV landslide images and the results of the image post-processing. The rationality of database design is related to system calculation speed and accuracy. The research content of this article involves the import and export of a large amount of data. There are many types of stored data.

After the air triple encryption is completed, the DTM is extracted. The software calculates the radiation adjustment on a single image during processing to compensate for the visual effect and then adjusts the matching degree of adjacent images to balance the color of the image in the measurement area. The orthophoto map generated by the software can be exported to TIFF format and can be processed in ArcGIS software.

5.3. Case Application and Analysis. During the application process of the example, the point cloud of the survey area in this article is automatically processed based on the Smart3D ContextCapture platform, and the two phases of field vertical and oblique images obtained by drone tilt photogrammetry technology are combined with camera parameters and field image control points.

Using aerial triangulation technology to calculate the high-precision external orientation elements of each image, generate dense image points with the same name through dense matching, and then, perform matching after adjustment, until the matching points meet the accuracy requirements, and then, high-precision point clouds can be generated:

$$\begin{cases} u_\beta(t) = \cos(wt) + \exp(wt), \\ u_\beta(t-1) = \sin(wt) + \exp(wt), \\ u_\beta(t+1) = \tan(wt) + \exp(wt). \end{cases} \tag{21}$$

The 4G module uses the SIM7100 chip, which is an LTE platform based on Qualcomm's MHD9215 multiplexer. It has multiple frequency bands, including TDD-LTE//GSM/GNSS SMT. Its processor is provided by Cortex A5 (550HZ) and three QD SP6 (up to 500 MHz) built up. SIM7100 has powerful expansion functions, rich interfaces, including UART, SPI, and 12C, and flexible application capabilities, including TCP sampan DP/FTP. During hardware design, the module in Figure 6 is embedded in the navigation communication board and connected to the UART pin of STM32.

The point cloud generated by the tilted image of the drone contains a lot of noise and nonground points. This article processes the point cloud based on the 3D point cloud processing software and forms a high-precision point cloud after processing vegetation and high-voltage towers. Through this module, data can be sent to the WIFI network so as to realize the control and management of the Internet of Things.

The output candidate region of the first step of dataset mode 1 is used as the input of the detection network. A simple configuration before use can realize the networking function of the UART device. The SIM7100 is connected to the physical pin of the serial port 3 of the STM32, and the USR-C322 is connected to the physical pin of the serial port 2 of the STM32:

$$\left\langle \begin{matrix} R(\sin x, \cos s) \\ \sum f(s,t) \times \exp s \end{matrix} \right| \left\langle \begin{matrix} -\langle (\partial f(x,t,s)/\partial s \partial t) \\ \sum f(s,t) \times \sin x \end{matrix} \right| = 0, \quad \text{for} \left\langle \begin{matrix} x > t \\ y < t \end{matrix} \right. \tag{22}$$

The landslide inspection management system needs to record and query the basic information of the landslide. In order to facilitate data entry, management, analysis, and query functions, corresponding data tables must be established to realize the above functions. A table is the most basic data storage unit of a database. It is a two-dimensional structure composed of rows and columns.

Columns are called fields and are used to define the structure of the table. Rows are called records and are used to store a piece of data in the table. We create corresponding fields in the table according to the characteristics of the data and set the appropriate data type for each field. In the landslide inspection management system, the types of data stored are different, and the fields in each table are also different.

According to Figure 7, the analysis of the results of the three sets of data under different image control point layout

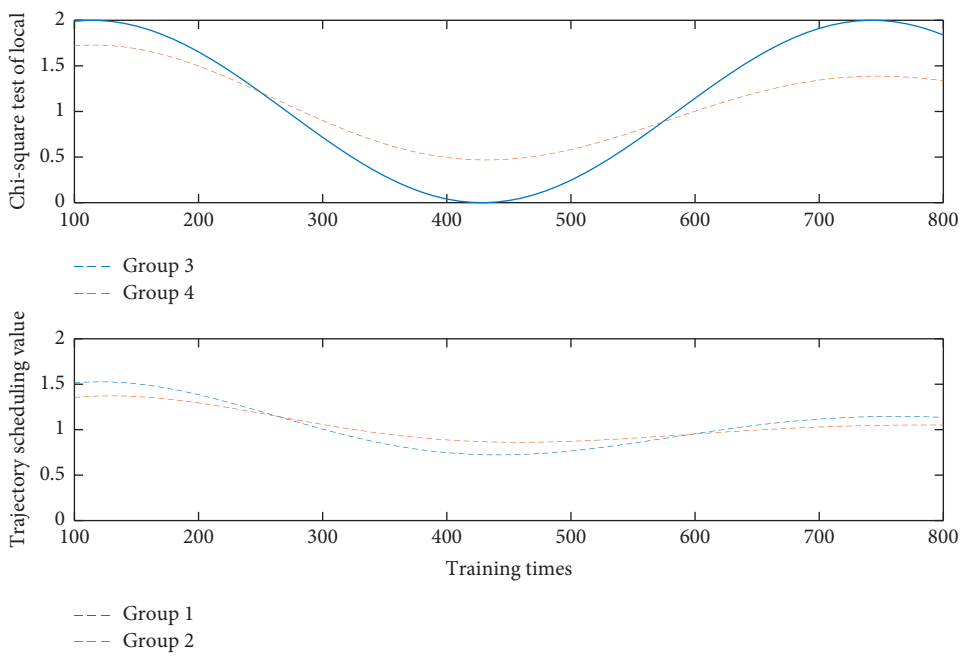


Figure 5: Dispatching distribution of 3D terrain trajectory of landslide.

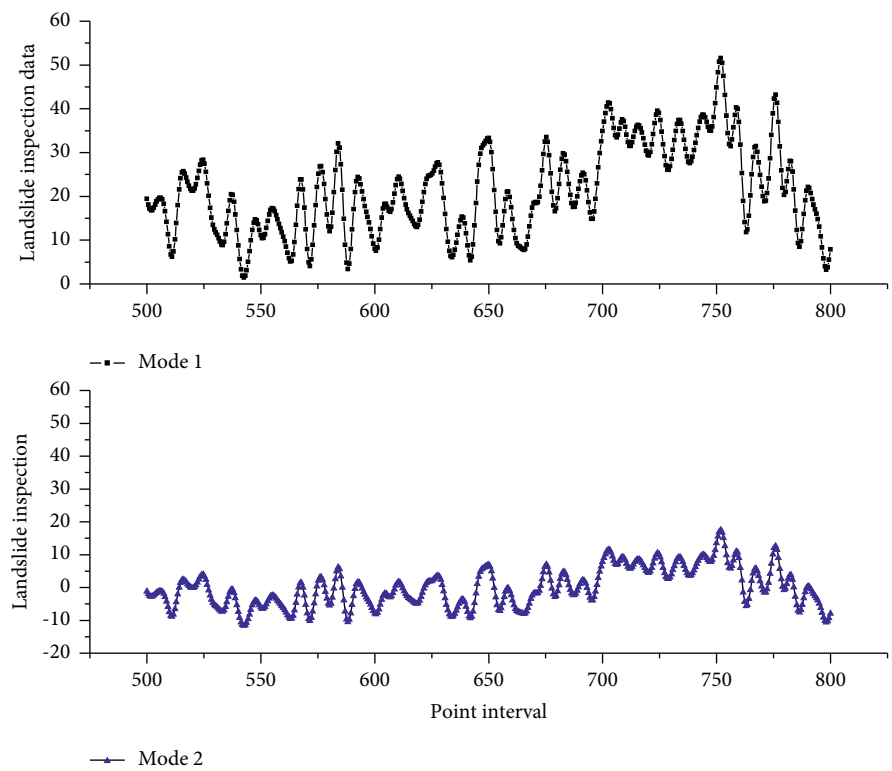


Figure 6: Data distribution of landslide inspection parameters.

schemes shows that the difference between the basic orientation point and the checkpoint of the data source I after using the differential technology is generally smaller than that of the data without using the differential technology. In terms of in-plane error, data source I is at least 2 times higher than data source II, and in elevation error accuracy, data source I can also be improved by more than 2 times compared to data source II.

The texture mapping based on the reconstructed TIN network mainly includes two parts: texture optimization and texture extraction. Due to the large overlap of the oblique images, most of the triangles reconstructed on the surface

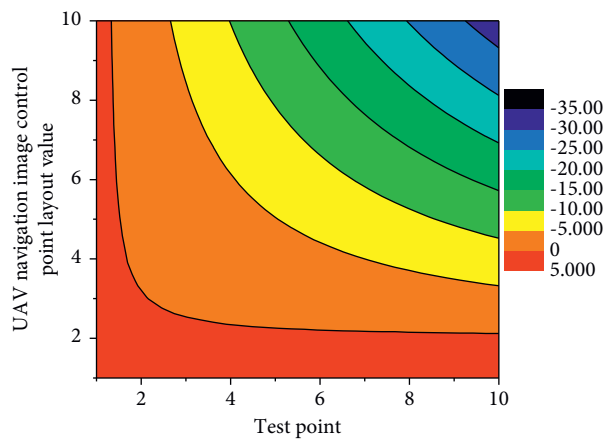


Figure 7: Layout of drone navigation image control points.

correspond to the corresponding textures of multiple images. The texture selection is based on the generated triangle mesh combined with the distortion parameters and the internal and external orientation elements of the image using photogrammetric computer vision methods:

$$\begin{Vmatrix} \cos t & -t \\ t & \sin t \end{Vmatrix} * \begin{Vmatrix} Ke^{-Kwt/2} & \tan t \\ -\tan t & -Ke^{-Kwt/2} \end{Vmatrix} = \begin{Vmatrix} \cos\left(1-K^2wt/2\right)U & 1 \\ 1 & \cos\left(1-K^2wt/2\right)U \end{Vmatrix}. \tag{23}$$

The monitoring data of the wooden piles showed that during the comparison time period, it shows a certain amount of deformation. The cumulative deformation of PM1 was 50 cm, and the cumulative deformation of PM2 was 4.4 cm. It can be seen that the trailing edge of the landslide is relatively broken and divided into pieces by cracks. Specifically, we output a candidate frame, intercept the original image through the candidate frame, pass the intercepted image through conVtool several times, and then output two branches through ROI-pooling, one for target classification and the other for target classification.

This is because the data source I adopts the post-difference technology, and the acquired image POS positioning information is processed by the difference software and the accurate latitude, longitude, and ellipsoid height at the time of camera exposure are calculated with higher accuracy, so the accuracy of the data source I will be higher than that of data source II; in addition, with the increase in the number of control points and the changes in positions, the error changes in the empty three planes have not changed significantly. This is because the nine points of the image control point scheme 1 are already available in this survey area. The plane accuracy is controlled to a higher level, but the error in the elevation is significantly improved, and the minimum is increased to more than 2 times.

6. Conclusion

This article studies the principle of landslide monitoring technology, and its double-difference positioning model introduces the core technology of UAV repair and detection and solving the ambiguity of the whole week, and derives the theoretical formula for the calculation of virtual observations and the establishment of the double-difference model. First of all, related software development is carried out on the network landslide monitoring technology, and the network landslide monitoring is applied to the positioning of the UAV. Secondly, it studied in detail the UAV surveying and mapping scheduling method based on complementary filtering, including Butterworth low-pass filtering technology and accelerometer gyroscope complementary filtering technology. On the basis of surveying and mapping scheduling, the vertical third-order complementary autonomous control algorithm with observer delay correction and the horizontal dual-observation autonomous control algorithm are studied, and simulations have verified the practicability. After that, according to the geometric attribute information of the ground fissures in the UAV image, the MF template and FDOG template that meet the vertical profile curve are constructed, and the UAV image is filtered and calculated separately to achieve the purpose of enhancing the ground fissure signal, and linear stretching is adopted. The method eliminates the difference in the value range obtained by the filtering operation and performs the difference operation between the two results to highlight the ground fissure signal in the image and weaken the edge signal of the ground feature. Finally, on the basis of autonomous control, the path planning algorithm is studied and the feasibility of the algorithm is verified by simulation, so that the UAV can avoid obstacles and complete the landslide monitoring task according to the optimal path under certain conditions.

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 in this article.

Acknowledgments

This study was supported by "The Special Fund Construction Project of Superior Subject in Shaanxi Province colleges and universities (Grant no. 0602)."

References

[1] S. Aggarwal and N. Kumar, "Path planning techniques for unmanned aerial vehicles: a review, solutions, and challenges," *Computer Communications*, vol. 149, pp. 270–299, 2020.

[2] L. Patané, "Bio-inspired robotic solutions for landslide monitoring," *Energies*, vol. 12, no. 7, p. 1256, 2019.

[3] M. Erdelj, M. Król, and E. Natalizio, "Wireless sensor networks and multi-UAV systems for natural disaster management," *Computer Networks*, vol. 124, pp. 72–86, 2017.

[4] D. Giordan, M. S. Adams, I. Aicardi et al., "The use of unmanned aerial vehicles (UAVs) for engineering geology applications," *Bulletin of Engineering Geology and the Environment*, vol. 79, no. 7, pp. 3437–3481, 2020.

[5] P. Arena, F. D. Pietro, A. L. Noce, S. Taffara, and L. Patanè, "Assessment of navigation capabilities of mini cheetah robot for monitoring of landslide terrains," *Research and Technology for Society and Industry (RTSI)*, vol. 9, pp. 540–545, 2021.

[6] D. Mishra and E. Natalizio, "A survey on cellular-connected UAVs: design challenges, enabling 5G/B5G innovations, and experimental advancements," *Computer Networks*, vol. 182, Article ID 107451, 2020.

[7] C. Kyrkou, S. Timotheou, P. Kolios, T. Theocharides, and C. Panayiotou, "Drones: augmenting our quality of life," *IEEE Potentials*, vol. 38, no. 1, pp. 30–36, 2019.

[8] H. Hildmann and E. Kovacs, "Review: using unmanned aerial vehicles (UAVs) as mobile sensing platforms (MSPs) for disaster response, civil security and public safety," *Drones*, vol. 3, no. 3, p. 59, 2019.

[9] H. Wang, H. Zhao, J. Zhang, D. Ma, J. Li, and J. Wei, "Survey on unmanned aerial vehicle networks: a cyber physical system perspective," *IEEE Communications Surveys & Tutorials*, vol. 22, no. 2, pp. 1027–1070, 2019.

[10] J. Besada, L. Bergesio, I. Campaña et al., "Drone mission definition and implementation for automated infrastructure inspection using airborne sensors," *Sensors*, vol. 18, no. 4, p. 1170, 2018.

[11] M. Labbadi and M. Cherkaoui, "Robust adaptive backstepping fast terminal sliding mode controller for uncertain quadrotor UAV," *Aerospace Science and Technology*, vol. 93, Article ID 105306, 2019.

[12] K. O. Said, M. Onifade, J. M. Githiria et al., "On the application of drones: a progress report in mining operations," *International Journal of Mining, Reclamation and Environment*, vol. 35, no. 4, pp. 235–267, 2021.

[13] A. W. A. Hammad, B. B. F. da Costa, C. A. P. Soares, and A. N. Haddad, "The use of unmanned aerial vehicles for dynamic site layout planning in large-scale construction projects," *Buildings*, vol. 11, no. 12, p. 602, 2021.

[14] M. M. Eltabey, A. A. Mawgoud, and A. Abu-Talleb, "The autonomy evolution in unmanned aerial vehicle: theory, challenges and techniques," *Advances in Intelligent Systems and Computing*, pp. 527–536, 2020.

[15] A. Khan, S. Gupta, and S. K. Gupta, "Multi-hazard disaster studies: monitoring, detection, recovery, and management, based on emerging technologies and optimal techniques," *International Journal of Disaster Risk Reduction*, vol. 47, Article ID 101642, 2020.

[16] R. Zanol, F. Chiariotti, and A. Zanella, "Drone mapping through multi-agent reinforcement learning," *IEEE Wireless Communications and Networking Conference (WCNC)*, vol. 4, pp. 6-7, 2019.

[17] H. Valente de Figueiredo, O. Saotome, E. H. Shiguemori, P. Silva Filho, and V. V. Estrela, "Image database of low-altitude UAV flights with flight condition-logged for photogrammetry, remote sensing, and computer vision," *Imaging and Sensing for Unmanned Aircraft Systems: Volume 2: Deployment and Applications*, vol. 2, pp. 91–111, 2020.

[18] K. Julge, A. Ellmann, and R. Köök, "Unmanned aerial vehicle surveying for monitoring road construction earthworks," *The Baltic Journal of Road and Bridge Engineering*, vol. 14, no. 1, pp. 14–17, 2019.

[19] H. Zhang, E. Aldana-Jague, F. Clapuyt, F. Wilken, V. Vanacker, and K. Van Oost, "Evaluating the potential of post-processing kinematic (PPK) georeferencing for UAV-based structure- from-motion (SfM) photogrammetry and surface change detection," *Earth Surface Dynamics*, vol. 7, no. 3, pp. 807–827, 2019.

[20] R. Basir, S. Qaisar, M. Ali, N. Ahmad Chughtai, M. Ali Imran, and A. Hashmi, "Performance analysis of UAV-enabled disaster recovery networks," *Autonomous Airborne Wireless Networks*, pp. 157–194, 2021.

[21] Y. Sahraoui, C. A. Kerrache, M. Amadeo et al., "A cooperative crowdsensing system based on flying and ground vehicles to control respiratory viral disease outbreaks," *Ad Hoc Networks*, vol. 124, Article ID 102699, 2022.

[22] D. Hermle, M. Keuschnig, I. Hartmeyer, R. Delleske, and M. Krautblatter, "Timely prediction potential of landslide early warning systems with multispectral remote sensing: a conceptual approach tested in the Sattelkar, Austria," *Natural Hazards and Earth System Sciences*, vol. 21, no. 9, pp. 2753–2772, 2021.

[23] Á. Török, Á. Barsi, G. Bögöly, T. Lovas, Á. Somogyi, and P. Görög, "Slope stability and rockfall assessment of volcanic tuffs using RPAS with 2-D FEM slope modelling," *Natural Hazards and Earth System Sciences*, vol. 18, no. 2, pp. 583–597, 2018.

[24] I. Jeelani and M. Gheisari, "Safety challenges of UAV integration in construction: conceptual analysis and future research roadmap," *Safety Science*, vol. 144, Article ID 105473, 2021.

[25] S. Ekici, "Design of unmanned helicopter equipped with turboshaft engine for agriculture spraying mission based on thermodynamic analysis," *Journal of the Institute of Science and Technology*, vol. 10, no. 1, pp. 532–546, 2020.

[26] L. Tan, J. Guo, S. Mohanarajah, and K. Zhou, "Can we detect trends in natural disaster management with artificial intelligence? a review of modeling practices," *Natural Hazards*, vol. 107, no. 3, pp. 2389–2417, 2021.

Hindawi
Journal of Robotics
Volume 2022, Article ID 8565913, 12 pages
https://doi.org/10.1155/2022/8565913

Research Article

Recognition and Localization of Target Images for Robot Vision Navigation Control

Muji Chen

College of Information Engineering, Henan Vocational College of Agriculture, Zhengzhou, Henan 451450, China

Correspondence should be addressed to Muji Chen; 2004110217@hnca.edu.cn

Received 20 January 2022; Accepted 5 March 2022; Published 24 March 2022

Academic Editor: Shan Zhong

This paper focuses on a visual navigation control system for mobile robots, recognizing target images and intelligent algorithms for the navigation system's path tracking and localization techniques. This paper examines the recognition and localization of target images based on the visual navigation control of mobile robots. It proposes an efficient marking line method for recognizing and localization target images. Meanwhile, a fuzzy control method with smooth filtering and high efficiency is designed to improve the stability of robot operation, and the feasibility is verified in different scenarios. The corresponding image acquisition system is developed according to the characteristics of the experimental environment, and the acquired images are preprocessed to obtain corrected grayscale images. Then, target image recognition and linear fitting are performed to obtain target image positioning. The system calculates the angle and distance of the mobile robot, offsetting the target image in real time, adjusting the output signal, and controlling the mobile robot to realize path tracking. The comparison of sensor data and path tracking algorithm results during the experiment shows that the path tracking algorithm achieves good results with an angular deviation of $\pm 1.5^{\circ}$. The application of RANSAC algorithm and improved Hough algorithm was analyzed in visual navigation control, and the two navigation line detection algorithms based on the image characteristics of the target image were improved in the optical detection area of the navigation line for the shortcomings of the two algorithms in visual navigation control, and the algorithms before and after the improvement were compared.

1. Introduction

The mobile robot is an essential branch of robotics. It is an intelligent robot control system capable of detecting and sensing the environment through various sensors and carrying out independent analysis, planning, and decision-making based on environmental information and its state [1]. The research field of mobile robotics involves many kinds of interdisciplinary theories and technologies, including computer vision, sensor information technology, communication technology, motion control theory, and mechanical engineering. The current hot wave of artificial intelligence also affects the research progress of mobile robotics [2]. With the rapid development of information technology, computer microelectronics, and network technology, mobile robotics has also developed rapidly, and more and more new robots with special functions have been introduced. The intelligence level of robots has been improving. In the 21st century, attention is on the robot's perception of the external environment and autonomy. The new direction of robotics is bound to develop toward practicality and intelligence. Mobile robots have been widely used in traditional industry and agriculture and will be further expanded to new sectors, services, defense and security, and medical services and will be commonly used in unsuitable and dangerous situations, such as deep sea and space. Therefore, the broad application prospect of mobile robots has made the research in this field receive widespread attention worldwide.

Autonomous mobility, which gives robots the ability to explore their environment more fully, dramatically increases the complexity of the tasks they can accomplish. State estimation during movement is a constant topic in mobile robotics research [3]. The primary consideration in

designing a reasonable and efficient state estimation method is the type of sensor the robot is equipped with and the characteristics of the data acquired by that type of sensor, i.e., the construction of a sensor observation model. The information that a mobile robot carries about itself and its environment is the source of all information in the subsequent navigation process and determines the form of information processing in the following global positioning and attitude tracking, map building, environment understanding, path planning, and motion control, and task execution. How to deal with the uncertainty contained in the perceptual information and how to design efficient cognitive methods to deal with the mental uncertainty based on the environmental information contained in the perceptual report are the significant challenges for building mobile navigation systems and must also be predicated on the construction of mobile robot observation models [4]. At the same time, the estimation of its positional attitude is the basis and prerequisite for performing other processes during the execution of tasks by mobile robots [5]. Therefore, the performance of the state estimation method will significantly affect the performance of the whole navigation system. Thus, the mobile robot observation model and the underlying state estimation model are introduced for two sensors, laser sensor, and RGB-D vision camera, respectively, to illustrate the robot state estimation process under different observation information and its uncertainty expression form further elaborates its problems. Specifically, the observation model is constructed for the laser sensor for the mobile robot equipped with the laser sensor [6]. Based on the observation model, various forms of observation similarity measures are given, and the characteristics of each form are analyzed. Based on this, a general model of the global localization process of the robot on the raster map is given. Finally, since global localization results are often uncertain and multi-hypothesis, the probability-based state tracking model is introduced in this chapter. For the visual observation model, the camera projection model is described, the projection model of spatial points to the camera plane is described, and the method to recover the spatial position of pixels is given. Based on this, a technique for global positional estimation based on the current observation and the feature matching results in the worldwide map is described. Due to the bias of feature observation, there is uncertainty in the global positional estimation results.

Robotics has been rapidly improved, thanks to the rapid development and maturity of microcomputer technology, sensors, and other related technologies. Intelligent robots have been popularized and applied in various fields such as civil, military, and scientific research. The research results of intelligent robots are more prominent in many developed countries and intelligent. Highly automated intelligent robots have been put into many fields such as aerospace, geological exploration, scientific exploration, rescue and disaster relief, such as China's lunar rover "Moon Rabbit". Some low-cost, clever robots are also coming into daily life and are used in many indoor environments in homes or offices, such as floor cleaning robots [7, 8]. We are now in a critical period of the modern manufacturing industry upgrading in the industrial field. More and more intelligent robots are needed to liberate labor, improve production efficiency, save energy consumption, etc. Visual inspection area image recognition is the basis of navigation line extraction. The quality of image segmentation affects the navigation line extraction and the size of the error in the measurement results of navigation parameters. In the navigation line region established by ultrasonic measurement, the navigation line visual detection region is set as the target operation domain for a series of image processing algorithms, and the detection region is dynamically tracked and set based on the detection results of adjacent frames; preprocessing image algorithms such as inverse color transformation and histogram equalization are specifically analyzed to enhance the different target images in the detection region differentially.

2. Related Works

Through the continuous development of electronic hardware technology and control disciplines, by the 1960s, some European countries already had various forms of mobile robots. With the rapid growth of processors in the 70s and 80s, mobile robots have made significant flexibility and stability. However, the main application scenarios are still the warehousing industry and logistics and transportation systems [9]. In the 1990s, the degree of intelligence and automation of mobile robots was further improved with the rapid development of computers, electronics, communications, and image processing technologies, and mobile robots adapted to various working environments were born, which have been widely used in the material assembly, home appliance production, chemical industry, food, and many other industries. The vision-based mobile robot navigation technology has been a new research boom in recent years and is one of the essential directions of mobile robot guidance technology research work. Research laboratories in universities in the countries that first researched visual navigation technology for mobile robots have achieved significant research results [10]. Many of the results have been applied to actual industrial production and even to the daily lives of the general public.

The Robot Vision Laboratory was the first to develop a vision-guided mobile robot based on map construction, which is accomplished through scene reconstruction using vision sensors to capture photos of the scene [11]. The laboratory at Purdue University has developed an active binocular stereo-based vision-guided mobile robot, Peter, which acquires 3D information of the operating environment and path obstacles, and combines 8 radar scanners, 24 ultrasonic sensors, 8 infrared distance sensors, and a passive infrared motion sensor to achieve flexible operation. The Intelligent Robotics Laboratory at Osaka University has conducted in-depth research on vision navigation and developed a mobile robot based on monocular vision navigation, which can detect the surrounding environment extensively by rotating its vision sensors and obtaining the positioning information, travel distance, and turning angle of the mobile robot by a rotary encoder and potentiometer

[12]. The vision navigation technology is widely used in advanced countries, such as autonomous lawnmowers, Mars landing vehicles, driverless vehicles, and Kiva robots for Amazon's unmanned warehouses.

Global localization of mobile robots is the basis of navigation. The localization process is based on observations to form feature representations and perform feature retrieval and area inference in a global map. Lasers and vision are two standard sensors used in the indoor navigation process [13]. Lasers can provide stable distance sensing information, which is advantageous in obstacle avoidance and motion planning tasks. However, it can only perceive flat information and is relatively homogeneous. Vision can perceive richer data, but its geometric perception range is lower, while laser has a more extended and more stable perception range. In the global localization process, there are multiple possible regions of positional distribution in the environment due to the existence of similar areas and the incompleteness of the perceived information. How to eliminate the spatial perceptual ambiguity to make an accurate estimation of the current poses is essentially a global optimization problem. The robot measures the degree of information consistency between the recent observation and the expected observation corresponding to the estimated position based on the constructed objective function. There is considerable international work on visual navigation motion control from image analysis to driving command generation [14]. In image segmentation, since the opening of the ImageNet challenge, there has been significant work based on the traditional feature point segmentation method to the currently popular deep learning methods: the use of deep learning to achieve semantic recognition of images and segmentation methods, with a very high number of references to achieve autonomous driving using images, or autonomous driving using reinforcement learning methods, the mainstream research in the field of autonomous driving. There is a large body of literature on autonomous driving, including the use of road segmentation and the use of environmental features for autonomous driving.

3. Model Design of the Target Image Recognition and Localization System Based on Robot Vision Navigation Control

3.1. Robot Vision Navigation Control System Construction. The visual navigation software is mainly used to process the images obtained from the camera in real time, then the navigation lines are extracted, navigation decisions are made after calculating the navigation parameters of the robot, and finally control signals are sent to the lower computer to control the weeding robot for automatic navigation. The visual navigation software in this paper is written on the Visual Studio 2015 development platform based on MFC with Open CV, where the algorithms are implemented in C++ and C languages. The functional framework of the visual navigation software is shown in Figure 1.

The visual navigation software comprises five parts: the information acquisition module, the image processing module, the navigation decision module, the information communication module, and the information storage module. The information acquisition module is the basis and preparation of the visual navigation. Its primary function is to obtain an accurate image with distortion correction after the calibration of the camera [15]. The IMU (Inertial Measurement Unit) detects the pitch angle of the camera in real time to correct the deviation of the camera pitch angle due to the vibration of the robot during the navigation process. The image processing module is the core part of visual navigation. The main functions of the image processing algorithm are image ROI construction, detection based on a deep learning model, detection frame clustering, image grayscale and smoothing filtering, corner point feature extraction in the detection frame, and navigation line fitting. The navigation decision module controls the motion of the robot based on the navigation information obtained after image processing. This study uses the fuzzy control method developed by our group to control the robot's movement. The primary process of the navigation decision module is to extract the navigation lines in the ROI after image processing, then to find out the dominant route, solve the position deviation and angle deviation of the robot, then input the navigation deviation parameters into the fuzzy control decision, and finally derive the control command of the robot. The function of the communication module is to realize the serial communication between the visual navigation upper computer software and the lower computer hardware control system. Through the configuration of the serial parameters and the development of the data transmission protocol, the control information is finally converted so that the digital signal output by the navigation software is converted to the level signal for controlling the robot motion [16]. The information saving module is responsible for saving the camera calibration parameters to avoid tedious, repetitive calibration work content, and saving the position deviation and angle deviation during the robot autonomous navigation process, which is used for the quantitative analysis of the accuracy of the image processing algorithm. It is divided into six parts: image processing display, camera parameter setting, camera calibration setting, serial port setting, production and saving of data, and robot motion control.

(1) The image processing display section displays the processing effects of the main stages of image processing in real time. The processing details of the image algorithm can be visually observed and analyzed for any problems that exist.

(2) The camera parameter setting section is used for camera ID selection, camera resolution setting, and image correction by calibrating internal camera reference.

(3) The camera calibration section is used to set the calibration board parameters, calibrate the camera, and save the calibration parameters. The calibration parameters saved locally can be read directly in the subsequent image correction operation, repeatedly avoiding tedious camera calibration work.

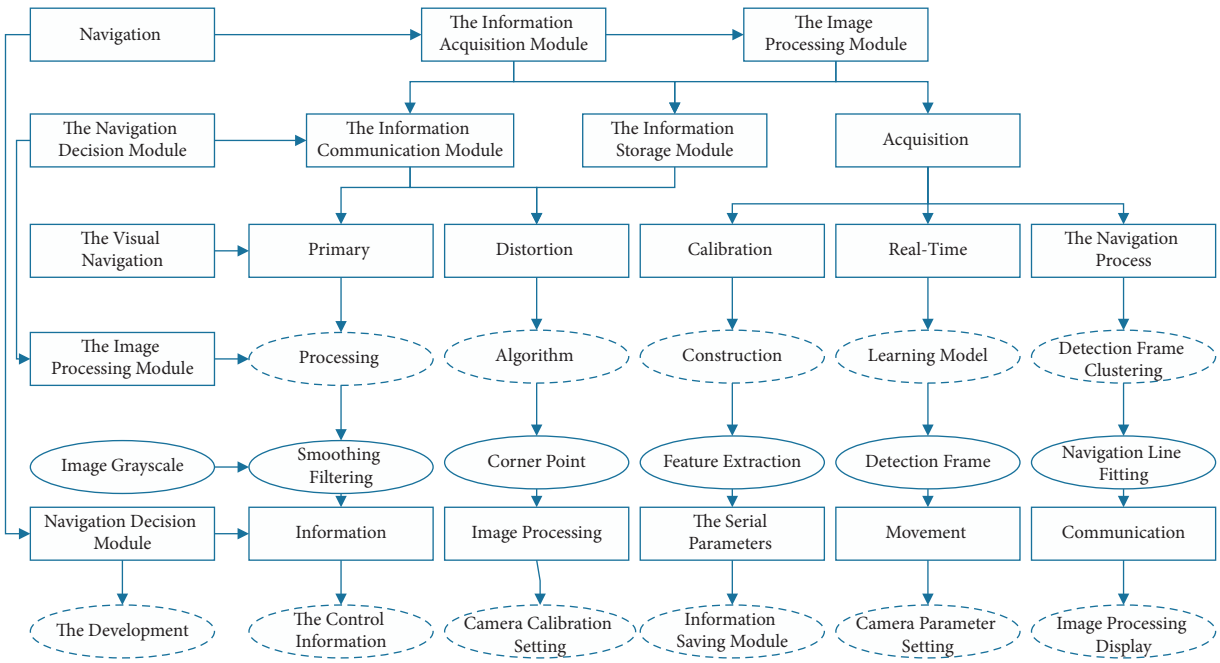


Figure 1: Functional framework of visual navigation software.

(4) The serial port selection section selects and configures the communication serial port between the upper and lower computers.

(5) The data display and saving section is used to display the pitch angle of the camera measured by the IMU in real time, to solve the position and angle deviation of the robot, to store the position and angle deviation of the weeding robot during automatic navigation in txt file format, and to select the size of the filter kernel.

(6) Robot motion control is divided into manual control mode and automatic navigation mode. The manual control mode is used to regulate the robot's position and adjust the part of the welding robot in the water field. The automatic navigation mode is for the robot to track forward along the seedling navigation line according to the control command of fuzzy decision. The robot vision navigation flow chart is shown in Figure 2.

Visual navigation of mobile robots is to collect road information through a camera to identify marking lines and guide the robot. Therefore, to accurately show the robot along the desired path, the body model must be built first to realize the conversion between the image space coordinate system corresponding to the camera and the world coordinate system centered on the robot's body. The mobile robot body structure is fixed, and its Kinect camera optical axis is parallel to the road surface, 34 cm high from the ground, and its maximum adequate vertical view is $a+b$ because there are shields above and below the camera, i.e., the bottom of the image taken by the camera is the road surface 66 cm in front of it. The Tourtellot mobile robot in this paper adopts a four-wheel structure, in which the left and right wheels are the driving wheels, and the front and rear wheels are the driven wheels. The wheeled mobile robot can accomplish a variety of motions, mainly by controlling the rotational speed of its left and right drive wheels, respectively. Therefore, to effectively manage the movement of the mobile robot, its kinematic model must be analyzed first. The position state of the mobile robot at two adjacent moments, with the x-axis forward as the robot's forward direction, where v_r and v_t are the velocities of the robot's left and right drive wheels, respectively, the angle the robot has turned at the adjacent moment t, r is the distance between the left and right drive wheels, and r is the radius of the circular arc motion at the adjoining moment where the forward velocity of the mobile robot is equal to the average rate of its left and right wheels, assuming that the steering angle is slight, the formula can be obtained as follows:

$$\begin{cases} v_t = \dfrac{v-1}{\sqrt{lw/2}}, \\ v_r = \dfrac{v+1}{\sqrt{2/lw}}. \end{cases} \tag{1}$$

Before the visual navigation line improvement RANSAC algorithm extraction, each pixel point in the optical detection area of the navigation line needs to be calculated in u_{IPM}, the directional gradient I_{upm} and v_{IPM}, directional gradient I_{vPM}, and the product $I^2_{u_{ipm}} I^2_{v_{iPM}}$ and sum $I_{u_{ipm}} I_{v_{ipm}}$ of the two gradient directions are calculated. Then, the Gaussian filtering is performed, and the template parameters are normalized; then, the corner point response value P of each image element is calculated, and the P value smaller than the threshold is set to zero. Finally, within the 3×3 neighborhood, the local nonmaximum values are

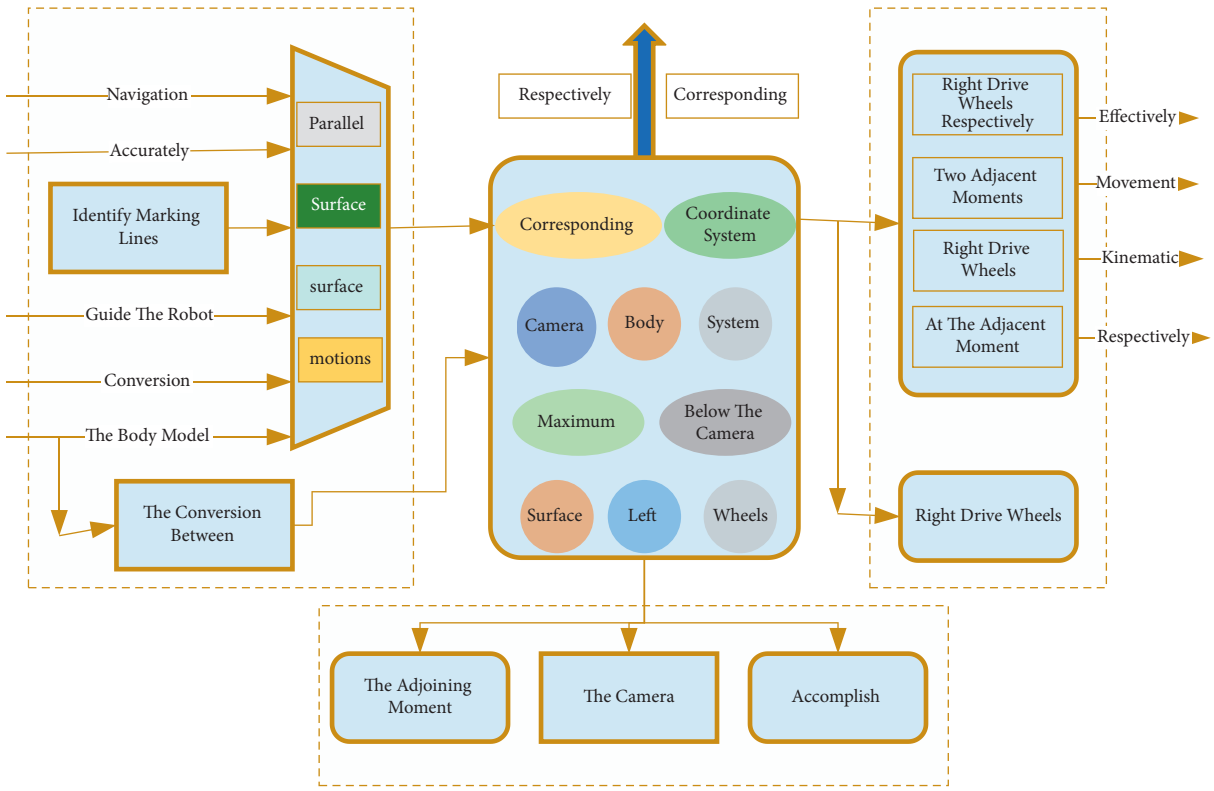


Figure 2: Flow chart of robot vision navigation.

suppressed, and the remaining local maximum values are output as corner points. The specific steps of the improved RANSAC algorithm for visual navigation line detection are as follows:

Step 1: A minimum of two data points is required for each random sampling. The number of samples in the data set You for each corner point needs to be guaranteed to avoid misfitting the harvesting navigation line due to too few corner points.

Step 2: Initial estimation of two randomly selected data points (u^1_{IPM}, v^1_{Ipm}) and (u^2_{IPM}, v^2_{Ipm}) models from the corner point data set U to obtain a linear model M.

$$v_{ip} = \frac{\sqrt{v^2_{ipm} + v^1_{ipm}}}{u^2 + u^1}. \quad (2)$$

Step 3: For the remaining data in the data set U, calculate the pixel distances d_m to the linear model M in turn, if the distance threshold A is d_m satisfied, but the point into the set U_s as an ingroup point together with the extracted sample points, and the other points as outgroup points.

Step 4: Count the number of ingroup points in the set U_s of ingroup points s. If s satisfies the threshold S_T of the number of ingroup points, refit the ingroup points U_s in the ingroup points using the least-squares method, and update the linear model M. If it does not satisfy, discard this linear model.

Step 5: Repeat the hypothesis and judge the mathematical model to find the ingroup points step N_R by step, compare the ingroup point set U_{SMax} with the most significant number of ingroup points, and output its corresponding linear model M to get the navigation path line L_M.

The essence of the Hough transform is to map the image to its parameter space, which requires the computation of all HM edge points and requires a large amount of memory space and operations. The improved Hough transform processes m_H $(m_H < M_H)$ only one edge point in the input image, and the selection of m_H this edge point is somewhat random. In addition, the enhanced Hough transform algorithm can obtain two endpoints of a straight line in the detection image and accurately locate the detected target straight line. The specific detection processes are as follows:

(1) In the image of the detection area, an edge point is randomly selected (u_{IPM}, v_{Ipm}) and mapped to the polar coordinate system to obtain a family of straight lines through the edge point. Suppose the edge point has been marked on a straight line. In that case, the random selection is continued among the remaining issues, and the polar coordinate equation where the family of straight lines through the edge point lies is obtained until all the edge points are randomly selected. The polar coordinates of the lines passing through the edge point are the opposite equation.

$$r = \frac{u_{ipm} \cdot \cos\theta}{\sqrt{v_{ipm} \cdot \sin\theta}}. \tag{3}$$

(2) The Hough transform of randomly selected edge points and calculation of the cumulative sum.

(3) In the Hough space, select the edge point that reaches the maximum value and continue to the next step when the issue is greater than the threshold, otherwise return to step1.

(4) The edge point where the maximum value reached by step3 is taken as the starting coordinate point. The displacement is carried out along the straight-line direction. The two endpoints of the line are detected to have coordinates of (u^b_{IPM}, v^b_{IPM}) and (u^e_{IPM}, v^e_{IPM}), respectively.

(5) When the length of the line obtained from the detection reaches a particular threshold value, the line is output as a result, and the detection is continued by returning to the initial step.

Compared with the standard Hough transform algorithm, the improved Hough transform has significantly improved memory consumption and computation. The improved probabilistic Hough algorithm can effectively avoid the interference of non-harvesting dividing lines and achieve accurate detection of visual navigation lines. In the complete detection process of optical navigation lines, the average processing time of a single frame for navigation line detection based on the original probabilistic Hough transform algorithm is 77.4 Ms. In comparison, the average processing time of a single frame based on the improved probabilistic Hough algorithm is 54.6 Ms. The enhanced algorithm also improves the processing speed compared with the original algorithm.

3.2. Target Image Recognition and Localization Model Design. The process of digital image acquisition and transmission can be disturbed by many factors, which can cause differences between the digital image and the real object scene and can affect the image processing of the vision system in the later stage. Preprocessing operations such as grayscale, image enhancement, and filtering must be performed on the original image captured by the camera at the beginning of the vision system. Then, the features of the preprocessed image are extracted, and then the parts are matched with those of the template image to identify the target strip of smoke. The surface of the identification target in this paper has robust texture features. The images captured by the camera are generally colorful. They contain much information, extending the time of processing images by the binocular vision system and considering the efficiency requirements of recognition and localization technology to improve the efficiency of the binocular vision system. The color images are to be converted into grayscale maps [17]. The R, G, and B components in the color image are converted into the same values. Since each pixel has a different value for the R, G, and B components, other colors can be displayed in the color image. The principle of target image visual localization is shown in Figure 3.

The amount of data in the image is also three times the number of pixels. When an image is grayed out, the information contained in the image becomes one-third of the original image [17]. The image pixels differ only in brightness and are all displayed in gray. The various colors in a color image are composed of the three base colors R, G, and B. In digital images, if the R, G, and B base colors are finely divided, the more colorful the image can be and the more information the image contains. The R, G, and B base colors are red grayscale, green grayscale, and blue grayscale. Weighted value method: according to different indicators to the original color image R, G, and B components multiplied by the corresponding weight to find the weighting, the expression is as follows:

$$\text{gray} = \sqrt{\sum W_R + R} \times \sqrt{\frac{W_G}{G} + \frac{W_B}{B}}. \tag{4}$$

Image enhancement is also critical to the overall binocular vision system and is necessary for processing images. Image enhancement highlights essential information in the image to meet the system's requirements and eliminate or weaken redundant details irrelevant to the system. Enhanced image processing makes the image more compatible with human visual habits and is intended for specific application purposes. After the image is enhanced, only the ability to distinguish information increases, while the background information is not added. The improved picture is more suitable for the application than the original image in specific scenarios. The standard methods for image enhancement are segmented linear transform enhancement and histogram equalization. The segmented linear transformation enhancement: Suppose the gray map function before the enhancement transformation is $f(r, c)$ and the grayscale range is $[O, M_f]$, after the segmented linear transformation enhancement, the gray map function is $g(r, c)$ and the grayscale range is $[O, M_g]$ and the formula.

$$g(r, c) = \sum_{g}^{f} \sqrt{\frac{d-c}{b-a}} + \frac{m_g - d}{m_f - b}. \tag{5}$$

A grayscale histogram is a statistical graph of the distribution of gray levels, representing the proportion of each gray level pixel in the total number of pixels in a digital image. The histogram can describe the general situation of a grayscale image, such as the degree of contrast between light and dark, the frequency of each gray level, the distribution of gray levels in the picture, etc. The gray histogram is a function of the gray level, with the gray value as the horizontal coordinate and the number of pixels as the vertical coordinate [18]. The gray histogram of an image has the following properties: the gray histogram is a statistical result of the number of occurrences of the gray value of all pixels in the picture, which does not reflect the specific position of the gray value pixels in the image, but only the number of occurrences of different gray values; a pair of images only corresponds to a couple of histograms. The histogram of a

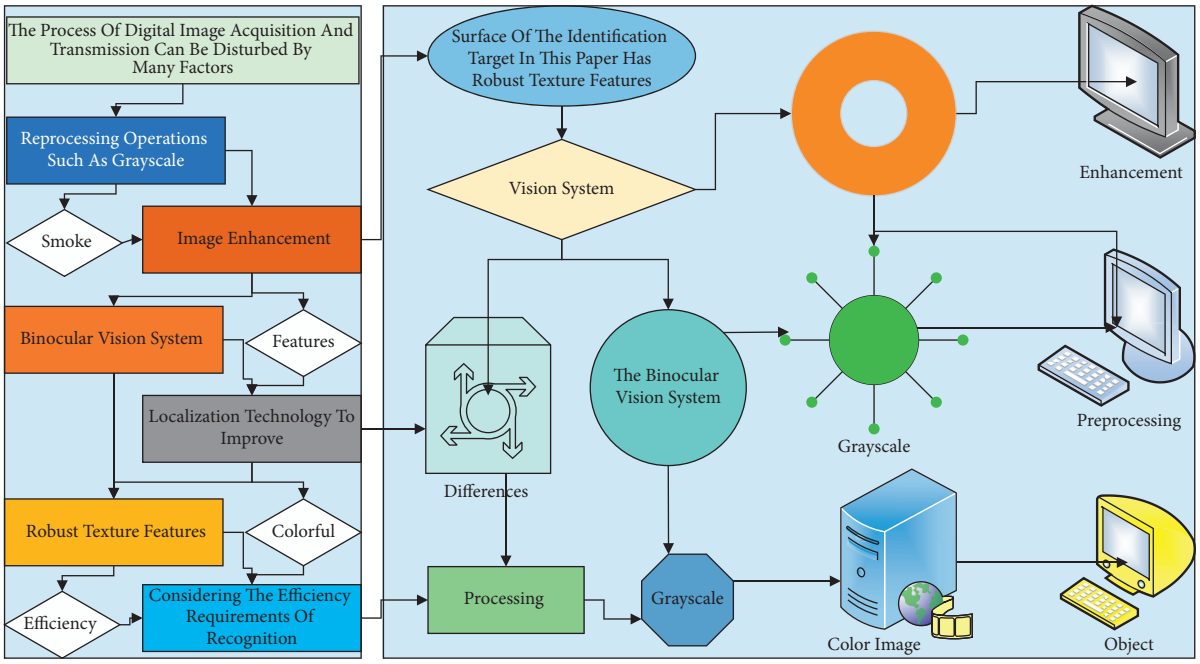


Figure 3: Target image visual localization principle.

team of images corresponds to only one histogram, but one histogram can be reversed to correspond to different ideas; the grayscale histogram counts the number of pixels with the same gray value in an image, so the grayscale histogram of a pair of images is equal to the sum of the histograms of all parts of the picture. The histogram with a $[O, L-1]$ range of gray values is a discrete function.

$$S = \int_{2}^{0} P_R + (w + dw). \tag{6}$$

The robot uses a trained feature point regression model and a topological structure regression model to estimate the robot's poses online. First, a pyramid is constructed for the acquired image, and SURF features and descriptors are extracted. Let $F = \{(p_i, v_i), i = 1\}$ be the location of the extracted feature points and their descriptors. Based on the depth map of the current observation, the coordinates corresponding to the feature points in the camera coordinate system can be obtained. Let $S = \{s \in R^3, i = 1\}$ denote the set of points corresponding to the features. Then, the topological positions corresponding to the current image are predicted using the topological position regression model to obtain multiple candidate topological positions. At the same time, the spatial coordinates in the world coordinate system are indicated for the features in F using the feature point regression model. Let $M = \{v_i = m_1 + m_u, i = 1\}$ denote the set of prediction results for feature points. M contains only the prediction results located at that topological node for each topological location. Due to the existence of erroneous predictions and the fact that each point has multiple projections, the global positional estimation problem of the camera can be expressed.

$$t = \arg\min \sum_{1} (\min \in |m - ts|). \tag{7}$$

4. Analysis of Results

4.1. Analysis of the Robot Vision Navigation Control System. In the process of autonomous navigation, the robot calculates the position deviation and angle deviation of the robot relative to the navigation line based on the seedling navigation line extracted by the vision system and continuously corrects the heading based on the variation during the forward motion [19]. Therefore, the positioning error of the weeding robot concerning the navigation line directly affects the navigation control process of the robot and must be measured and analyzed. The angular deviation error was measured by fixing the deviation of the robot's center on the centerline of the seedling row and rotating the robot. The angular deviation between the robot's centerline and the centerline of the seedling row changed from [15–15°°] to [−15°°]. When the centerline of the robot is parallel to the centerline of the seedling column, the angular deviation is defined as positive when the weeding robot is turned counterclockwise, i.e., the centerline of the weeding robot deviates to the left concerning the centerline of the seedling, the angular deviation is defined as unfavorable when the weeding robot is turned clockwise, i.e., the centerline of the weeding robot deviates to the right concerning the centerline of the seedling column. The measured and calculated values of angular deviation were recorded every 5° as a set of data. The experiment was repeated thrice for each group to improve the reliability of the experimental data, as shown in Figure 4. The mean error of the angular deviation was calculated to be 0.11°, and the standard deviation was 0.04°.

To verify the feasibility and reliability of the visual navigation and path tracking designed in this paper, the following scenes are set up: straight line, turning path, and obstacle occlusion. In the experiment, its forward speed is set as $V = 0.2$ m/s. The angular velocity w is in the

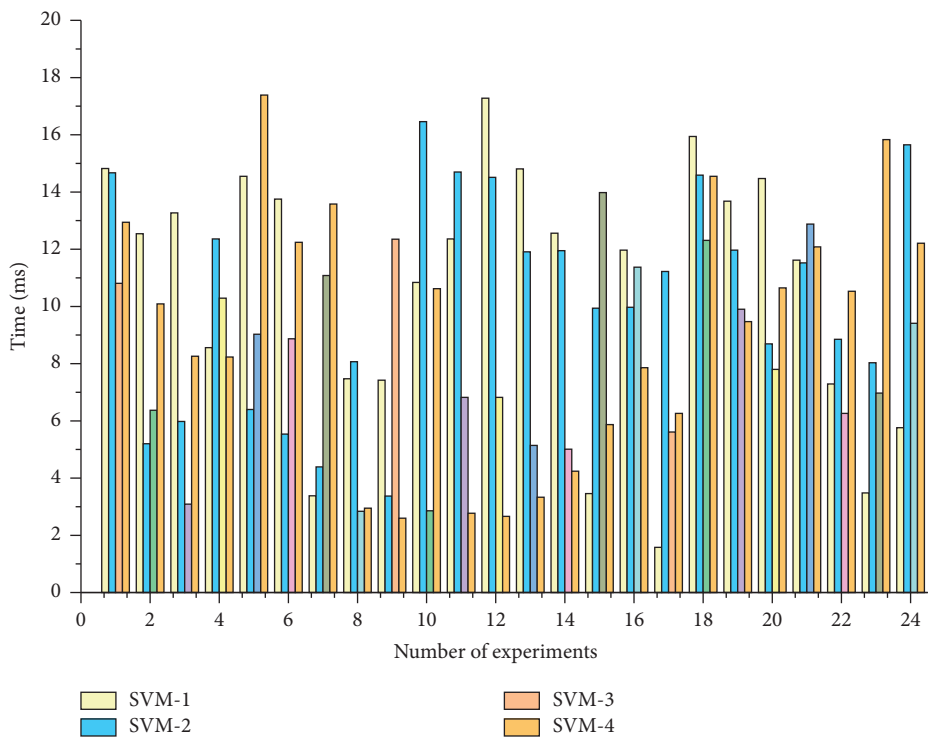


Figure 4: Experimental test comparison chart.

[−1.0 rad/s, 1.0 rad/s] range, in each frame after processing, the recognized center line is displayed with a straight blue line, and if a digital road sign is detected, the connection between the presighting point and the center of the digital road sign is displayed with a red straight line. The initial position of the Tourtellot robot is slight to the right of the path in the straight section at startup, and the identification line recognition results under several moments at intervals during operation. During the process of the robot, the angle of the centerline of the marking line obtained through image processing and recognition is plotted against the distance from the pre-sighting point to the centerline, as shown in Figure 5, where the dark blue curve is after the Kalman filtering. The light blue curve is the angle θ^{pre} without the Kalman filtering. It can be seen that θ the fluctuation is relatively stable during the robot's operation without significant abrupt changes. In contrast, the θ^{pre} instability of the unfiltered red curve in the figure is the distance deviation d after Kalman filtering, while the purple curve is the unfiltered distance deviation d^{pre}. The result of the distance deviation depends on the angle to a certain extent, so the fluctuation θ^{pre} leads to a more significant distance deviation from the pre-sighting point to the marker line also varies more.

The lane lines were laid in the lab, and the corresponding QR codes were set at each node location to place the mobile robot at any initial node location. First, we start the mobile robot management system and wireless communication server, set the relevant network parameters, establish the wireless LAN server, and wait for the lower computer to access. We start the embedded development board TX2, the

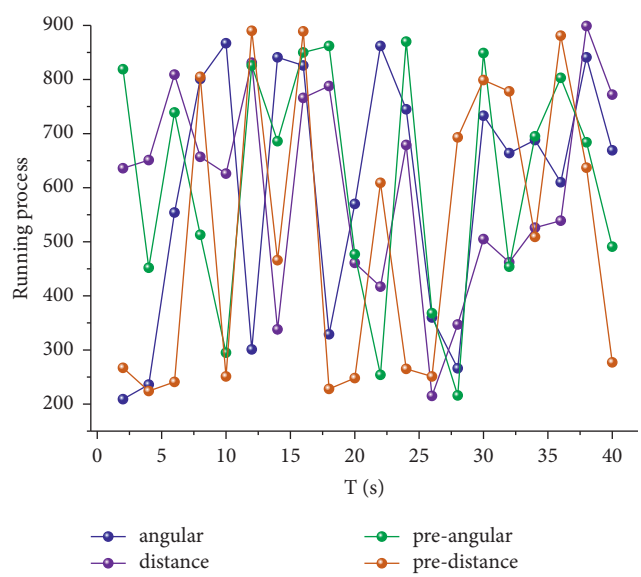


Figure 5: Running process identification lines θ and d curves.

industrial CCD camera, and the execution component Turtle Bot2, and enable the wireless network access function after checking and confirming the connection of each hardware of the lower computer, setting the network parameters, and accessing the LAN established by the wireless server; the upper computer reads the map file and reads the information to the vehicle controller according to the readings. The upper computer reads the map file, reads the data from the vehicle controller, initializes the coordinates of the node

where the mobile robot is located according to the read information, and displays its position in the map in real time; the management system will automatically number the tasks after receiving the task command, and then compare the current mobile robot location information with the target location information to get the shortest path information by path planning and transmit the path information to the lower computer through wireless communication. The path information will be transferred to the lower computer through wireless communication. After receiving the path information, the lower computer will execute the task and walk on the specified path through lane line tracking. The QR code recognition will be used for positioning correction and steering guidance, and the real-time data will be uploaded to the upper computer [20]. To verify the accuracy of the path tracking algorithm based on visual navigation, we randomly select a time point in the mobile robot operation test, extract the measured angle data of the gyroscope in the actuator at this time point, and compare it with the data obtained from the path tracking algorithm at the corresponding time. The deviation of the gyroscope angle data from the angle data obtained from image processing is calculated, and the relationship between the selected time nodes and the angle deviation is plotted as shown in Figure 6. From the results, the deviation value between the path tracking algorithm and the actual angle is ±1.5°, indicating that this paper's path tracking algorithm can achieve high accuracy.

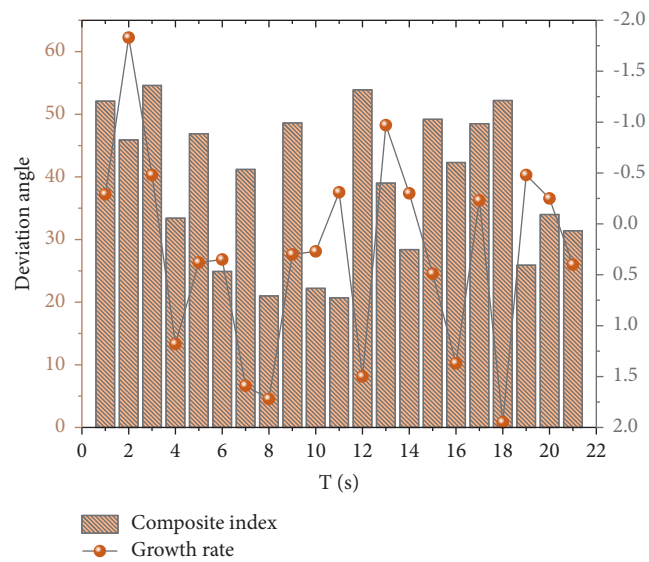


Figure 6: Deviation angle change.

4.2. Results of Target Image Recognition and Localization Analysis. The validation criteria for the measurement accuracy of visual navigation parameters were first established, followed by the analysis of the overall structure of the optical navigation detection system in the field and the experimental locations and experimental methods. Then, experiments on navigation line detection were conducted for the improved RANSAC algorithm and the enhanced probabilistic Hough transforms algorithm under different harvesting environments. The success rate of detection and image processing operation speed were statistically analyzed. Finally, experiments are conducted to verify the success rate of navigation line detection for the image pyramid optical flow tracking algorithm under different environments, and multiple sets of displacement deviation and angle deviation benchmark values are set to verify the measurement accuracy of navigation parameters according to the visual navigation parameter measurement accuracy verification standard. The measurement accuracy of angular deviation of navigation parameters of the image pyramid optical flow tracking algorithm is verified. The experiments on the error measurement of the angular variation of the navigation parameters of the image pyramid optical flow tracking algorithm were performed in 6 groups in turn, with the base value of the angular deviation being 0°. As shown in Figure 7, the average value of the maximum error of the actual measurement value of the angular deviation of the harvester was 10.57°, the average value of the mean error was 3.73°, and the average value of the standard deviation was 2.98°. In the actual detection process, the factors affecting the angle deviation measurement error are mainly: due to the influence of factors such as the influence of driving vision, the manually operated intelligent rice, and wheat harvester cannot always keep the straight line where the left-hand divider is located parallel to the harvest navigation line, i.e., the actual reference value floats around 0°; the incomplete segmentation of harvested and unharvested areas in the image causes the detected navigation line and the straight line where the left-hand divider is located to additional angle exists.

The training process of the detection algorithm includes data set preparation, formal training, and network generalization error evaluation. Then, the deep learning detection algorithm is compared with the traditional SVM algorithm for classification experiments. The results surface that although the deep learning detection check-all rate is slightly lower than the SVM algorithm, its check-accuracy rate and accuracy rate are higher than the SVM algorithm. And the detection rate of the deep learning detection algorithm has a more significant advantage, which can meet the basic requirements of real-time QR code detection for mobile robots. Finally, the operation experiment of the whole system is carried out to verify the feasibility. Moreover, comparing the gyroscope angle data with the path tracking algorithm data can be obtained. Various experiments were designed according to different given algorithm parameters to influence the environment model rasterization resolution parameter on the confidence occupancy meter. The test image size varies. The algorithm automatically transforms the original test map to 448×448 resolution size before input. The output result of the positioning frame will also be altered accordingly. The partial detection results of the deep learning detection algorithm are shown in Figure 8. The results show that the deep learning detection algorithm performs well for QR codes with simple backgrounds, complex backgrounds, vestiges, small deformations, partial occlusions, or multiple QR codes in a single image.

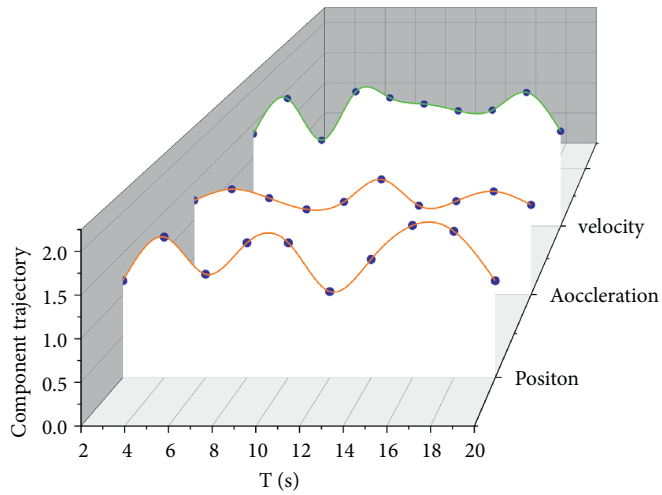


Figure 7: Component trajectories of the rotational path, velocity, and acceleration.

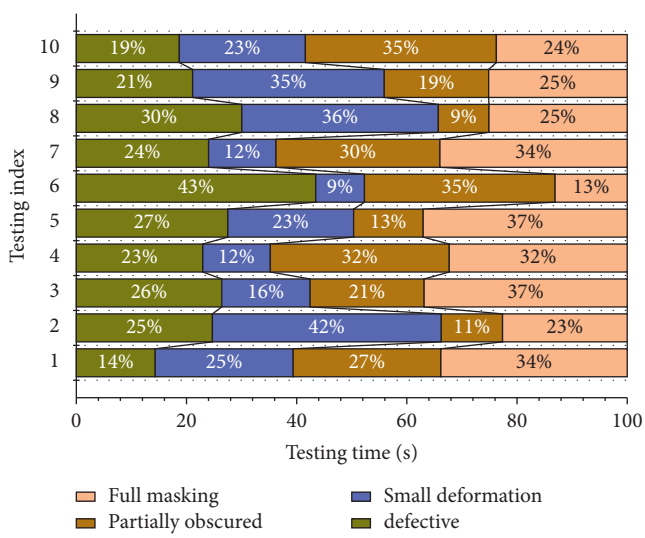


Figure 8: Comparison results of testing indicators.

Among the detection results, the deep learning detection algorithm can make accurate judgments on the images of the experimental scenes of the mobile robot in this paper. The accuracy and rapidity of the integrated detection algorithm can meet the basic requirements of real-time detection of QR codes during the execution of the mobile robot tasks in this paper's experiments. Moreover, the corresponding QR codes are set at each node location, and the mobile robot is placed at any initial node location. First, we start the mobile robot management system and wireless communication server, set the relevant network parameters and establish the wireless LAN server, and wait for the lower computer to access; then, we start the embedded development board TX2, the industrial CCD camera, and the execution component Turtle Bot2 and enable the wireless network access function after each hardware connection of the lower computer is checked and confirmed, set the network parameters and access the LAN established by the wireless server; the upper computer reads the QR code and reads the QR code. The upper computer reads the map file, reads the information from the vehicle controller, initializes the coordinates of the node where the mobile robot is located according to the read information, and displays its position in the map in real time. The gyroscope is corrected once after the QR code, so it is assumed that the gyroscope data is the actual angle. The deviation of the gyroscope angle data from the angle data obtained from image processing is calculated, and the relationship between the selected time nodes and the angle deviation is plotted. From the results, the variation of the path tracking algorithm from the actual angle is ±1.5°, which indicates that the path tracking algorithm in this paper can achieve high accuracy.

5. Conclusion

With the continuous development and progress of science and technology, the intelligent level requirements are gradually increasing, and the automatic control of mobile robots has become an important direction in the development of robot systems, whose visual navigation system is one of the hotspots of research today. In this paper, an in-depth study is conducted on the problem of robot optical navigation path detection. For the problem that the RANSAC algorithm first establishes the linear mathematical model of the path in the navigation line detection and then performs the remaining corner point model verification, which leads to more iterations of the algorithm and more considerable computation, model verification criteria are added to avoid the problem of time-consuming and detection errors caused by continuing iterative verification in the case of model errors. By limiting the range of edge point probability extraction and setting the success criterion of straight line detection, the improved Hough transform algorithm can effectively solve the problem of fast and accurate identification of navigation lines caused by the probability extraction of edge points in the whole detection area; finally, the image pyramid optical flow tracking algorithm is used to realize the tracking detection of robot visual navigation and tracking measurement of visual navigation parameters. The

robot is experimented with setting straight, turning, and obstacle occlusion scenes. From the trajectory graphs obtained from the experimental simulation, the robot can better fit the marking line operation during the operation. The charts of parameters and d show that the bit-posture relationship of the robot body relative to the marking line is closed during the process, which also illustrates the effectiveness and accuracy of target image recognition and localization.

Based on the research work carried out in this paper, we briefly analyze the potential research points with further depth in this study, taking into account the current trends in computer vision and robotics.

The introduction of deep neural network descriptors as the front-end data matching for VSLAM to achieve visual data matching with illumination and viewpoint invariance is a current research hotspot and trend in VSLAM-related work. A considerable amount of work has been carried out in this area, including the extraction of intermediate descriptors using existing network models and the design of a network structure of visual feature descriptors specifically for VSLAM. However, the gap between the current research in this area and the application of VSLAM lies in the visual feature point extraction and the real-time nature of the generation. In addition, deep neural network descriptors tend to have higher dimensionality and take longer to perform feature point matching and distance calculation. Therefore, it is not easy to guarantee the online performance of deep vision descriptors for system applications with high real-time requirements such as VSLAM. Therefore, there is work to be done on a downscale and speed up the depth vision descriptors to meet the demand of VSLAM real-time applications.

Data Availability

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

Conflicts of Interest

The author declares no conflicts of interest.

Acknowledgments

This work was supported by the 2021 school level scientific research and innovation team project, Innovation and Entrepreneurship Education Scientific Research and Innovation Team, (No. HNACKT-2021-01) and 2021 Research Projects of Educational Science (No. HNACJY-2021-15).

References

[1] C. Sampedro, A. Rodriguez-Ramos, H. Bavle, A. Carrio, P. de la Puente, and P. Campoy, "A fully-autonomous aerial robot for search and rescue applications in indoor environments using learning-based techniques," *Journal of Intelligent and Robotic Systems*, vol. 95, no. 2, pp. 601–627, 2019.

[2] S. Wang, F. Jiang, B. Zhang, R. Ma, and Q. Hao, "Development of UAV-based target tracking and recognition systems," *IEEE Transactions on Intelligent Transportation Systems*, vol. 21, no. 8, pp. 3409–3422, 2019.

[3] L. Qiu, C. Li, and H. Ren, "Real-time surgical instrument tracking in robot-assisted surgery using multi-domain convolutional neural network," *Healthcare Technology Letters*, vol. 6, no. 6, pp. 159–164, 2019.

[4] A. Devo, G. Mezzetti, G. Costante, M. L. Fravolini, and P. Valigi, "Towards generalization in target-driven visual navigation by using deep reinforcement learning," *IEEE Transactions on Robotics*, vol. 36, no. 5, pp. 1546–1561, 2020.

[5] Y. Xiong, Y. Ge, L. Grimstad, and P. J. From, "An autonomous strawberry-harvesting robot: design, development, integration, and field evaluation," *Journal of Field Robotics*, vol. 37, no. 2, pp. 202–224, 2020.

[6] P. Neubert, S. Schubert, and P. Protzel, "A neurologically inspired sequence processing model for mobile robot place recognition," *IEEE Robotics and Automation Letters*, vol. 4, no. 4, pp. 3200–3207, 2019.

[7] W. J. Heerink, S. J. S. Ruiter, J. P. Pennings et al., "Robotic versus freehand needle positioning in CT-guided ablation of liver tumors: a randomized controlled trial," *Radiology*, vol. 290, no. 3, pp. 826–832, 2019.

[8] S. G. Mathisen, F. S. Leira, H. H. Helgesen, K. Gryte, and T. A. Johansen, "Autonomous ballistic airdrop of objects from a small fixed-wing unmanned aerial vehicle," *Autonomous Robots*, vol. 44, no. 5, pp. 859–875, 2020.

[9] K. M. Abughalieh, B. H. Sababha, and N. A. Rawashdeh, "A video-based object detection and tracking system for weight sensitive UAVs," *Multimedia Tools and Applications*, vol. 78, no. 7, pp. 9149–9167, 2019.

[10] K. Lee, J. Gibson, and E. A. Theodorou, "Aggressive perception-aware navigation using deep optical flow dynamics and PixelMPC," *IEEE Robotics and Automation Letters*, vol. 5, no. 2, pp. 1207–1214, 2020.

[11] C. H. G. Li and Y. M. Chang, "Automated visual positioning and precision placement of a workpiece using deep learning," *The International Journal of Advanced Manufacturing Technology*, vol. 104, no. 9, pp. 4527–4538, 2019.

[12] D. Fielding and M. Oki, "Technologies for targeting the peripheral pulmonary nodule including robotics," *Respirology*, vol. 25, no. 9, pp. 914–923, 2020.

[13] P. M. Kumar, U. Gandhi, R. Varatharajan, G. Manogaran, R. Jidhesh, and T. Vadivel, "Intelligent face recognition and navigation system using neural learning for smart security in internet of things," *Cluster Computing*, vol. 22, no. 4, pp. 7733–7744, 2019.

[14] V. Vasilopoulos, G. Pavlakos, S. L. Bowman et al., "Reactive semantic planning in unexplored semantic environments using deep perceptual feedback," *IEEE Robotics and Automation Letters*, vol. 5, no. 3, pp. 4455–4462, 2020.

[15] V. Vasilopoulos, G. Pavlakos, S. L. Bowman et al., "Reactive semantic planning in unexplored semantic environments using deep perceptual feedback," *IEEE Robotics and Automation Letters*, vol. 5, no. 3, pp. 4455–4462, 2020.

[16] J. W. Martin, B. Scaglioni, J. C. Norton et al., "Enabling the future of colonoscopy with intelligent and autonomous magnetic manipulation," *Nature machine intelligence*, vol. 2, no. 10, pp. 595–606, 2020.

[17] M. Ma, H. Li, X. Gao et al., "Target orientation detection based on a neural network with a bionic bee-like compound eye," *Optics Express*, vol. 28, no. 8, pp. 10794–10805, 2020.

[18] J. Yang, C. Wang, B. Jiang, H. Song, and Q. Meng, "Visual perception enabled industry intelligence: state of the art,

challenges and prospects," *IEEE Transactions on Industrial Informatics*, vol. 17, no. 3, pp. 2204–2219, 2020.

[19] W.-H. Su, "Advanced machine learning in point spectroscopy, RGB- and hyperspectral-imaging for automatic discriminations of crops and weeds: a review," *Smart Cities*, vol. 3, no. 3, pp. 767–792, 2020.

[20] A. A. Zhilenkov, S. G. Chernyi, S. S. Sokolov, and A. P. Nyrkov, "Intelligent autonomous navigation system for UAV in randomly changing environmental conditions," *Journal of Intelligent and Fuzzy Systems*, vol. 38, no. 5, pp. 6619–6625, 2020.

Hindawi
Journal of Robotics
Volume 2022, Article ID 2183229, 9 pages
https://doi.org/10.1155/2022/2183229

Research Article

Mobile Robot Path Planning Based on Enhanced Dynamic Window Approach and Improved A* Algorithm

Hongxia Yang and Xingqiang Teng

College of Business, Yantai Nanshan University, Shandong, Longkou 265701, China

Correspondence should be addressed to Hongxia Yang; yanghongxia1@nanshan.edu.cn

Received 28 January 2022; Accepted 5 March 2022; Published 22 March 2022

Academic Editor: Shan Zhong

Path planning is one of the most popular researches on mobile robots, and it is the key technology to realize autonomous navigation of robots. Aiming at the problem that the mobile robot may collide or fail along the planned path in an environment with random obstacles, a robot path planning scheme that combines the improved A* algorithm with an enhanced dynamic window method is proposed. In the improved A* algorithm, in order to improve the algorithm efficiency, so that a single planning path can pass through multiple target points, the search point selection strategy and evaluation function are optimized. In order to achieve local obstacle avoidance and pursuit of dynamic target points in dynamic and complex environments, an online path planning method combining enhanced dynamic window algorithm and global path planning information is proposed. The preview deviation angle tracking method is used to successfully capture moving target points. It also improves the efficiency of path planning and ensures that on the basis of the global optimal path, the random obstacle can be avoided in real time so that the robot can reach the target point smoothly. The simulation results show that compared with other methods, the proposed method achieves excellent global and local path planning performance, the planned trajectory is smoother, and the search efficiency is higher in complex environments.

1. Introduction

Mobile robot is an important branch of robotics research [1]. With the continuous development of the world economy and technology, mobile robots appear more and more frequently in scientific research, production environment, and daily life [2]. At the same time, with the development of computers and control technology, the application fields of mobile robots are becoming more and more extensive [3].

Mobile robot needs to plan a path from the initial position to the target position in the work scene. The path should meet a series of requirements such as short length, high efficiency, and high safety, and it must be able to avoid static and dynamic obstacles along the way [4]. At the same time, mobile robots should have certain computing capabilities to calculate the shortest and safest route in real time to save time and reserve energy [5].

According to the characteristics of mobile robot path planning, it can be divided into the intelligent search algorithm, artificial intelligence algorithm, geometric model algorithm, and local obstacle avoidance algorithm [6]. The intelligent search algorithm uses randomly generated initial solutions or sampling points to approach the optimal solution through multiple iterations. The biggest characteristic of this type of algorithm is randomness, so its solution is not unique [7]. Algorithms based on artificial intelligence include methods such as Q learning and deep learning, but such methods require a large number of different types of training samples, which limit their practicality in the real world [8].

The geometric model-based path planning method is to construct a geometric model on the basis of known environment, then select an appropriate path, and adjust the feasible solution based on the optimal strategy in real-time [9]. The paths obtained by such methods are all nonsmooth

paths, so optimization is needed to achieve the smooth turning of the mobile robot. Such methods include A* algorithm, Voronoi graph, and so on. In the global path planning, the A* algorithm will select the square with the lowest current cost value until the search reaches the end, thereby planning the path with the lowest total costs. However, the traditional A* algorithm suffers from poor global optimization capability, and there are too many redundant points and inflection points in the planned path [10].

The local obstacle avoidance algorithm is used to enhance the obstacle avoidance ability of mobile robots and improve safety. It aims to move the robot away from obstacles and plan a safe collision-free path. Commonly used local obstacle avoidance algorithms include artificial potential field method, dynamic window approach (DWA), and so on [11]. The DWA is a method of sampling the surroundings at the current moment to obtain the robot's action state at the next moment. This method can quickly reach the target point while avoiding collisions between the robot and obstacles in the search space. However, it is highly dependent on global parameters and is prone to failure in unknown environments [12].

In order to solve the above problems, this paper proposes a random obstacle avoidance method for mobile robots that combines the improved A* algorithm with an enhanced dynamic window approach (EDWA), which takes into account the smoothness of the planned path and the search efficiency of the A* algorithm. Then based on the global planning path information and dynamic environment changes, the EDWA is used to complete the local path planning, so as to make up for the shortcomings of the traditional A* algorithm with poor timeliness. The proposed method pursues dynamic obstacles on the basis of global path planning information, thereby planning the optimal path and improving real-time performance and flexibility.

The rest of this paper is organized as follows: the second part is the related work, which introduces the existing advanced robot path planning methods. The third part is the improved A* algorithm. The fourth part is the enhanced version of the dynamic window method. The fifth part is the experimental part. The sixth part is the conclusion.

2. Related Methods

The A* algorithm has the advantages of short planned paths and fast calculation speed, and it is a widely used global path planning method. Yuan and Han [13] successfully implement the improved A* algorithm to achieve robot path planning in a static environment. Yu et al. [14] use an improved Dijkstra algorithm to complete global path planning, but its local path search speed is very slow, and a smooth and safe optimal path cannot be quickly planned. Kaplan et al. [15] introduce an improved multiorder B-spline curve for path smoothing, but this method cannot complete the smoothing of the dynamic path, and the running process is cumbersome. The efficiency of the algorithm is poor. When the environment changes or the robot is affected by other external factors, the established path will fail.

Artificial potential field method and DWA are currently widely used local path planning methods. Aiming at the problem that the robot is regarded as a single point in most planning, which leads to the problem of being unable to pass the narrowband, Kiss and Tevesz [16] propose a global dynamic window navigation scheme based on model predictive control in which the weighted objective function is omitted. Aiming at the energy consumption problem in the path planning process, Henkel et al. [17] propose an efficient local path planner suitable for omnidirectional mobile robot navigation in a dynamic environment. The experimental results show that compared with the DWA algorithm, the energy consumption is reduced by 9.79%.

In practical applications, robots often encounter unexpected obstacles such as unknown obstacles during the planning process, which may cause the robots to collide. Therefore, many methods combine the strong obstacle avoidance ability of local path planning algorithms with global path planning methods to improve performance. Aiming at the problem that the classic A* algorithm cannot be applied in a dynamic environment, Zhong et al. [18] propose a hybrid method combining improved A* algorithm and DWA algorithm. Experimental results show that the path planned by the improved algorithm is smoother and more efficient. But the disadvantage of this method is that it is only applicable to target points in the same direction. Wu et al. [19] propose a hybrid algorithm combining beetle antenna search algorithm and artificial potential field algorithm for the dynamic path planning problem. The experimental results verify the effectiveness and superiority of the method. The method proposed by Kashyap et al. [20] combines the DWA algorithm and teaching and learning optimization. Path planning and obstacle avoidance can be successfully implemented in both the dynamic terrain and the dynamic terrain.

For the path planning problem of mobile robots in complex environments, Ma et al. [21] propose a method combining improved A* algorithm, minimum snap trajectory generation, and the timed elastic band, which improves the efficiency and flexibility of mobile robot path planning. The experimental results show that this method can shorten the length of the planned path, reduce the total turning angle, and effectively complete the global optimal path planning and local real-time obstacle avoidance.

3. Improved A* Algorithm

3.1. Traditional A Algorithm.* The A* algorithm is a heuristic search algorithm that obtains the planned path of the mobile robot by continuously searching for the path approaching the destination. This method is simple and fast, and the heuristic search is very targeted. It only needs to search a part of the state space of the problem, so it can achieve the purpose of narrowing the search range and reducing the complexity of the problem, and the path finding efficiency is very high.

The A* algorithm can be expressed as follows:

$$f(n) = g(n) + h(n), \tag{1}$$

where n is the current node, $f(n)$ is the cost value of the optimal path from the starting node to the target point, and $g(n)$ is the depth factor, which is the cost of the optimal path from the starting node to node n found by the heuristic search algorithm. $h(n)$ is an estimated cost value of the optimal path from node n to the target node. The value of each decision is evaluated according to the evaluation function, and it is decided which scheme to try first in order to obtain the shortest path.

3.2. Improved A Algorithm.* The traditional A* algorithm has the following two shortcomings: (1) there are too many searching nodes, which affect the computational efficiency and (2) the path obtained by the A* algorithm has redundant collinear nodes and redundant turning points. The robot walking along such a path will affect its movement continuity.

In view of the above problems, this paper proposes two improvements to the A* algorithm: (1) in order to improve the searching efficiency, an optimized search point selection strategy is proposed, and the heuristic function is improved; (2) in order to further optimize the path, the redundant point deletion strategy is introduced. Redundant nodes are eliminated so that the final path contains only the starting point, end point, and key turning points.

3.2.1. Optimization of Search Point Selection Strategy. The traditional A* algorithm expands the eight neighborhood grids of the current node during path planning. As shown in Figure 1, the pink grid is the current point, and n_1 to n_8 are the eight directions in which the current grid can move. Too many options of the orientation for the target point will cause unnecessary grids in the eight-neighbor grids, resulting in a waste of computing time and storage space.

In order to further improve the searching efficiency, an optimized search point selection strategy is introduced. According to the relative position of the target point and the current point, three search directions are discarded, and five search directions are retained. Let the angle between the segment from the target point to the current point and the direction of n_2 as ε, then the corresponding relationship between the angle ε and the 3 discarded directions is shown in Table 1. After the above optimization, the number of search nodes can be reduced, and the computational efficiency can be improved.

3.2.2. Heuristic Function of the Optimization Algorithm. The traditional A* algorithm will traverse many unnecessary nodes during the path searching process, which affects the searching efficiency. The key factor causing this problem is the heuristic function design of the A* algorithm. If the estimated value of the heuristic function is less than the actual cost value, too many search nodes will be generated, reducing the computational efficiency, though the optimal path can be obtained. However, if the estimated value of the heuristic function is greater than the actual cost value, a relatively smaller number of searching nodes will be generated, which will improve efficiency, but it is difficult to obtain an optimal path. When the estimated value of the heuristic function is equal to the actual cost value, the highest search efficiency can be achieved. Since the heuristic function adopts Euclidean distance, the value of the heuristic function will always be less than or equal to the actual distance from the current point to the target point. When the current point is far from the target point, the estimated value of the heuristic function is much smaller than the actual value, which will cause the algorithm to search for more nodes and reduce efficiency. At this time, the weight of the estimated value should be increased to improve efficiency. When the current point is gradually approaching the target point, the estimated value gradually approaches the actual value. In order to prevent the estimated value from being too large and to avoid failure in searching for the optimal path, the estimated value weight should be reduced accordingly. In summary, we propose to improve the cost function as follows:

$$f(n) = g(n) + \left(1 + \frac{r}{R}\right)h(n), \tag{2}$$

where $f(n)$ is the integrated cost value, $h(n)$ is the estimated cost value from the current point to the target point, $g(n)$ is the actual cost value from the starting point to the current point, R is the distance from the starting point to the target point, and r is the distance between the current point and the target point.

3.2.3. Multiobjective Optimization. The traditional A* algorithm is only suitable for searching a single target point, and it has the shortcomings of low searching efficiency and

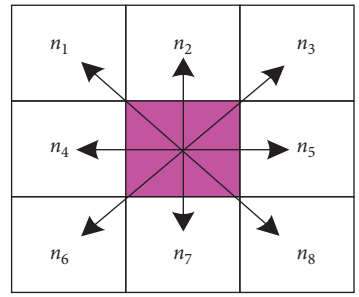


Figure 1: Schematic plot of the robot movement.

Table 1: The relationship between angle ε and the discarded directions.

ε	Retained directions	Discarded directions
[337.5°, 360°)∪ [0°, 22.5°)	$n_1n_2n_3n_4n_5$	$n_6n_7n_8$
[22.5°, 67.5°)	$n_1n_2n_3n_5n_8$	$n_4n_6n_7$
[67.5°, 112.5°)	$n_2n_3n_5n_7n_8$	$n_6n_7n_8$
[112.5°, 157.5°)	$n_3n_5n_6n_7n_8$	$n_1n_4n_6$
[157.5°, 202.5°)	$n_4n_5n_6n_7n_8$	$n_1n_2n_4$
[202.5°, 247.5°)	$n_1n_4n_6n_7n_8$	$n_2n_3n_5$
[247.5°, 292.5°)	$n_1n_2n_4n_6n_7$	$n_3n_5n_8$
[292.5°, 337.5°)	$n_1n_2n_3n_4n_6$	$n_5n_7n_8$

poor compatibility. Nowadays, the work situation is complicated, and the tracking of multiple target points in one planning can greatly improve efficiency. Therefore, the proposed method first samples and compares the location information of the target points and then improves the searching algorithm to consider multiple target points in one planning.

The priority of different target points is an important prerequisite for multitarget path planning. The Euclidean distance is used as the cost function of selecting the optimal target point. The lower the selection cost, the higher the priority level. Specific steps are as follows:

(1) Calculate the cost function of the target points T_i, $i = 1, 2, \ldots, n$; (2) compare the priority levels of the target points T_i; (3) preferentially plan the path for the target point T_i with the highest priority; (4) take the current target point T_{i-1} as the starting point of the mobile robot and plan the path of the secondary target point T_i; and (5) judge whether to reach the n-th target point. If it arrives, stop searching; otherwise, go back to the last step.

4. Enhanced Dynamic Window Approach

In this paper, an enhanced dynamic window algorithm (EDWA) is proposed, which combines global path information to avoid obstacles and pursue dynamic targets. Even part of the environment information is unknown, the mobile robot can still perform local path planning, avoid obstacles, and reach the designated target point.

4.1. Kinematics Model of Mobile Robot. DWA mainly samples the speed space of the mobile robot in the window area and simulates the feasible motion trajectory within time t in the speed space (v_t, ω_t). The changes of the linear velocity v_t and the angular velocity ω_t in the velocity space represent the motion states of the mobile robot. Among all feasible trajectories, the optimal trajectory is obtained with the evaluation function. Therefore, assuming that the mobile robot moves within the time interval Δ_t, the kinematics model can be formulated as follows [22]:

$$\begin{cases} x = x + v_x \Delta t \cos\theta_t - v_y \Delta t \sin\theta_t, \\ y = y + v_x \Delta t \sin\theta_t + v_y \Delta t \cos\theta_t, \\ \theta_t = \theta_t + \omega_t \Delta t. \end{cases} \tag{3}$$

4.2. Speed Sampling of Mobile Robot. There are an infinite number of groups in the velocity space, but in actual operation, it is necessary to constrain the sampling velocity range based on the constraints of the mobile robot as well as environmental constraints.

The speed constraints of mobile robots can be expressed as follows:

$$V_m = \{(v, \omega) | v \in [v_{\min}, v_{\max}]\}, \quad \omega \in [\omega_{\min}, \omega_{\max}]\}. \tag{4}$$

In the moving time interval of the dynamic window, the speed constraint of the mobile robot caused by the addition and subtraction constraints of the motor can be calculated as follows:

$$\begin{aligned} V_d = \{ ((v, \omega) | v \in [v_c - \dot{v}_b \Delta t, v_c + \dot{v}_a \Delta t], \\ \omega \in [\omega_c - \dot{\omega}_b \Delta t, \omega_c + \dot{\omega}_a \Delta t]\}, \end{aligned} \tag{5}$$

where v_c and ω_c denote current speed, $\dot{v}_a$ and $\dot{\omega}_a$ denote the maximum acceleration of the mobile robot, and $\dot{v}_b$ and $\dot{\omega}_b$ denote the maximum deceleration of the mobile robot.

Mobile robot braking distance constraint: when avoiding obstacles in a local environment, the safety of the mobile robot needs to be ensured. Under the constraint of maximum deceleration, the speed can be reduced to 0 m/s before the impact. The braking constraint is expressed as follows:

$$V_d = \left\{(v, \omega) | v \le \left(2d(v, \omega)\dot{v}_b\right)^{1/2}, \quad \omega \le \left(2d(v, \omega)\dot{\omega}_b\right)^{1/2}\right\}, \tag{6}$$

where $d(v, \omega)$ represents the shortest distance between the trajectory (v, ω) and the obstacle.

4.3. Evaluation Function Considering Global Path Information. After simulating several trajectories based on the range of motion speed and the robot motion model, the optimal trajectory is obtained through the evaluation function. The traditional DWA often falls into the local optimum. In order to fix this problem, an evaluation function considering the global path information is proposed to ensure that the final local path is in accordance with the global optimal path. The improved evaluation function is as follows:

$$G(v, \omega) = \sigma\left(\alpha PH(v, \omega) + \beta d(v, \omega) + \gamma v_e(v, \omega)\right), \tag{7}$$

where $PH(v, \omega)$ is the angle difference between the directions of the trajectory end points and the current target point, $d(v, \omega)$ is the distance between the trajectory and the nearest obstacle, $v_e(v, \omega)$ is the current speed evaluation function, σ is the smoothing function, and α, β, and λ are weighting coefficients. With the evaluation function, the path planned by the robot can avoid random obstacles and fit the global optimal path.

4.4. Dynamic Target Pursuing with Preview Deviation Angle Algorithm. In the current complex working environment, it is necessary for mobile robots to realize online path planning while avoiding obstacles and complete the task of pursuing dynamic targets as quickly as possible. In the process of pursuing the target, not only the dynamic path planning problem needs to be solved, but also the mobile robot needs to use its own sensors to continuously adjust the yaw angle and combine the weighted obstacle search algorithm to improve the chasing efficiency.

During the operation, the position information of the target point is constantly updated with the robot sensors; the center point of the target is regarded as the preview point; and the distance between the center of the mobile robot and the center of the target is the preview distance. The angle between the moving directions of the target point and the

mobile robot is the preview angle. The schematic diagram of chasing strategy based on the preview deviation angle is shown in Figure 2. In Figure 2, L is the preview distance, and M_d is the vertical distance between the mobile robot and the dynamic target point trajectory. Given $L \neq 0$ and $M_d \neq 0$, then M_d is always less than L, θ_1 is the preview deviation angle, and it is stipulated that the preview deviation angle is positive when the preview point is in the frost-left of the moving direction of the mobile robot, and the center of the moving target point is the preview point.

It can be seen from Figure 2 that when $\theta_1 \longrightarrow 0$ and $L \longrightarrow 0$, the mobile robot can catch up with the moving target, so we get:

$$\phi_1 = \arctan \frac{M_d}{N} - \theta_1. \tag{8}$$

In the case of $L \neq 0$ and $M_d = 0$, it can be seen from Figure 2 that the robot is always behind the target point during the pursuit; we get $N = L$, $\arctan(M_d/N) = 0$, then $\theta_1 = 0$, and $\phi_1 = 0$; and the robot accelerates in a straight line to chase the target point. In the case of $L \neq 0$ and $M_d \neq 0$, it can be derived that during the pursuing process $N \neq 0$, $\arctan(M_d/N) \longrightarrow 0$ and $\theta_1 \longrightarrow 0$, so $\phi_1 \longrightarrow 0$. In Figure 2, if an arc is drawn with O as the center and tangent to the trajectory of the moving target, then the arc is the planned trajectory, and the arc radius R can be calculated as follows:

$$R = \frac{L^2}{2M_d}. \tag{9}$$

From the sensor readings of the mobile robot, the coordinates of the moving target point (x_g, y_g) and the mobile robot (x, y) can be measured, then the coordinates of the center O (x_r, y_r) can be obtained. The center O that is closer to the mobile robot will be selected:

$$\begin{cases} -x + \tan\theta_g \cdot y - b \cdot \tan\theta_g = 0, \\ \dfrac{\left|1 - x_r + \tan\theta_g \cdot y_r - b\theta_g \tan\theta_g\right|}{\sqrt{1 + \tan^2\theta_g}} = R. \end{cases} \tag{10}$$

Then the arc trajectory is expressed as follows:

$$(x - x_r)^2 + (y - y_r)^2 = R^2. \tag{11}$$

5. Experiment

In order to verify the global and local path planning performance of the proposed method, we compare the proposed hybrid method with other methods in different environments. The experiment runs on a computer with a CPU of I5-9400 and a RAM of 16 GB, and the algorithm is implemented through MATLAB programming simulation. The parameters are set as follows: the maximum velocity is 1 m/s, the maximum angular velocity is 20°/s, the velocity resolution is 0.01 m/s, the angular velocity resolution is 1°/s, the acceleration is 0.2 m/s^2, and the angular acceleration is 50°/s^2. The parameters of the evaluation function are: $\alpha = 0.1$, $\beta = 0.05$, and $\gamma = 0.2$.

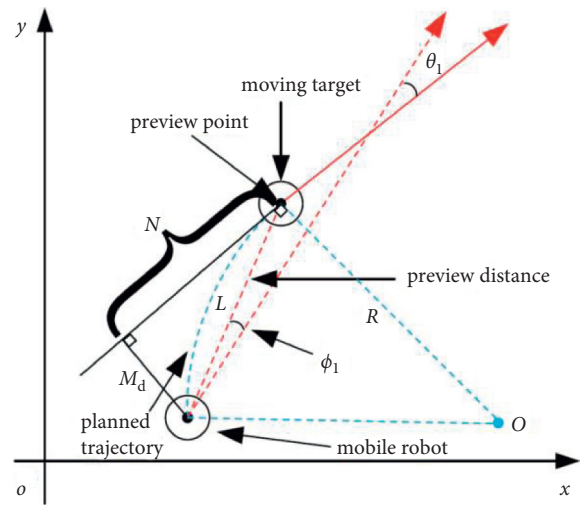


Figure 2: Preview deviation angle tracking.

5.1. Map and Environment Construction. In order to facilitate performance analysis, this article uses the following settings for mobile robots and environmental maps:

(1) The grid method is used to mathematically model the environment map, and the feasible and infeasible areas are represented by white grids and black grids, respectively.

(2) The starting point and ending point of the map are given, with some unknown obstacles in the map.

(3) The mobile robot and dynamic obstacles are represented as dilated points.

(4) The dynamic obstacle in the map moves in a straight line at a uniform speed, but the moving direction and position of the obstacle are unknown.

(5) The mobile robot is equipped with sensors, which can perceive the information of the map within a limited range, such as the position and speed of obstacles.

(6) The mobile robot moves at a constant speed and can move in all directions.

5.2. Trajectory Comparison. On the static global path planning problem, compared with the traditional A* algorithm, the proposed algorithm mainly optimizes the planned path trajectory. In a 20×20 grid environment, global path planning is performed with (1, 1) as the starting point and (20, 20) as the ending point, and the performance of the proposed method is compared with the classic A* algorithm [18]. The reference standards for optimizing the trajectory include the cumulative turning angles, the average turning angles, and the path length. Figure 3 shows the optimization process of path nodes, and the red line represents the generated global path. Figure 3(a) is the path trajectory generated by the traditional A* algorithm, Figure 3(b) is the path trajectory generated with [18], and Figure 3(c) is the path trajectory generated by the proposed method. The statistics are shown in Table 2.

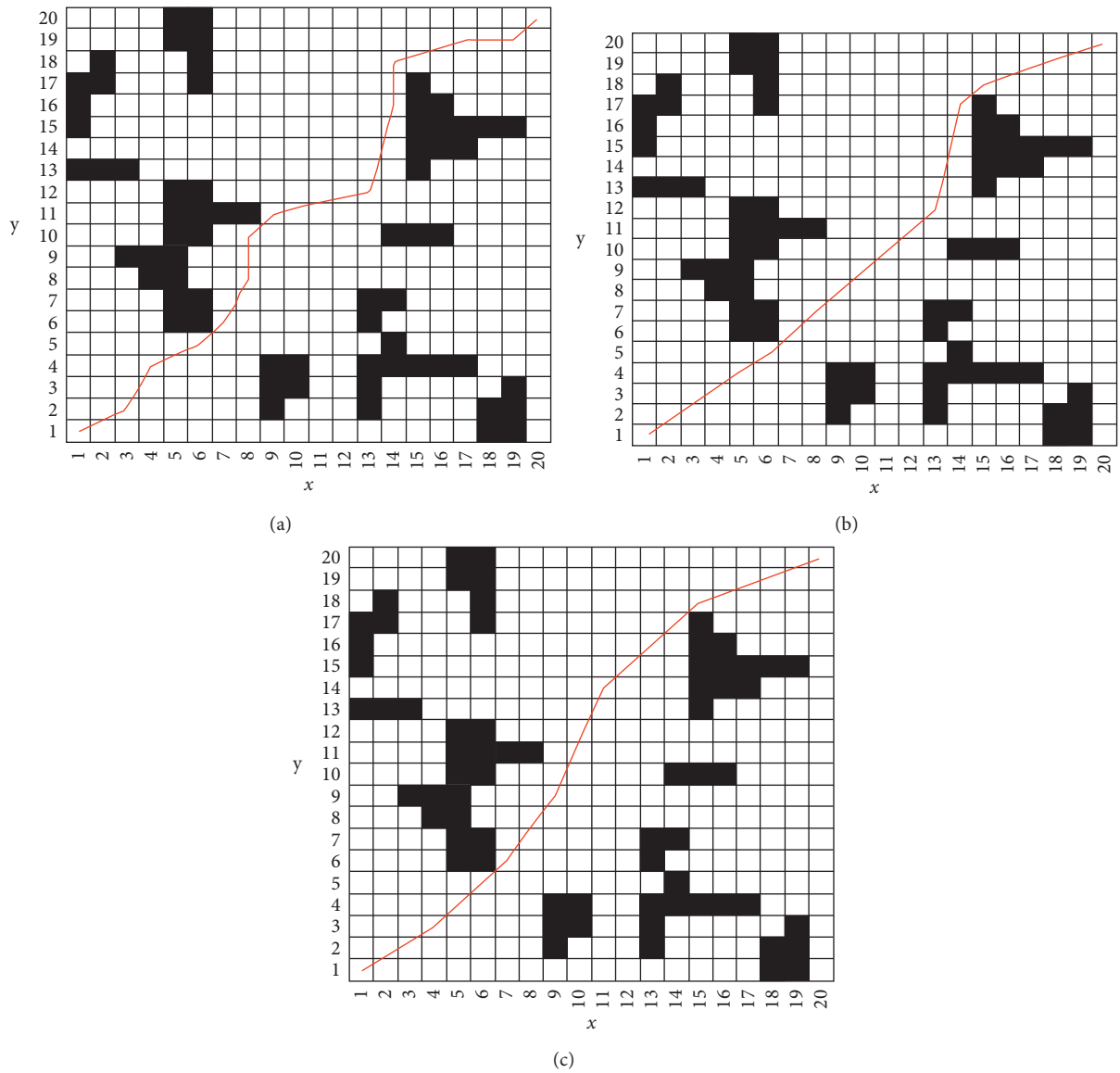


Figure 3: Path planning results. (a) The path trajectory generated by the traditional A* algorithm. (b) The path trajectory generated with [18]. (c) The path trajectory generated by the proposed method.

Table 2: Planned paths with different methods.

Method	Nodes	Cumulative turning angles (°)	Average turning angles (°)	Path length
Classic A*	15	488.49	32.56	33.77
[18]	5	128.92	25.78	25.89
Proposed method	6	85.41	14.24	26.05

It can be seen from Table 2 that compared with the traditional A* algorithm, the proposed method has significantly reduced the cumulative turning angles, average turning angles, and path length through path optimization strategy. The method in [18] only deletes redundant points for the trajectory. It can be seen from Table 2 that although the proposed algorithm has a slight increase in path length compared to [18], the cumulative turning angle and average turning angle have been reduced, making the planned trajectory smoother and more in line with the kinematics law of the mobile robot. In addition, the multiobjective strategy adopted in this paper shortens the average distance between nodes, which is more conducive to the selection of the subsequent local path planning range.

5.3. Comparison of Algorithm Efficiency. In order to alleviate the impact of environmental changes on the experiments and further improve the evaluation fairness, different data are used to simulate different algorithms on the same

Table 3: Comparison under different environments.

Method	20 × 20		50 × 50		100 × 100	
	Path planning time (T)	Path length	Path planning time (T)	Path length	Path planning time (T)	Path length
Classic A*	22	28.7	26	82.8	155	169.7
[18]	18	25.6	22	67.2	85	77.3
Proposed method	10	26.1	16	61.9	42	74.8

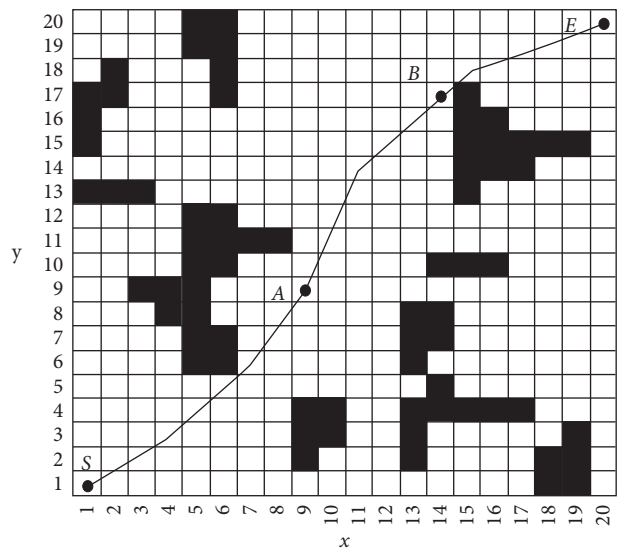
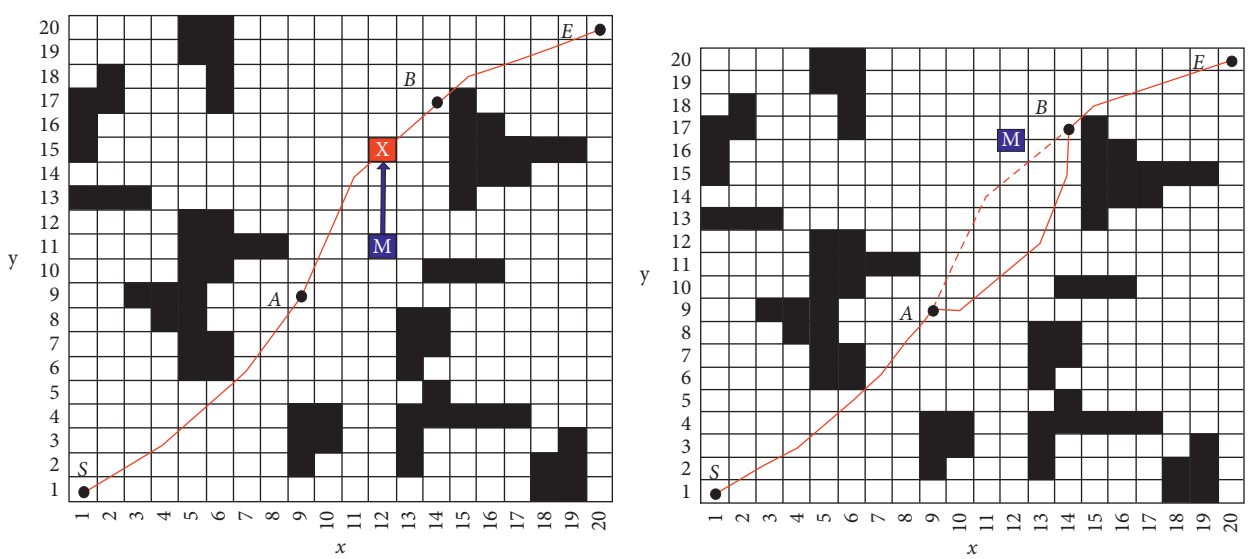


Figure 4: Path planned under given environment.



(a) (b)

Figure 5: Dynamic obstacle avoidance: (a) the robot is located in A and (b) the robot is located in B.

computer, and three environments of different complexity are established. The complexity of the environment is set as the size of the environment and the coverage of obstacles.

Table 3 compares the proposed algorithm with other methods from the path planning time (the number of cycles T) and the planned path length. From the results, it can be found that the proposed method has achieved good results in various environments and significantly reduced the computational complexity. The reason is that in the proposed method, a lot of the search nodes have been simplified compared with the traditional A* algorithm and the method in [18]. In terms of the number of turning points, the

Table 4: Results under dynamic environment.

Method	Reach destination	Planning time (T)	Path length
Classic A*	No	—	—
[18]	Yes	35	55.8
Proposed method	Yes	26	47.6

proposed algorithm can reduce the number of turning points in most cases. And through the improvement of multiobjective optimization and heuristic function, the time and space cost of path planning is greatly reduced.

5.4. Dynamic Environment Simulation. In order to verify the dynamic programming ability of the proposed method, unknown dynamic obstacles are introduced in the simulation, and a simple environment is established for simulation verification. In this environment, according to the given map information, with S (1, 1) as the starting point and E (20, 20) as the end point, global path planning is performed. After optimization, the path shown in Figure 4 is obtained, which is represented by a red line. Select appropriate local subtarget points on the path and add two local subtarget points A and B.

In this environment, the path is divided into three subpaths, namely SA, AB, and BE. When the robot is at the starting position, the DWA does not detect any changes in the map, and the robot travels from point S to the first local subtarget point A along the path. At this time, the robot detects a dynamic obstacle M (blue square). According to the information feedback from the sensors, the robot will collide with the obstacle at the X position (red grid), as shown in Figure 5(a).

According to the sensor information, the traveling speed of the dynamic obstacle M is less than the robot, and the moving trajectory of M has an intersection with the path trajectory of the robot. The EDWA is used to replan the path in the local area and update the subpath AB. As shown in Figure 5(b), the original path is represented by a dashed line, and the path planned by the proposed method is represented by a solid line. When the second subtarget point B is reached, the obstacle is located on the left side of the robot, completing the dynamic obstacle avoidance process.

Table 4 shows the quantitative results of different algorithms when unknown dynamic obstacles are introduced in the 20×20 grid environment shown in Figure 4. The traditional A* algorithm does not have the ability to avoid obstacles. After encountering unknown obstacles, the path planning fails, and the target location cannot be reached. The proposed algorithm and [18] both have the ability to replan the path when encountering unknown obstacles, and both reach the target position. The processing time and path length of the proposed algorithm are better than those of [18].

Figure 6 shows the simulation results of the proposed method for avoiding dynamic obstacles and chasing dynamic target points. The proposed hybrid algorithm can realize local path planning and at the same time use the preview deviation angle algorithm to complete the pursuit of dynamic target points. And after combining the global path information, the hybrid algorithm plans a path with a shorter length and higher smoothness.

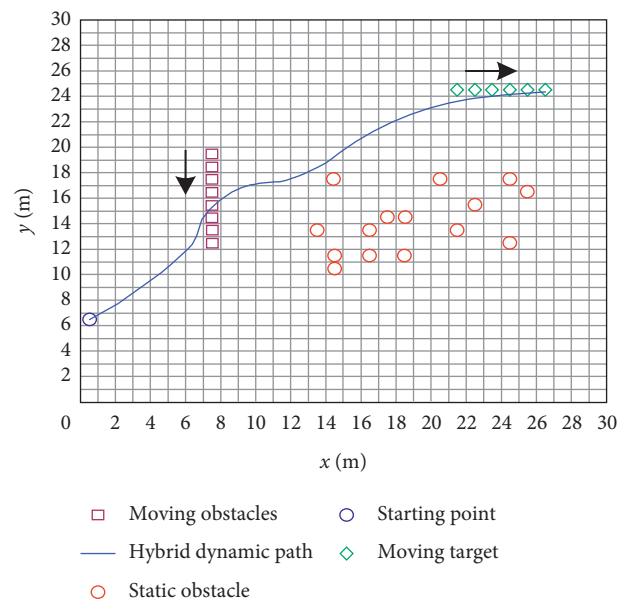


Figure 6: Path planning with dynamic obstacles and moving target.

6. Conclusion

This paper proposes a random obstacle avoidance method for robot path planning that combines the improved A* algorithm with EDWA. By optimizing the search point selection strategy and evaluation function of the traditional A* algorithm, the search efficiency is improved, and a redundant node deletion strategy is introduced to reduce the path length. The EDWA is used to realize the real-time path planning and make up for the shortcomings of the traditional A* algorithm with poor timeliness. While chasing dynamic obstacles, the optimal path is planned in accordance with the global path planning information. Through comparative analysis, the effectiveness of the proposed hybrid algorithm is proved. In the future, we plan to incorporate machine vision techniques to further explore the path planning of mobile robots in complex environments and multitask conditions.

Data Availability

The data included in this paper are available without any restriction.

Conflicts of Interest

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

References

[1] B. K. Patle, G. Babu L, A. Pandey, D. R. K. Parhi, and A. Jagadeesh, "A review: on path planning strategies for navigation of mobile robot," *Defence Technology*, vol. 15, no. 4, pp. 582–606, 2019.

[2] F. Rubio, F. Valero, and C. Llopis-Albert, "A review of mobile robots: concepts, methods, theoretical framework, and

applications," *International Journal of Advanced Robotic Systems*, vol. 16, no. 2, Article ID 1729881419839596, 2019.

[3] G. Fragapane, D. Ivanov, M. Peron, F. Sgarbossa, and J. O. Strandhagen, "Increasing flexibility and productivity in Industry 4.0 production networks with autonomous mobile robots and smart intralogistics," *Annals of Operations Research*, vol. 308, pp. 1–19, 2020.

[4] A. Pandey, A. K. Kashyap, D. R. Parhi, and B. K. Patle, "Autonomous mobile robot navigation between static and dynamic obstacles using multiple ANFIS architecture," *World Journal of Engineering*, vol. 16, 2019.

[5] M. Javaid, A. Haleem, R. P. Singh, and R. Suman, "Substantial capabilities of robotics in enhancing industry 4.0 implementation," *Cognitive Robotics*, vol. 1, 2021.

[6] H. Jahanshahi and N. N. Sari, "Robot path planning algorithms: a review of theory and experiment," 2018, https://arxiv.org/abs/1805.08137.

[7] B. K. Patle, D. R. K. Parhi, A. Jagadeesh, and S. K. Kashyap, "Matrix-binary codes based genetic algorithm for path planning of mobile robot," *Computers & Electrical Engineering*, vol. 67, pp. 708–728, 2018.

[8] E. S. Low, P. Ong, and K. C. Cheah, "Solving the optimal path planning of a mobile robot using improved Q-learning," *Robotics and Autonomous Systems*, vol. 115, pp. 143–161, 2019.

[9] B. Li, H. Liu, and W. Su, "Topology optimization techniques for mobile robot path planning," *Applied Soft Computing*, vol. 78, pp. 528–544, 2019.

[10] M. Kusuma and C. Machbub, "Humanoid robot path planning and rerouting using A-star search algorithm," in *Proceedings of the 2019 IEEE International Conference on Signals and Systems (ICSigSys)*, pp. 110–115, IEEE, Bandung, Indonesia, July 2019.

[11] J.-H. Jung and D.-H. Kim, "Local path planning of a mobile robot using a novel grid-based potential method," *International Journal of Fuzzy Logic and Intelligent Systems*, vol. 20, no. 1, pp. 26–34, 2020.

[12] A. Özdemir and V. Sezer, "A hybrid obstacle avoidance method: follow the gap with dynamic window approach," in *Proceedings of the 2017 First IEEE International Conference on Robotic Computing (IRC)*, pp. 257–262, IEEE, Taichung, Taiwan, April 2017.

[13] Q. Yuan and C.-S. Han, "Research on robot path planning based on smooth A* algorithm for different grid scale obstacle environment," *Journal of Computational and Theoretical Nanoscience*, vol. 13, no. 8, pp. 5312–5321, 2016.

[14] L. Yu, D. Kong, X. Shao, and X. Yan, "A path planning and navigation control system design for driverless electric bus," *IEEE Access*, vol. 6, pp. 53960–53975, 2018.

[15] A. Kaplan, N. Kingry, P. Uhing, and R. Dai, "Time-optimal path planning with power schedules for a solar-powered ground robot," *IEEE Transactions on Automation Science and Engineering*, vol. 14, no. 2, pp. 1235–1244, 2016.

[16] D. Kiss and G. Tevesz, "Advanced dynamic window based navigation approach using model predictive control," in *Proceedings of the 2012 17th International Conference on Methods & Models in Automation & Robotics (MMAR)*, pp. 148–153, IEEE, Miedzyzdroje, Poland, August 2012.

[17] C. Henkel, A. Bubeck, and W. Xu, "Energy efficient dynamic window approach for local path planning in mobile service robotics," *IFAC-PapersOnLine*, vol. 49, no. 15, pp. 32–37, 2016.

[18] X. Zhong, J. Tian, H. Hu, and X. Peng, "Hybrid path planning based on safe A* algorithm and adaptive window approach for mobile robot in large-scale dynamic environment," *Journal of Intelligent and Robotic Systems*, vol. 99, no. 1, pp. 65–77, 2020.

[19] Q. Wu, Z. Chen, L. Wang et al., "Real-time dynamic path planning of mobile robots: a novel hybrid heuristic optimization algorithm," *Sensors*, vol. 20, no. 1, p. 188, 2020.

[20] A. K. Kashyap, D. R. Parhi, M. K. Muni, and K. K. Pandey, "A hybrid technique for path planning of humanoid robot NAO in static and dynamic terrains," *Applied Soft Computing*, vol. 96, Article ID 106581, 2020.

[21] Z. Ma, H. Qiu, H. Wang, L. Yang, L. Huang, and R. Qiu, "A* algorithm path planning and minimum snap trajectory generation for mobile robot," in *Proceedings of the 2021 4th International Conference on Robotics, Control and Automation Engineering (RCAE)*, pp. 284–288, IEEE, Wuhan, China, December 2021.

[22] L. Chang, L. Shan, C. Jiang, and Y. Dai, "Reinforcement based mobile robot path planning with improved dynamic window approach in unknown environment," *Autonomous Robots*, vol. 45, no. 1, pp. 51–76, 2021.

Hindawi
Journal of Robotics
Volume 2022, Article ID 1256083, 9 pages
https://doi.org/10.1155/2022/1256083

Research Article

Personalized Product Recommendation Model of Automatic Question Answering Robot Based on Deep Learning

Jie Peng [1] **and Jianhui Xu** [2]

[1] *Center for Faculty Development, Sichuan Engineering Technical College, Deyang, Sichuan 618000, China*
[2] *Department of Economics and Management, Sichuan Engineering Technical College, Deyang, Sichuan 618000, China*

Correspondence should be addressed to Jie Peng; pj@scetc.edu.cn

Received 30 December 2021; Revised 16 February 2022; Accepted 8 March 2022; Published 21 March 2022

Academic Editor: Shan Zhong

The collaborative filtering algorithm widely used in recommendation systems has problems with the sparsity of scoring data and the cold start of new products. A personalized product recommendation model for automated question-answering robots using deep learning is proposed. First, a personalized attention mechanism at the word level and the comment level is proposed, and the comments and users are individually coded. Then, the bidirectional gated recurrent unit (Bi-GRU) is used to construct the score prediction matrix, and through the dynamic collaborative filtering algorithm to integrate the time characteristics of the user's interest changes. Finally, the feature codes of the users and products are input into the Bi-GRU model for learning, so as to output the recommendation list of personalized products of the automated question answering robot. Experimental results based on the JD and Tianchi datasets show that the training loss of the proposed model is lower than 45 and 23, respectively. And HR@15 and MRR@15 exceed 48 and 15, respectively, which are better than other comparison models. It can better adapt to the actual needs of automatic question-answering robots.

1. Introduction

The deep popularization of the Internet and the rapid development of communication technology have enriched online service businesses increasingly. The subsequent data scale also increased sharply, and the information overload was serious. Whether for physical goods or virtual services, the recommendation system is an important technical means to solve the problem of service information data overload [1]. Automatic question-answering technology is a new intelligent retrieval system that allows users to take natural language query as input, and the system can find and return the exact answer from the relevant documents. Especially for self-service platforms, such as automatic question-answering robots, the quality of the recommendation algorithm is directly related to the long-term benefits of its own development [2, 3]. Personalized recommendation is based on the needs of the users, mining products, or services that users are interested in from a large amount of information and according to the score prediction results for personalized display.

The recommendation algorithm is the core of personalized recommendation, which directly determines the recommendation performance of the recommendation system, and has become a hot issue in current research [4]. Reference [5] explored the consumption habits of hyper-personalized health products as unconventional luxury goods. Research shows that consumers think hyper-personalized products are worth their money, whether they want to own them or not. Therefore, potential consumption habits are one of the leading factors in the shopping process. Although the traditional recommendation algorithm has developed to a certain extent, it still has certain shortcomings. Among them, collaborative filtering is a technology commonly used in personalized recommendation systems. The basic idea is to use users with the same interests in the past to choose similar products in the future [6]. However, the algorithm relies too much on historical data,

has a cold start problem, and performs poorly in the face of new or unpopular products. A typical collaborative filtering algorithm is a collaborative filtering recommendation model based on matrix factorization. Although this model has good recommendation performance, the sparsity of scoring data has always restricted the bottleneck of traditional collaborative filtering [7]. Reference [8] proposed a novel graph-based ranking-oriented recommendation algorithm based on the influence of paired preferences on recommendation diversity. It uses the users' explicit and implicit feedback to improve the resource allocation model, and matches the target users with users with similar preferences to achieve personalized recommendations. However, due to the structure of the model itself, the cold start problem still exists. In addition, because content-based recommendation only relies on the single content information of the product, the accuracy of recommending mature products is not high. In general, the performance is much lower than the collaborative filtering algorithm. It is usually used as a supplement to the collaborative filtering algorithm when the data are sparse. Reference [9] proposed a new method of personalized recommendation using rule-based semantic reasoning. It can easily and quickly generate practical solutions to personalized recommendations. It establishes a connection between the customer and the store by building a recommendation system to provide seamless information exchange. Reference [10] introduced the design and implementation of a personalized product recommendation model based on user interests. The "shopping basket analysis" function model with the Apriori algorithm as the core uses the sales data in the transaction database. It can dig out all kinds of interesting connections between the products purchased by customers and help businesses develop marketing strategies. The shelves can be reasonably arranged to guide sales and attract more customers. Although the above algorithm has achieved certain results in the field of personalized recommendation, data sparseness and poor handling of heterogeneous data still exist.

With the continuous development of computer technology, deep learning has become increasingly mature and gradually applied to personalized recommendation services. Deep learning can obtain useful information from massive amounts of data to obtain a connection between items and users. Secondly, it is also possible to pass all kinds of data through the same hidden space to ensure the consistent representation of the data [11, 12]. Therefore, deep learning can solve or alleviate the impact of these problems on the performance of the recommendation system, and improve the efficiency and accuracy of information acquisition. Reference [13] designed a personalized recommendation system using machine learning, which can recommend that students strengthen their leadership and become unique among their peers. The model was built using Python flask and Jupyter notebook, and tested using a public dataset and a private dataset. The results show that the model has good accuracy. However, the algorithm does not work well in actual application scenarios. Reference [14] proposed an intelligent humanoid robot with self-learning ability based on deep learning and big data knowledge base to communicate with people. By adopting a deep learning method based on recurrent neural network encoder, convolutional neural network encoder, and bidirectional attention flow, better human–computer interaction performance is achieved. However, it has not been able to improve the quality of personalized product recommendation services for the time being. Reference [15] proposes a personalized ranking of neural graphs. By incorporating the user–item interaction diagram into the embedded learning and directly using the user–item interaction information in the embedded learning, more complex structures can be used in interaction modeling. The experimental results prove the effectiveness of the proposed method. However, the applicability of the platform is poor, and it cannot be generally applied to the robot personalized recommendation service platform for automatic question answering.

In summary, in view of the data sparsity and cold start problems of traditional recommendation models, a personalized recommendation model based on deep learning is proposed. It is applied to the automatic question-answering robot to generate an accurate personalized product recommendation list. Most recommendation algorithms based on reviews ignore the personalized information of users or products. The proposed model uses a personalized attention mechanism to individually encode the users (commodities) and comments. It can extract deep hidden features and effectively reduce the impact of data sparsity.

2. Questions Raised

The recommendation system can accurately locate user interests and commodity characteristics. It is a bridge between the users and commodities, and realizes the perfect match between information producers and information consumers. By analyzing the behavior data generated by users' online consumption, the recommendation system can model users' interests and recommend appropriate goods under unsupervised training, which is more intelligent and personalized. The recommendation system is more and more applied in network services. After querying and browsing movies on the Douban film platform, the platform will recommend movies that users may be interested in according to users' tastes. Today's headline news recommendation system can identify users' interests and hobbies according to the news users usually browse. After users browse Taobao, Taobao's recommendation system will push "guess you like" products to users. Suppose a user wants to buy a pair of basketball shoes from Taobao, and enters keywords in the search bar, the recommendation system will recommend some products that the user may be interested in.

Recommendation algorithm directly affects the performance of the recommendation system. A good recommendation algorithm can help users quickly locate goals, save a lot of time, and improve user experience. Although the recommendation algorithm has been widely used in major websites and software, there is still much room for improvement. Traditional algorithms have shortcomings. For example, the collaborative filtering algorithm relies on

historical interactive data, has a cold start problem, performs poorly in the face of new or unpopular products, and the recommendation performance decreases significantly when the interactive data of user products are very sparse. In addition, because content-based recommendation only depends on the single content information of goods, the accuracy of recommendation for mature goods is not high, and the performance is far lower than that of the collaborative filtering algorithm in general.

In order to improve the accuracy of the product recommendation scheme, the proposed model incorporates time series in the collaborative filtering process. This reduces the impact of the information process in the product recommendation process. At the same time, using the Bi-GRU for data learning can further ensure the effectiveness of the automatic question-answering robot recommendation scheme.

3. Proposed Personalized Commodity Recommendation Model

3.1. Overall Framework. The key to the current collaborative filtering product recommendation algorithm is the prediction of product scores, and there are the following three problems: First, there is the problem of sparseness of the scoring matrix. The sparsity of review data is a key factor that affects the accuracy of the final score prediction of the personalized product recommendation system. This is because the number of products on the platform is far greater than the number of comments made by a single user. The comment data present a very obvious sparseness, which in turn affects the insufficient amount of relationship between the users and commodities, and between users and users. The accuracy of the prediction results of commodity ratings has been reduced [16]. Second, the cold start problem. There are two main reasons for the cold start of the recommendation system, namely new users and new products. In the product recommendation platform, new users register every day, and new products are also on the shelves. The system cannot clarify the interest level of the new products or new users [17, 18]. Finally, there is the issue of information expiration. Over time, a user's interest preferences or the popularity of the product will change. Traditional recommendation algorithms do not consider the impact of this change.

To solve the above problems, a personalized product recommendation model based on the Bi-GRU and the dynamic collaborative filtering is proposed. Its overall structure is shown in Figure 1.

First, the text information of user reviews and product reviews is processed by the Bidirectional Encoder Representations from Transformers (BERT) model and the Bi-GRU from the transformer to extract the hidden feature vectors of the users and commodities, respectively. Then, the user's scoring behavior is introduced, and the final product scoring prediction is realized by the TimeSVD++ algorithm. The model uses the Bi-GRU to capture the hidden feature vectors of the users, and combines the image data to reduce the data sparsity and cold start problems in the process of predicting the commodity scoring matrix. Incorporating time series into the collaborative filtering process reduces the impact of the information process in the product recommendation process [19, 20].

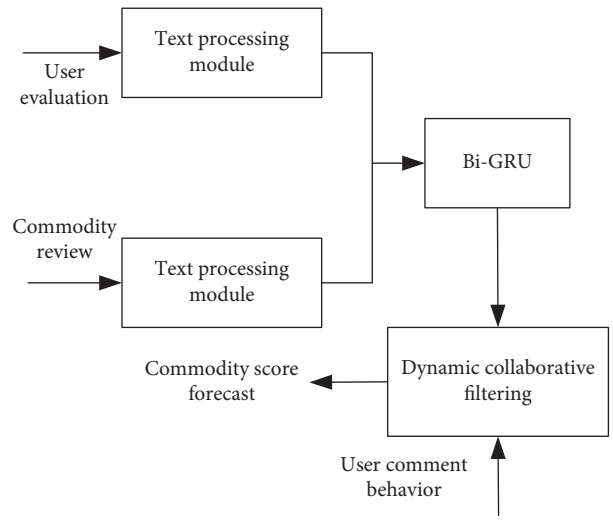


Figure 1: Recommended model architecture.

3.2. User and Product Hidden Feature Extraction

3.2.1. Personalized Comment Encoder. In the user–comment network, a comment set $R_u = \{r_1, r_2, \ldots, r_D\}$ of a user u is given, where D represents the maximum number of comments in the user–comment set. In particular, each comment r_d retains only d' words.

Using the pre-trained word vector, R_u is sent to the word vector mapping layer to obtain $\mathbf{R}_u' \in R^{d \times d' \times k}$, where k is the dimension of the word vector. In order to make the word information comprehensively consider the context information in the forward and backward directions in the comment, $\mathbf{R}_u'$ is sent to the Bi-GRU for encoding, and $\mathbf{H}_u = (h_1, h_2, \ldots, h_d) \in R^{d \times d' \times 2o}$ is obtained. o represents the output dimension of GRU. Since it is the Bi-GRU here, its output dimension is $2o$.

Since each user (commodity) has a unique identity (ID), first, use the first Multilayer Perceptron (MLP) to map the ID into a low-dimensional vector $\mathbf{u}_l \in R^n$. This vector is used to capture the personality information of the user's word level, and its expression is:

$$\mathbf{u}_l = \mathrm{ReLU}\left(\boldsymbol{\omega}_1 u_q + \mathbf{b}_1\right), \tag{1}$$

where ω_1 represents the weight of the first MLP. $\mathbf{b}_1$ is the bias term. u_q is the ID of user u.

Each user's word habits and the polarity expressed by words have individual characteristics when they post comments. In order to make word latent vectors have personalized characteristics, it is first necessary to learn a word level attention vector for a certain user u. The specific calculation is as follows:

$$\mathbf{s}_l = \mathrm{softmax}\left(\mathbf{H}_u \mathbf{P}_l \mathbf{u}_l^T\right), \tag{2}$$

where $\mathbf{u}_l^T \in R^{n\times 1}$ is the transposed vector of $\mathbf{u}_l$. $\mathbf{P}_l \in R^{2o\times n}$ is the transition matrix. $\mathbf{s}_l \in R^{d\times d'\times 1}$ is the attention score corresponding to d' words in d comments. Next, continue to use $\mathbf{s}_l$ with personalized information to adjust the words of the comment, and get the implicit expression of d comments:

$$\mathbf{R} = \mathbf{s}_l^T \otimes \mathbf{H}_u, \tag{3}$$

where transpose the last two dimensions of $s_l \in R^{d\times d'\times 1}$ to get $\mathbf{s}_l^T \in R^{d\times 1\times d'}$. $\otimes$ represents the batch matrix multiplication, such that after $\mathbf{s}_l$ and $\mathbf{H}_u$ are multiplied, the first dimension d remains unchanged, and only the second and third dimension matrix multiplications are performed. Therefore, the implicit expression of the final d comments is $\mathbf{R} \in R^{d\times 1\times 2o}$. In order to facilitate subsequent calculations, the dimension is converted to $\mathbf{R} \in R^{d\times 2o}$.

3.2.2. Personalized User (Commodity) Encoder. Considering that not all information in $\mathbf{R}$ is conducive to constructing a user preference vector, there is a small amount of irrelevant information. Therefore, before gathering d comments, add a gating mechanism to control the flow of information. Specifically, the input of the gating mechanism is $\mathbf{R}$, and its output is a gating weight matrix $\vartheta \in R^{d\times 2o}$:

$$\vartheta = \sigma\left(\mathbf{R}\boldsymbol{\omega}_\vartheta + \mathbf{b}_\vartheta\right), \tag{4}$$

where σ the sigmoid function. $\omega_\vartheta \in R^{2o\times 2o}$ is the weight matrix. $\mathbf{b}_\vartheta$ is the bias term. Next, use ϑ to control the amount of information that each dimension in $\mathbf{R}$ can flow into the next layer:

$$\mathbf{R}^\vartheta = \mathbf{R} * \vartheta, \tag{5}$$

where $*$ is the multiplication of the corresponding elements. Multiply the corresponding elements in $\mathbf{R}$ and ϑ to obtain the adjusted expression of d comments as $\mathbf{R}^\vartheta \in R^{d\times 2o}$.

In reality, the same expression or similar comments will produce different emotional polarities for different users [21, 22]. In order to be able to gather d comments on the user's preference vector based on the user's personalized information, first, use the second MLP to map the user ID to a low-dimensional vector $\mathbf{u}_r \in R^n$ of the review level:

$$\mathbf{u}_r = \mathrm{ReLU}\left(\boldsymbol{\omega}_2 u_q + \mathbf{b}_2\right), \tag{6}$$

where ω_2 is the weight matrix of the second MLP. $\mathbf{b}_2$ is the bias term. Since different reviews have different contributions to the modeling of user preferences [23], it is necessary to learn the personalized attention vector of the review level:

$$\mathbf{s}_r = \mathrm{softmax}\left(\mathbf{R}^\vartheta \mathbf{P}_r \mathbf{u}_r^T\right), \tag{7}$$

where $\mathbf{u}_r^T \in R^{n\times 1}$ is the transposed vector of $\mathbf{u}_r$. $\mathbf{P}_r \in R^{2o\times n}$ is the transition matrix. $\mathbf{s}_r \in R^{d\times 1}$ is the attention score of each comment. Next, according to the attention score, the user preference vector $\mathbf{U} \in R^{1\times 2o}$ can be obtained by gathering each comment:

$$\mathbf{U} = \mathbf{s}_r^T \mathbf{R}^\vartheta. \tag{8}$$

The above has introduced the processing process from the user–comment set to the user preference vector $\mathbf{U}$ in the user–comment network. Similarly, in the product review network, the product feature vector $\mathbf{I} \in R^{1\times 2o}$ can also be obtained from the product review collection.

3.3. Commodity Recommendation Model Based on Dynamic Collaborative Filtering. The dynamic collaborative filtering algorithm is based on the Singular Value Decomposition (SVD)++ algorithm and adds time series items, and hence called the TimeSVD++ algorithm. The TimeSVD++ algorithm is evolved from a simple factorization model. Assume that the user's score prediction matrix $\mathbf{y} \in R^{N\times M}$ for the product has been obtained. N and M, respectively, represent the number of users and the number of products. y_{ui} in the matrix represents the predicted value of user u's rating of product i.

The deep features of the users and products obtained by the Bi-GRU processing review text information are F_U and F_I, respectively. After coupling with the sharing layer and using a factorization machine, the predicted value $\widehat{y}_{ui}$ of the user's scoring matrix for the product is obtained. The TimeSVD++ algorithm first needs to reduce the error between $\widehat{y}_{ui}$ and y_{ui} to obtain the best score prediction. The specific process can be expressed as:

$$\begin{aligned} \widehat{y}_{ui} &= F_I F_U \\ \min Q &= \sum_{y_{ui}} \left(\widehat{y}_{ui} - y_{ui}\right)^2. \end{aligned} \tag{9}$$

Adding the bias term in the scoring prediction process constitutes the SVD model:

$$\widehat{y}_{ui} = F_I F_U + \delta + b_u + b_i. \tag{10}$$

where δ is the average value of the predicted value of the user's product rating during the training process. b_u and b_i represent the user bias item and the product bias item, respectively, and represent the average value of the score prediction value of a user or a product.

Based on the SVD model, the SVD++ algorithm adds user interest information through implicit feedback information. In other words, as long as any user has commented on a certain product, no matter how high or low the scoring prediction value of the review content is, it means that the user is interested in the product. The degree of interest is expressed as $\gamma_j = \left\{\gamma_{j1}, \gamma_{j2}, \ldots, \gamma_{jQ}\right\}$ by a hidden factor. At this time, the user's rating prediction model for the product is revised as follows:

$$\widehat{y}_{ui} = \left[F_U + \frac{\sum_{j\in N(u)} \gamma_j}{\sqrt{|N(u)|}}\right] F_I + \delta + b_u + b_i, \tag{11}$$

where $N(u)$ represents the set of all products that the user u has evaluated.

The basic idea of the TimeSVD++ algorithm is that with the passage of time, the user's preference for the product

changes, i.e., F_U, b_u, and b_i in equation (11) are no longer fixed quantities, but a function of time. However, the product feature vector F_I does not change with time. It can be seen that the TimeSVD++ algorithm incorporates time series items. Therefore, it is necessary to divide multiple time periods along the time axis in the process of predicting the user's rating of the product. Predict product scores in various time periods [24, 25]. The divided time period is represented by $e(t)$. The values of F_U, b_u, and b_i are the same in the same time period, but are different in different time periods. The scoring prediction process of the TimeSVD++ algorithm considering time series items is as follows:

$$\hat{y}_{ui} = \left[F_U(t) + \frac{\sum_{j\in N(u)}\gamma_j}{\sqrt{|N(u)|}}\right]F_I + \delta + b_u(t) + b_i(t), \tag{12}$$

where $b_i(t)$ and $b_u(t)$, respectively, represent the bias of the product and the user at time T, and both consist of a static part and a dynamic part, namely:

$$\begin{aligned} b_i(t) &= b_i + b_{i,e(t)} \\ b_u(t) &= b_u + \beta_u \cdot \operatorname{sign}(t - \bar{s}_u) \cdot |t - \bar{s}_u|^{\beta}, \end{aligned} \tag{13}$$

where $b_{i,e(t)}$ represents the offset of the product in the time period $e(t)$. $\bar{s}_u$ represents the average value of all the ratings given by the user. β represents dynamic weight.

The TimeSVD++ algorithm minimizes the error between the predicted value and the true value after obtaining the predicted value of the score. The calculation is as follows:

$$\begin{aligned} \min Q = &\sum_{y_{ui}} (\hat{y}_{ui} - y_{ui})^2 + \\ &\lambda\left[\|F_U\|^2 + \|F_I\|^2 + \|b_u(t)\|^2 + \|b_i(t)\|^2 + \|\gamma_j\|^2\right], \end{aligned} \tag{14}$$

where λ is the conversion factor.

4. Experiment and Analysis

In order to verify the effect of the proposed model, an experimental platform was built using the deep learning framework Tensorflow provided by Google. The framework integrates in-depth models such as GRU, which can make development simple and easy to understand. Therefore, it has become a more popular learning framework. A three-layer self-attention network will be used in the experiment, and other experimental parameter values used are shown in Table 1.

4.1. Dataset. The experiment uses two real datasets, namely the JD dataset under JD and the Tianchi dataset under Alibaba Cloud. The JD dataset is provided by the e-commerce company Jingdong, one of the largest online B2C retailers in the country. It contains 370,878,895 interactions with 28,710 products from 105,180 customers in 75 days. The Tianchi dataset is a public dataset provided by the Ali Mobile recommendation algorithm. It is a real user–commodity dataset based on Alibaba's mobile commerce platform. This dataset provides 23291027 interactions of 20,000 customers on 4,758,484 products in one month.

Table 1: Experimental training parameters.

Experimental parameters	Specific settings
Optimization function	Adam
Discarding rate	0.6
Learning rate	0.001
Embedded vector dimension	256

Before the start of the experiment, the above two datasets were preprocessed. First, products with less than 5 appearances and users with less than 10 interactions are screened out. Then, the two datasets are divided into a training set and a test set according to time. 85% of the interactions are the training set, and the rest are the test set, i.e., for the JD dataset, 64 days of data are used for training and 11 days of data for testing. For the Tianchi dataset, 26 days of data are used for training, and the remaining 4 days of data are used for testing.

At the same time, in order to control the uniqueness of independent variables, when conducting comparative experiments, consider that the most commonly used collaborative filtering in recommendation cannot recommend products that have not appeared before. Therefore, product interactions and users that have not appeared in the training set are screened out from the test set. The two datasets after being preprocessed are shown in Table 2.

4.2. Evaluation Index. When the recommendation system makes recommendations, a limited number of products will be recommended each time. If the recommendation is effective, then the products that meet the user's needs should be included in the recommendation list. Two commonly used top-k evaluation indicators are selected to evaluate the quality and effect of the recommendation list generated by the recommendation model.

(1) Hit Ratio (HR): HR is calculated as follows:

$$\mathrm{HR@}K = \frac{W@K}{\Omega}. \tag{15}$$

The meaning of the denominator is all test sets. The meaning of the numerator is the sum of the number of test sets in the top-k list recommended to each user. HR and Recall have the same functions, and both can evaluate the recall rate of the recommendation system in the recall phase. It can be seen from the definition that the larger the HR, the better the recommendation effect.

Mean Reciprocal Rank (MRR): MRR is calculated as follows:

$$\mathrm{MRR} = \frac{1}{N}\sum_{i=1}^{|N|}\frac{1}{\mathrm{rank}_i}, \tag{16}$$

where rank_i is the position of the first item in the ground truth result of the recommendation list for the ith user. Take MRR@20 as an example. If the item actually clicked by the user appears in the recommendation list and is ranked n th,

Table 2: Preprocessed dataset parameters.

Dataset	JD	Tianchi
Total users	103864	18205
Total commodity	24759	674263
Total interaction	38012329	8665937
Average number of user interactions	451.30	360.94
Training set interaction	32310478	7366046
Test set interaction	5701851	1299891

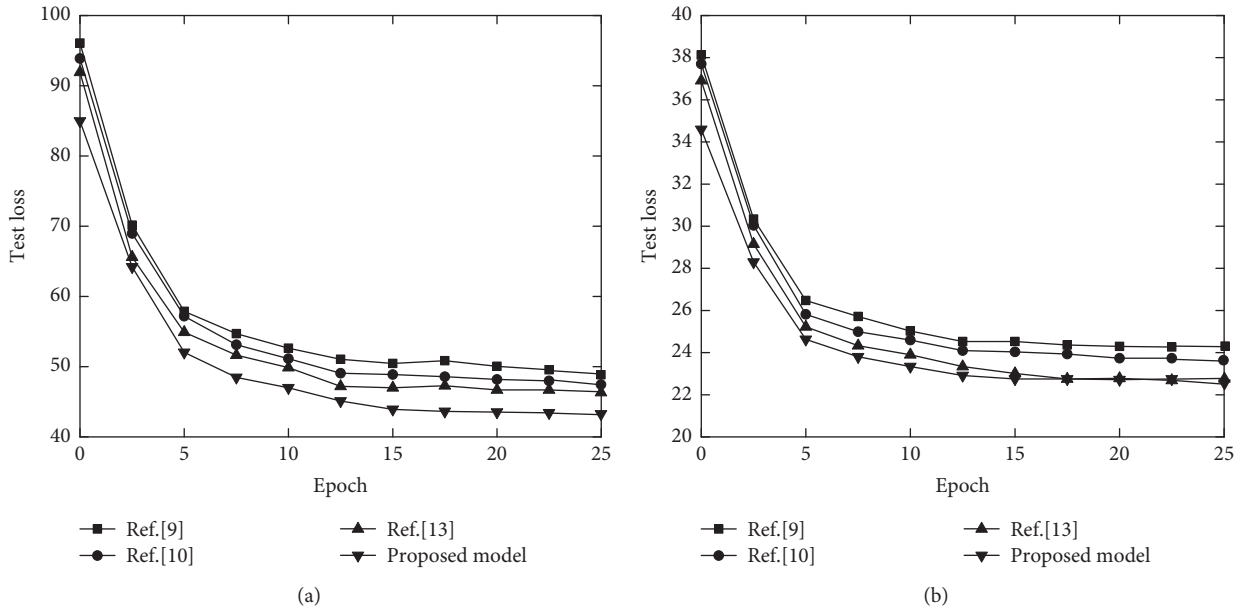


Figure 2: Test loss comparison results of different models. (a) JD dataset. (b) Tianchi dataset.

then $n \geq 1$ and if $n \leq 20$, then the MRR is equal to $1/n$. If this item does not appear in the first 20, the MRR is equal to 0. Therefore, the larger the MRR, the better. The MRR is used to evaluate the order of the items in the recommended list. Therefore, if evaluating the quality of the recommendation sequence of the recommendation model, MRR is a very important indicator.

4.3. Comparison of Loss Functions. The test loss results of the proposed model and reference [9, 10, 13] on the twodata sets are shown in Figure 2. The ordinate represents the superimposed loss of all data in each epoch.

It can be seen from Figure 2 that all four models have reached convergence. However, the test loss of the proposed model on the JD and Tianchi datasets is the smallest, which is lower than 45 and 23, respectively. The proposed model combines the Bi-GRU and the dynamic collaborative filtering algorithm to achieve personalized product recommendation. The hidden attention vector is learned in a targeted manner, which can further improve the accuracy of the recommendation. Reference [9] emphasizes rules-based semantic reasoning and reference [10] recommends products based on user interests. All lack dynamic information filtering, and hence the test loss is relatively large. Reference [13] uses machine learning to generate a product recommendation list, which has improved recommendation performance compared to reference [9, 10]. However, there is a lack of dynamic prediction of user interests, and the accuracy of the recommendations needs to be improved.

4.4. Performance Comparison with Other Algorithms. In order to demonstrate the performance of the proposed model, compare it with reference [9, 10, 13]. The comparison of the evaluation results of selecting HR@15 and MRR@15 on the two datasets is shown in Figure 3. In order to see the change trend of the index more clearly and intuitively, the ordinate is enlarged by 100 times.

On the whole, it can be seen from Figure 3 that the four models have great trends in the first 5 iterations. After 5 iterations, the indicator changes stabilized. In the end, the proposed model's indicators surpass other models, and the effect is the best. Taking HR@15 as an example, the values of the proposed models on the two datasets exceed 68 and 48, respectively. Taking MRR@15 as an example, the values of the proposed models on the two datasets exceed 25 and 15, respectively. Reference [9] proposed a rule-based semantic inference personalized recommendation model, which can quickly generate personalized recommendations and practical solutions. However, the rules are pre-set and lack dynamic updates. Therefore, the application effect of the

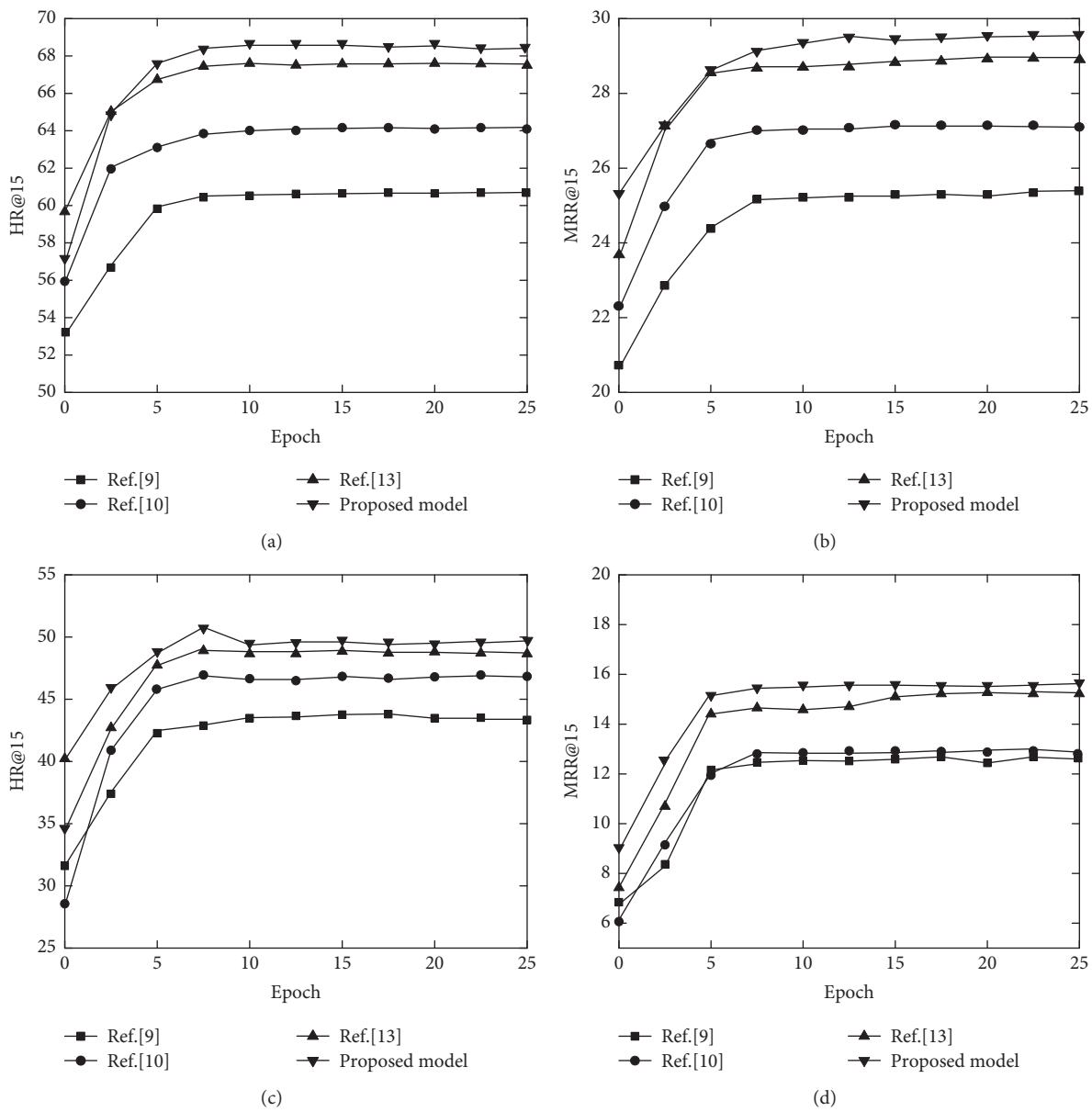


Figure 3: Comparison results of the evaluation indexes of different models. (a) HR@15 comparison results of the JD dataset. (b) MRR@15 comparison results of the JD dataset. (c) HR@15 comparison results of the Tianchi dataset. (d) MRR@15 comparison results of the Tianchi dataset.

automatic question-answering robot is not good. Taking MRR@15 as an example, which is lower than 25 and 13, respectively, reference [10] uses user interests to design personalized product recommendation models. Sales data in the transaction database is used to mine various interesting connections between the products purchased by customers. However, it lacks a powerful learning algorithm as a support, and hence the overall performance is not much different from reference [9]. Reference [13] uses machine learning to build a personalized recommendation system to generate a personalized recommendation plan. Although there are good learning algorithms for data analysis, there are still some problems in the actual application process because of the lack of dynamic user interest learning. Therefore, the maximum HR@15 should not exceed 67. The proposed model uses the Bi-GRU network to learn text features, and uses the dynamic collaborative filtering algorithm to generate product recommendations. The fused model can better adapt to the actual needs of automatic question-answering robots.

Separately, on the JD dataset, the proposed model has obvious advantages, while on the Tianchi dataset, the advantage is slightly smaller. It may be because different models of the dataset have different expressiveness because the size of the dataset and the data distribution of each latitude in the data are different, resulting in different training effects. The proposed model should also be widely used in different datasets to enhance the robustness of the

model. On the other hand, the average number of interactions in the JD dataset is slightly larger than the Tianchi dataset. A larger number of interactions can capture more information. However, this can also be regarded as a research direction in the future. It is more universal and advantageous to choose and explain the number of interactions through continuous experimentation. In addition, the existing research experience shows that using data that are closer to the test set time to train the model will get better results. The experimental results also show that the effect on the JD dataset is better and more obvious. It also proves the correctness of the previous experience again.

5. Conclusion

With the rapid growth of users and product content, automatic question-answering robots using traditional recommendation algorithms have obvious shortcomings in the face of sparse historical interactive data or new product recommendations. The use of deep neural networks in the recommendation field can effectively solve the above problems. To this end, a personalized product recommendation model for automated question-answering robots using deep learning is proposed. A personalized attention mechanism is used to personalize comments and users and we input the Bi-GRU score prediction model for learning. At the same time, the dynamic collaborative filtering algorithm is used to integrate the time characteristics of user interest changes so as to get the recommendation plan of personalized products. The experimental results based on the JD and Tianchi datasets show that:

(1) The proposed model uses the attention mechanism to obtain in-depth features of products and reviews, which can reduce the error of the recommended solutions. Its training loss is lower than 45 and 23, respectively.

(2) The proposed model combines the Bi-GRU and the dynamic collaborative filtering algorithm to achieve personalized product recommendation. Comprehensive consideration of the timeliness of user interest can improve the reliability of the recommended scheme. Take MRR@15 as an example; its values on the two datasets exceed 25 and 15, respectively. The overall performance is better than other models.

Experiments have found that due to the complex model of the deep learning network, training takes a lot of time and requires high computer configuration. Therefore, next, we will consider the optimization and upgradation of the deep neural networks and the combination with the recommendation system based on efficiency to ensure that it can be more efficient in practical applications.

Data Availability

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

Conflicts of Interest

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

References

[1] P. Nitu, J. Coelho, and P. Madiraju, "Improvising personalized travel recommendation system with recency effects," *Big Data Mining and Analytics*, vol. 4, no. 3, pp. 139–154, 2021.

[2] D. De Venuto and F. Bellotti, "A CNN based personalized product recommendation model," *International Journal of Smart Business and Technology*, vol. 8, no. 1, pp. 21–32, 2020.

[3] J. R. Bock and A. Maewal, "Adversarial learning for product recommendation," *AI*, vol. 1, no. 3, pp. 376–388, 2020.

[4] W. Deng, "Leveraging consumer behaviors for product recommendation: an approach based on heterogeneous network," *Electronic Commerce Research*, vol. 5, no. 1, pp. 1–27, 2020.

[5] M. S. Rosenbaum, G. Contreras Ramirez, J. Campbell, and P. Klaus, "The product is me: hyper-personalized consumer goods as unconventional luxury," *Journal of Business Research*, vol. 129, no. 6, pp. 446–454, 2021.

[6] B. R. Sreenivasa and C. R. Nirmala, "Hybrid location-centric e-Commerce recommendation model using dynamic behavioral traits of customer," *Iran Journal of Computer Science*, vol. 2, no. 3, pp. 179–188, 2019.

[7] B. Sheng and G. Sun, "Matrix factorization recommendation algorithm based on multiple social relationships," *Mathematical Problems in Engineering*, vol. 2021, Article ID 6610645, 8 pages, 2021.

[8] W. Zhou and W. Han, "Personalized recommendation via user preference matching," *Information Processing & Management*, vol. 56, no. 3, pp. 955–968, 2019.

[9] X. Chen, Y. Xue, and Y. Shiue, "Rule based semantic reasoning for personalized recommendation in indoor O2O e-commerce," *International Core Journal of Engineering*, vol. 6, no. 1, pp. 309–318, 2020.

[10] J. Zhang, "Personalised product recommendation model based on user interest," *Computer Systems Science and Engineering*, vol. 34, no. 4, pp. 231–236, 2019.

[11] X. Bu, J. Zhu, and X. Qian, "Personalized product search based on user transaction history and hypergraph learning," *Multimedia Tools and Applications*, vol. 79, no. 11, pp. 1–19, 2020.

[12] B. Wu, Y. Ye, and Y. Chen, "Visual appearance or functional complementarity: which aspect affects your decision making?" *Information Sciences*, vol. 476, no. 8, pp. 19–37, 2019.

[13] G. Suganya, M. Premlatha, P. Dubey, A. R. Drolia, and S. Srihari, "Subjective areas of improvement: a personalized recommendation," *Procedia Computer Science*, vol. 172, no. 5, pp. 235–239, 2020.

[14] W. Budiharto, V. Andreas, and A. Gunawan, "Deep learning-based question answering system for intelligent humanoid robot," *Journal of Big Data*, vol. 7, no. 1, pp. 77–92, 2020.

[15] Z. Hu, J. Wang, Y. Yan, P. Zhao, J. Chen, and J. Huang, "Neural graph personalized ranking for Top-N recommendation," *Knowledge-Based Systems*, vol. 213, no. 8, pp. 106426–106437, 2020.

[16] W. Dai, Y. Wang, J. Ma, and Q. Jin, "BTR: a feature-based Bayesian task recommendation scheme for crowdsourcing system," *IEEE Transactions on Computational Social Systems*, vol. 7, no. 3, pp. 780–789, 2020.

[17] B. Wu, X. He, Z. Sun, L. Chen, and Y. Ye, "ATM: an attentive translation model for next-item recommendation," *IEEE*

Transactions on Industrial Informatics, vol. 16, no. 3, pp. 1448–1459, 2020.

[18] S. Kang, C. Jeong, and K. Chung, "Tree-based real-time advertisement recommendation system in online broadcasting," *IEEE Access*, vol. 8, no. 3, pp. 192693–192702, 2020.

[19] T. Schreiner, A. Rese, and D. Baier, "Multichannel personalization: identifying consumer preferences for product recommendations in advertisements across different media channels," *Journal of Retailing and Consumer Services*, vol. 48, no. 5, pp. 87–99, 2019.

[20] M. Asenova and C. Chrysoulas, "Personalized micro-service recommendation system for online news," *Procedia Computer Science*, vol. 160, no. 3, pp. 610–615, 2019.

[21] B. Huynh, A. Ibrahim, Y. S. Chang, H. Tobias, and J. O'Donovan, "User perception of situated product recommendations in augmented reality," *International Journal of Semantic Computing*, vol. 13, no. 3, pp. 289–310, 2019.

[22] C. Yu, Z. Zhang, C. Lin, and Y. J. Wu, "Can data-driven precision marketing promote user ad clicks? Evidence from advertising in WeChat moments," *Industrial Marketing Management*, vol. 90, no. 5, pp. 481–492, 2020.

[23] W. Li and B. Xu, "Aspect-based fashion recommendation with attention mechanism," *IEEE Access*, vol. 8, no. 8, pp. 141814–141823, 2020.

[24] R. S. Sreepada and B. K. Patra, "Mitigating long tail effect in recommendations using few shot learning technique," *Expert Systems with Applications*, vol. 140, no. 2, pp. 112887.1–112887.17, 2020.

[25] S. Hu, A. Kumar, F. Al-Turjman, S. Gupta, S. Seth, and Shubham, "Reviewer credibility and sentiment analysis based user profile modelling for online product recommendation," *IEEE Access*, vol. 8, no. 1, pp. 26172–26189, 2020.

Hindawi
Journal of Robotics
Volume 2022, Article ID 2576303, 11 pages
https://doi.org/10.1155/2022/2576303

Research Article

Research and Implementation of Turbo Coding Technology in High-Speed Underwater Acoustic OFDM Communication

Yarang Yang [1] **and Yunpeng Li** [2]

[1] *College of Physics and Electrical Engineering, Kashi University, Kashi, Xinjiang 844006, China*
[2] *Qingdao Vocational and Technical College of Hotel Management, Qingdao, Shandong 266100, China*

Correspondence should be addressed to Yunpeng Li; liyunpeng@qchm.edu.cn

Received 28 December 2021; Accepted 17 February 2022; Published 15 March 2022

Academic Editor: Shan Zhong

It is demonstrated that the fully parallel turbo decoding algorithm can achieve an approximate error correction decoding performance when 36 iterations are used and when the log-map algorithm with 6 iterations is used. By comparison, it is shown that it can achieve much higher decoding rates than the log-map algorithm for various frame lengths of LTE standard turbo codes at the cost of higher hardware resource requirements. According to the fully parallel turbo decoding algorithm, this paper proposes a scheme for implementing a fully parallel turbo decoder on FPGA, detailing the overall structure and processing of the decoder hardware implementation, the design of the algorithm block processing unit, and the interleaving module. The performance of the decoder is tested by fixed-point simulation for different frame lengths of turbo coding in LTE standard, and it is proved that the fully parallel turbo decoder can be applied to turbo coding of various frame lengths. Both simulation and experimental results show that the distributed cancellation method and the joint estimation cancellation method have good results for both time-domain impulse noise and large-amplitude single frequency noise cancellation, while the joint estimation cancellation method of large-amplitude single frequency noise cancellation first has better performance.

1. Introduction

Hydroacoustic communication is a rapidly developing field of scientific research; its engineering applications used to be limited to military aspects to solve the problem of mine remote control, submarine to submarine, mother ship and submarine, or other underwater unmanned combat platform transmissions to obtain battlefield information, and limited bandwidth, multipath, transmission delay, and channel structure of the rapid time change are currently difficult problems [1].

After years of development, OFDM has fully demonstrated its advantages, but some problems based on the hydroacoustic OFDM communication system are still not well solved. Many communication methods are controversial, there is no unified standard, and different experiments are usually used in different system parameters, such as coding modulation, communication bandwidth, and carrier frequency. The problems in hydroacoustic OFDM communication need to be studied in more depth, such as how to eliminate external noise interference; how to eliminate and compensate the impact of UAC on the OFDM system to ensure high system reliability; how to better use the sparsity characteristics of UAC to improve system performance; how to adapt to the hydroacoustic communication channel by using coding techniques to improve system performance; and how to improve the utilization of frequency band to achieve high-speed hydroacoustic communication.

The hydroacoustic channel is very complex, with strong multipath and complex noise interference, and the channel parameters are time varying, so it is very difficult to get the state information of the hydroacoustic channel at the transmitter side. The transmitter side achieves adaptive transmission based on the obtained CSI, but the hydroacoustic channel is time varying and there is a certain delay

in the process of feeding CSI back to the transmitter side. Therefore, this paper investigates the prediction of hydroacoustic channels based on the OFDM system and proposes a channel prediction technique based on the sparsity of hydroacoustic channels to compensate for the time delay of underwater transmission, so that the obtained channel state information can better reflect the current channel condition. Because of the channel characteristics such as space-time frequency variation and narrow-band, high-noise, strong multipath interference, and large transmission delay, Turbo codes combine convolutional codes and random interleavers to realize the idea of random coding, while soft-output iterative decoding is used to approximate the maximum likelihood decoding, which can help achieve the channel coding performance limited by Shannon theory over Gaussian channels [2, 3].

2. Key Technologies for High-Speed Hydroacoustic Communication

The key technologies for high-speed hydroacoustic communications include six main directions, namely, new modulation methods to reduce the effects of multipath; coding techniques, including compression coding required for image transmission and error correction coding techniques that can improve system reliability; receiver architecture, as reflected in the use of powerful signal processors and algorithms; underwater network systems; various explicit and implicit diversity techniques applied to fading channels; and hydroacoustic channel physics. The topics include the simulation and measurement of channels [2].

After decades of development, the hydroacoustic communication method has gradually tended to move from noncoherent communication to coherent communication, with new technologies and advances in signal modulation and demodulation and signal processing, such as spatial modulation technology and blind equalization technology. According to historical materials, it was Leonardo da Vinci who pioneered a method of transmitting and receiving information underwater, which can be traced back to the earliest hydroacoustic communication system. Several frequency points are selected in the frequency band, and the transmission of information works at the transmitting end by transmitting carriers of different frequencies. The frequency shift keying system is more stable than the ASK system communication performance, but the same communication rate is slower, and the frequency band utilization is not high, the information of adjacent frequency points is easily affected by Doppler shift, and the demodulation is easy to generate false codes.

Multipath propagation causes intercode interference in single-carrier digital communication systems, introduces multipath expansion between symbols, and to some extent limits the use of single-carrier communication technology for high-speed hydroacoustic communication systems. In shallow sea channels, the time-varying nature of the seawater medium and the relative motion between the transmitter and receiver are the two factors that cause Doppler frequency shift. The application of OFDM communication technology to hydroacoustic communication requires overcoming the Doppler shift. Due to the above complex characteristics of the shallow sea hydroacoustic channel, various OFDM-based communication technologies must be improved accordingly to apply to the hydroacoustic channel, and a lot of experimental validation is needed after the improvement to be put into use.

3. OFDM for Hydroacoustic Communication Systems

3.1. OFDM for Hydroacoustic Communication Systems. The basic principle of OFDM is to split the original signal into N subsignals by serial-parallel conversion: if the serial code rate is R, the converted code rate is R/N, the subsignal period is Δt, and the theoretical value of the complete period of the original signal is $T = N * \Delta t$. Then the N subsignals are modulated on N mutually orthogonal subcarriers, and finally the N modulated signals are summed to get the transmit signal. At the receiving end, the input signal is divided into N branches, and then the N subcarriers are mixed and integrated to recover the original data. At the receiving end, the input signal is divided into N branches, which are mixed and integrated with N subcarriers to recover the parallel data, and then the original data can be recovered through parallel-serial conversion and demodulation [4].

$$[T_1 - T_{\min} + k \cdot T] < T_2 + \Delta t < [T_1 + T_{\min} + (k+1) \cdot T]. \tag{1}$$

The phenomenon of self-interference exists in the hydroacoustic communication system, that is, the signal sent by a terminal after mechanical vibration and another short-range low-attenuation propagation method to reach its receiving transducer, and its amplitude is much larger than the target signal propagated from other terminals over long distances. Therefore, the frequency division multiplexing of the hydroacoustic system in the simultaneous transmission and reception of signals, the need to filter out the self-interference signal, and the system's filtering performance put forward high requirements, and engineering implementation is also more difficult; time-division multiplexing system in the reception of signals also needs to send a signal after a certain interval, waiting for the elimination of the self-interference component before collecting data.

$$T_2 + \Delta t \leq [T_1 + T_{\min} + k \cdot T]. \tag{2}$$

Different terminals have different requirements for communication indicators, which include communication type, communication target, communication speed, and communication delay. Take an exploration activity that includes a diver, an underwater probe, an underwater beacon, and a mother ship as an example: the diver and the mother ship communicate with each other through commands, and the mother ship sends commands to the diver to direct him to conduct underwater exploration. This process requires a small communication delay; the underwater probe sends data back to the mother ship in one direction, which

does not require high communication delay, but high communication speed; the underwater beacon sends commands back to the mother ship in one direction, which requires high real-time communication and does not require high communication speed.

$$v_0 = \frac{k \cdot T}{2\left[T_1 + T_{\max} + (k+1) \cdot T - 2\Delta t\right]},$$
$$a = A\frac{Sf_r f^2}{f_r^2 - f^2} - B\frac{f^2}{f}. \tag{3}$$

The overlapping time slot allocation method is still applied under the condition that the propagation time of the acoustic signal is greater than the minimum time slot. The first time slot is allocated for the host and the slave to send signals simultaneously, while the second time slot is allocated for the host and the slave to receive signals simultaneously. When communication is performed in this way, the highest communication rate per unit time is achieved because as many data frames as possible are transmitted, but the communication delay of the system is larger. If the host or slave needs to send data at moment t, the earliest it can receive this signal is at moment $3t$, and the longest communication delay is twice the propagation time.

3.2. Key Technologies of a Hydroacoustic Communication System with OFDM. The key technologies of OFDM's hydroacoustic communication system mainly include spatial diversity.

$$\psi(t) = \exp\left(j\lambda t + \gamma t^a - j\beta \operatorname{sgn}(t)\varphi(t,a)\right),$$
$$\varphi(t,a) = \begin{cases} \tan\left(\dfrac{a\pi}{3}\right), & a \neq 1, \\ \ln t, & a = 1. \end{cases} \tag{4}$$

Having$\lambda, \gamma, \alpha, \beta$ as four parameters, the characteristic function of the steady-state distribution can be obtained by these four parameters. The accurate measurement of the pressure expansion ratio of the signal effectively compensates for the Doppler of the signal, while the compensation of the consistent Doppler frequency bias is usually required again after the variable sampling due to the measurement and compensation accuracy.

Compared with single-carrier systems, OFDM signals behave in the time domain as a superposition of N mutually orthogonal subcarrier signals. When all these N signals are exactly summed at the peak point, the OFDM signal produces the maximum peak, which is N times the average power, and the peak-to-average power ratio is correspondingly larger. As the number of subcarriers N increases, the maximum value of peak-to-average power ratio (PAPR) also increases linearly, which puts high requirements on the linear range of the transmitter front-end amplifier, so a simple and effective technique to reduce the peak-to-average power ratio becomes an important research direction for OFDM technology.

$$h(t,\tau) = \sum_{p=1}^{N_{pa}} \xi_p(t)\delta\left(\tau + \tau_p(t)\right), \tag{5}$$

where N denotes the number of paths, $\tau_p(t)$ denotes the amplitude of the path, and $\xi_p(t)$ denotes the delay of the amplitude of the path. Coherent communication of hydroacoustic communication systems must perform channel estimation and compensation for correct demodulation. Using hydroacoustic channel characteristics, such as the sparse nature of the channel, the design of a simple structure, low computation, and high estimation accuracy of the channel estimation, an equalization algorithm is a key technology for multicarrier hydroacoustic communications.

As much as possible, the bits in the information sequence input to the interleaver are disordered so that the information bits nearby before interleaving are permuted to positions farther apart after interleaving, especially not to have information bits that are originally adjacent to each other still in adjacent positions after being permuted. Try not to have the bits in the information sequence of an input interleaved, after being permuted by the interleaver and when it does not go through interleaving, corresponding to the check bits in the output check sequence of different component encoders that are most correlated with it and are in the position where the censoring operation will be performed. The theory of ray propagation is expounded, and, on this basis, the sound ray span model and the search algorithm for eigenrays in the layered ocean based on the model are introduced. The algorithm can quickly and accurately search for important eigenrays. Then, the eigenpath model of the underwater acoustic channel is given in this paper, and the eigenpath channel is analyzed in detail with mathematical formulas from the aspects of propagation attenuation and delay, multipath components, and Doppler frequency shift, along with the rationality and feasibility of the model. The length of the interleave is increased as much as possible so that the correlation of the information sequence before and after interleaving is as low as possible, thus obtaining a turbo coded output sequence with a larger minimum Hamming distance and a minimum Hamming distance code word, which can reach a lower BER limit.

In the practical application of communication systems, the choice of interleaver used in turbo encoder requires a balance between the error correction decoding performance and the complexity it requires for hardware implementation. In the LTE standard, a quadratic permutation polynomial interleaved is chosen for turbo coding. The redundant bits can be removed by the redundancy process because the conventional turbo encoder uses two fractional convolutional encoders to obtain two sequences of check bits equal to the length of the information bit sequence as the redundant bits of the coded output, which is twice as many as the check bits needed in general, and this extra check bits can be removed by the redundancy process, which can reduce the redundancy of the turbo coded output sequence and obtain higher performance. These extra parity bits can be removed through redundancy processing, thus reducing the

redundancy of the turbo coding output sequence and obtaining a higher code rate.

3.3. Advantages and Disadvantages of a Hydroacoustic Communication System with OFDM. The advantages of the OFDM system are strong resistance to intercode interference, suitability for high-speed data transmission, high-frequency utilization, and strong antifading ability. OFDM is strong against intercode interference due to the use of cyclic prefix [5].

The shortcomings of OFDM technology include sensitivity to frequency bias and phase noise and large power peak-to-mean ratio. The power efficiency of the OFDM signal is lower, and the OFDM signal is summed by multiple independent modulated subcarrier signals [6].

4. Turbo Code-Based High-Speed Hydroacoustic OFDM Communication Coding Scheme Design

4.1. High-Speed Hydroacoustic OFDM-Based Communication System Structure Design. The architecture of the high-speed hydroacoustic OFDM-based communication system shown in Figure 1 includes coding, interleaving, and modulation modules.

High-speed hydroacoustic OFDM-based communication system architecture support all mandatory and optional data rates: 6, 9, 12, 18, 24, 36, 48, and 54 Mb/s, including adaptive modulation and multipath fading channel dispersion, i.e., simulating dynamically changing data rates encoded with a high channel fading rate and allowing data rates to change more rapidly, and random data generation at a bit rate simulates a variety of code rate environments that hold at different data rates for some time so that the required data rate depends on the expected value of the module; it uses 1/2, 2/3, and 3/4 convolutional coding code rates, uses data interleaving design, and supports BPSK, QPSK modulation, 16QAM, and 64QAM modulation [7]. See Figure 2 for an OFDM frame fusion module based on the architecture of a high-speed hydroacoustic OFDM communication system.

The high-speed hydroacoustic OFDM-based communication system architecture corrects the number of data symbols and omitted bits in each packet and uses an uninterrupted frame structure, omitting the tail part; the decoder has a reset state and fixes the transmit power level instead of the average signal-to-noise ratio of the different channels.

In general, a high input SNR implies a high output SNR and a high prediction accuracy. However, the length of the cyclic prefix determines the size of the input SNR. If the cyclic prefix is too long, fewer noise coefficients will be removed and the accuracy of the time-domain channel prediction will be affected. If the cyclic prefix is too short, it will make the actual channel tap coefficients potentially leaky and lead to a serious loss of prediction performance [8–11]. Therefore, in time domain channel prediction, choosing the appropriate length of the cyclic prefix has a certain impact on the improvement of the prediction performance.

The Turbo frequency domain equalizer structure is shown in Figure 3. Both demodulator and decoder use processing operates a priori on the input and output values, providing a Turbo excitation for the next iteration. The same Turbo processing principles apply to coded CDMA systems, producing Turbo multiuser detection.

Amplifier circuits are divided into Class A, Class B, Class A-B, and Class *D* according to their efficiency, of which Class A amplifiers have an amplifier tube conduction angle of 360°. Class B amplifiers have a conduction angle of 180°; the amplifier tube only works in half a cycle and can only amplify the positive half-cycle signal [12–14]. Raised cosine filters (also known as Nyquist filters) generate bandwidth-limited signals without intermember interference, making them the most commonly used filter type for OFDM modulation. In most communication systems, the raised cosine overall response is obtained by evenly distributing the filtering between the transmitter and receiver, resulting in a square root raised cosine filter. For the OFDM modulation scheme, where the amplitude is constant and the data being transmitted are used to control the carrier frequency, a Gaussian filter is typically used to modulate the signal (rather than the RF signal). The common push-pull amplification structure is the use of two amplifier tube combinations, respectively, amplifying the positive and negative half-cycle signal and then synthesizing as the output. Class A and B amplifiers are slightly less efficient than class B amplifiers, but they solve the problem of crossover distortion of class B push-pull amplification. Class *D* amplifier tube is switching, and the efficiency is very high, but the circuit design and impedance matching are very tedious and easy to lead to signal nonlinear distortion. This system selects the high-efficiency Class A and B amplifier structure to build the power amplifier circuit.

4.2. Design of Turbo Code Coding Scheme in the High-Speed Hydroacoustic OFDM Communication. The original coding structure of the Turbo code encoder is Parallel Concatenated Convolution Codes (PCCC), as shown in Figure 4 [8], which cleverly combines convolutional codes and random interleavers to realize random coding while constructing long codes from short codes through the interleaved.

The Turbo code has a much better error correction decoding performance than traditional channel coding schemes such as packet codes and convolutional codes because, on the one hand, the interleave is used to obtain two check sequences with low correlation as the coding output and, on the other hand, the decoder uses an iterative decoding method by exchanging soft information between the component decoders. Since a hard decision will lose some information in the information sequence, the decoder uses two soft-input and soft-output component decoders to preserve as much as possible the information of the system information sequence contained in the different check sequences.

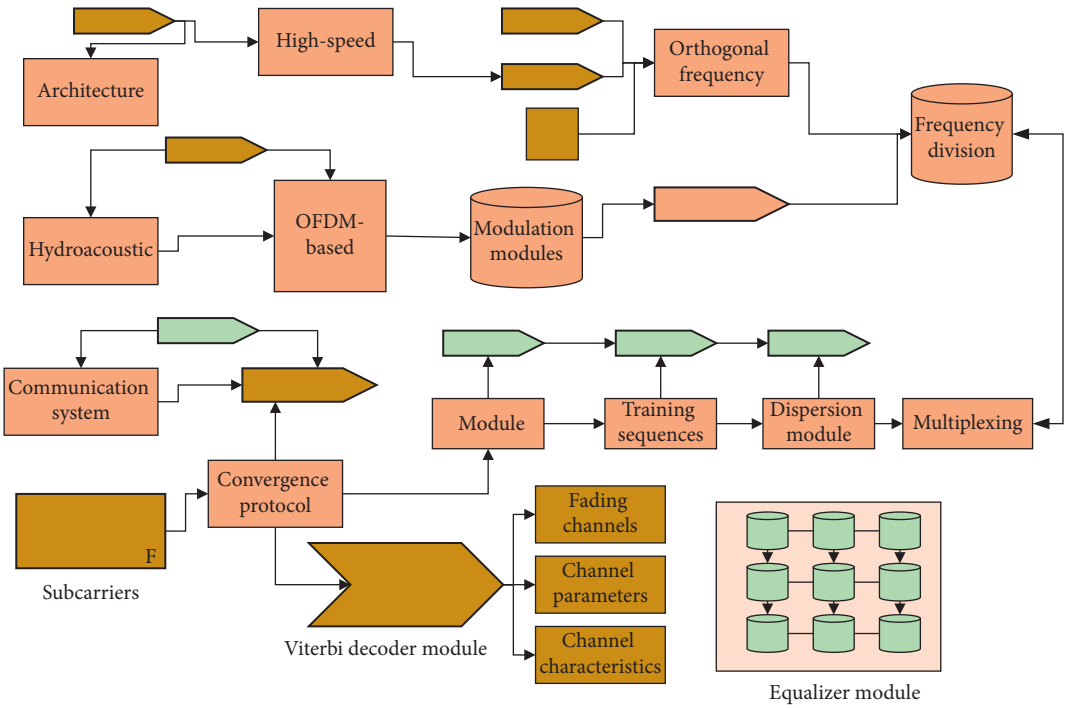


Figure 1: Structure of high-speed hydroacoustic OFDM-based communication system.

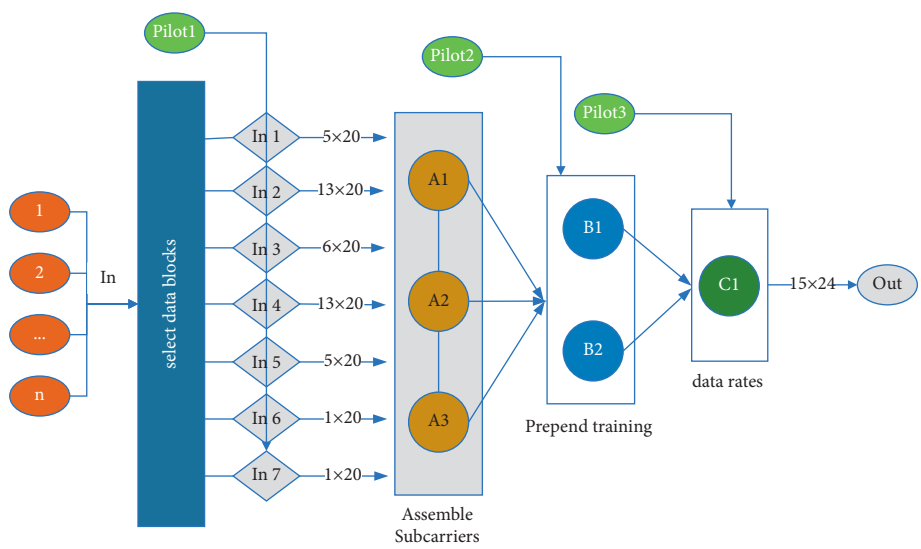


Figure 2: OFDM frame fusion module based on the structure of high-speed hydroacoustic OFDM communication system.

Since the output of the two ends of the receiver transducer is a differential signal with a DC bias, direct amplification of this differential signal with an uncertain reference level tends to overload the input of the amplifier circuit [15]. The error of our simulation and experimental results is small, less than 1%, which can illustrate the accuracy of our theory. Therefore, a differential to single-ended amplifier circuit is introduced to convert the differential signal into a single-ended signal whose reference level is the ground level in the circuit for poststage processing. The differential to single-ended circuit also suppresses the common-mode component of the output signal of the hydroacoustic transducer. The differential amplifier circuit is built using TI's low-power instrumentation amplifier INA118.

As shown in Figure 5, the Turbo encoder is usually designed to consist of two or more conventional decoders connected in parallel. To ensure that the data received by the encoder is statistically independent, an interleaver is added between the two conventional decoders.

The same is fed to the convolutional encoder in bit format. This encoder obtains this data, ready to be sent over the physical channel, by competing for information,

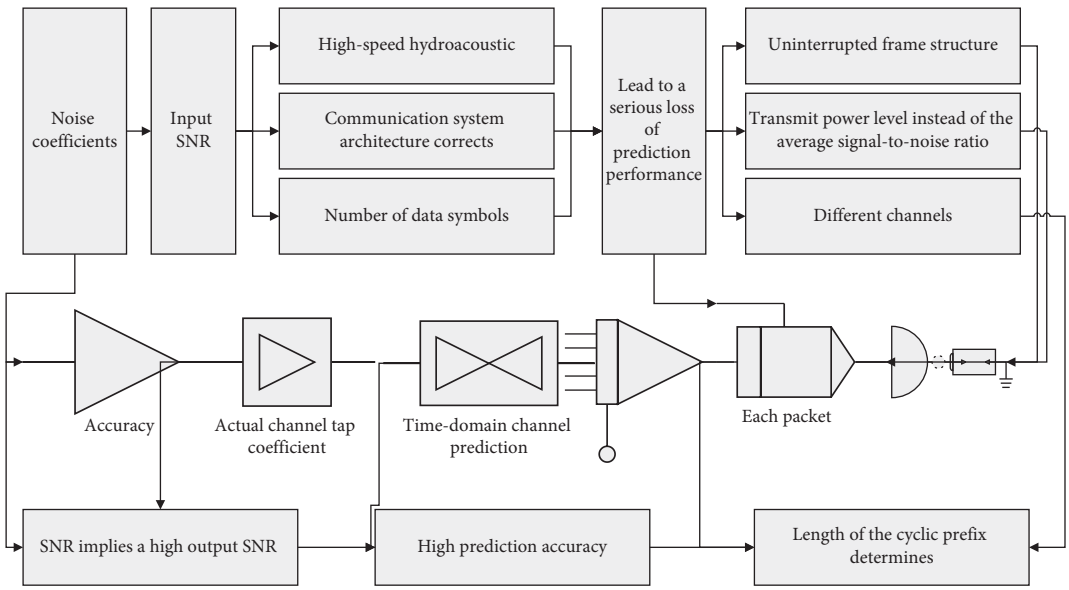


Figure 3: Frequency domain equalizer structure.

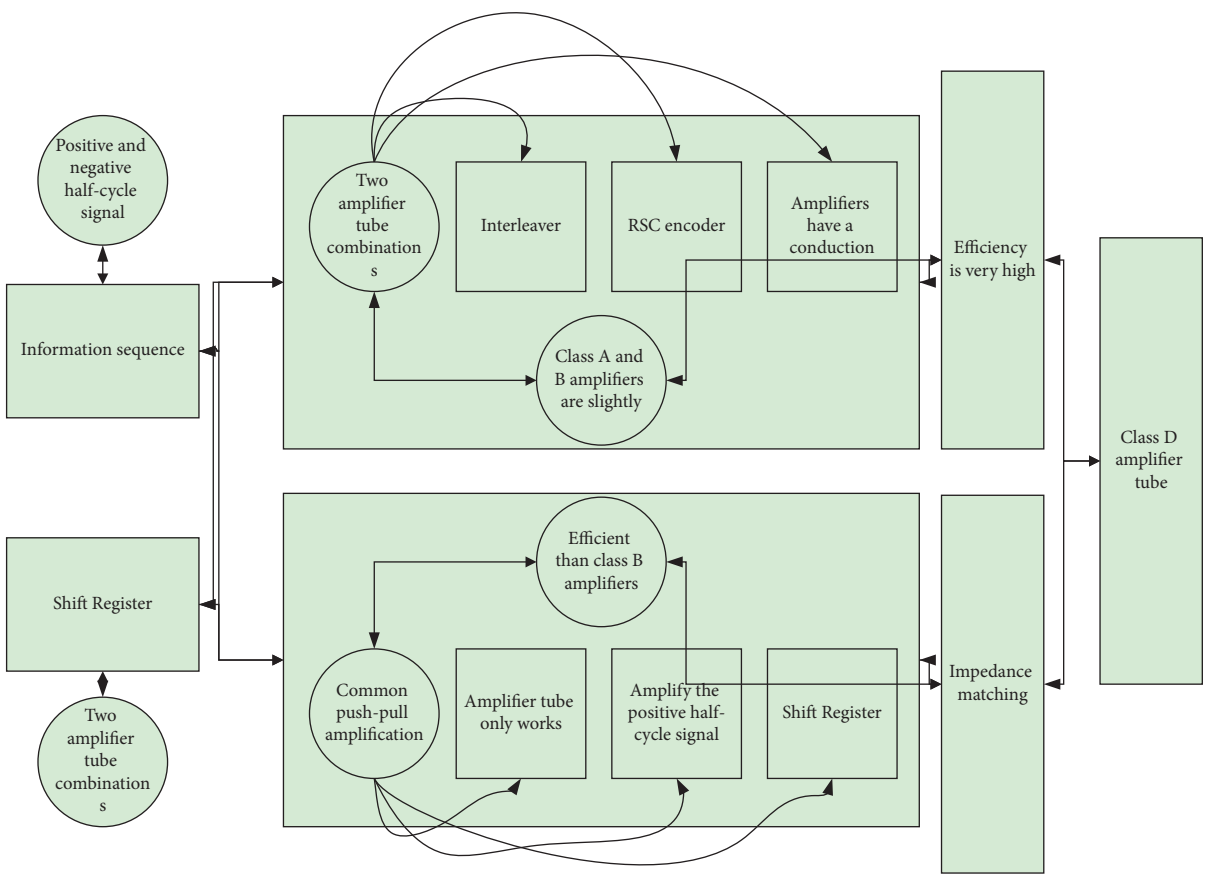


Figure 4: Parallel cascaded convolutional code encoder.

extracting information systems and recursive bits that can ensure that the data reach the end-user terminal accurately. Figure 5 depicts a schematic diagram of a Turbo cyclotron encoder consisting of two separate interleaved encoders.

The two independent check information sequences enter into the two-component decoders respectively, and each subdecoder decodes the system information sequence and check information sequence respectively to get different soft

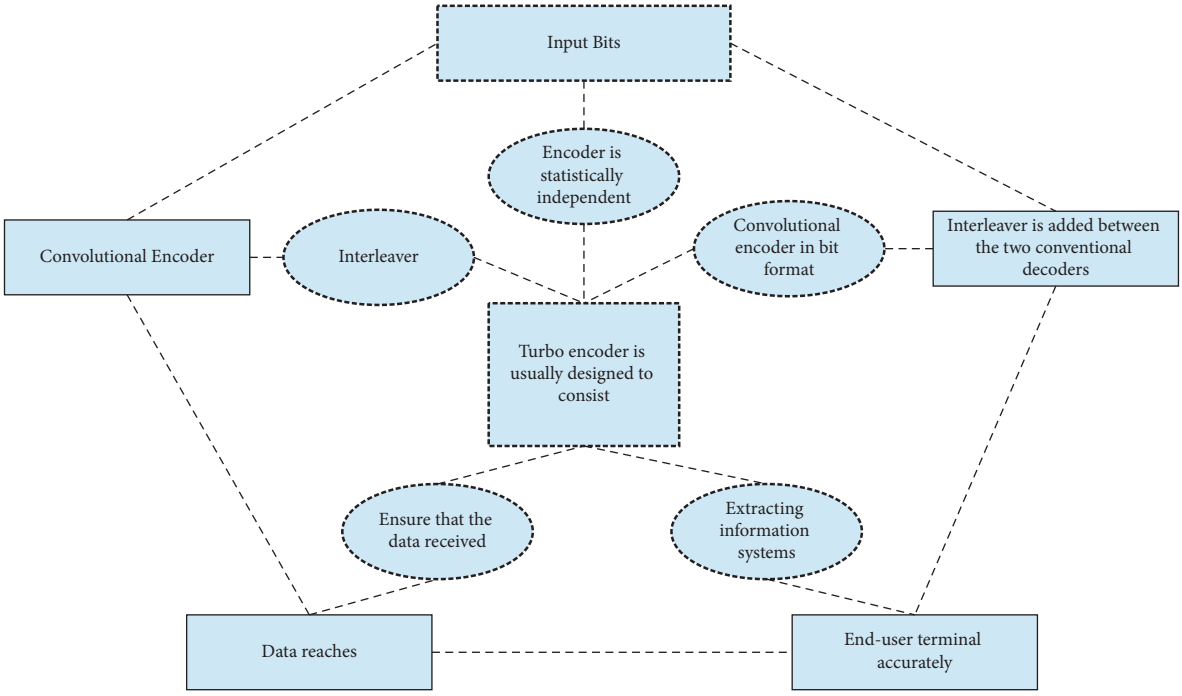


Figure 5: Structure of Turbo code encoder.

information about the information sequence, and then they exchange the relatively independent two sets of output soft information as to their respective new inputs, and after several iterations, the error bits in the received sequence are reduced to a certain degree, as well as the output soft information [16]. After several iterations, the error bits in the received sequence are reduced to a certain level, and the output soft information can be hard-judged as the decoded output so that a good error correction performance is achieved.

The encoded check bits are processed to produce codewords of different code rates. The component encoder can be either a recursive systematic convolution (RSC) code or a nonsystematic convolution (NSC) code. Given an RSC code, it is always possible to find an NSC code that generates polynomials corresponding to it and vice versa; see Figure 6 for the structure of the Turbo receiver of ISI channel.

Because the calculation of the forward state metric and backward state metric of each information bit in the log-map decoding algorithm of turbo code depends on serial recursive operations, the processing time of the whole decoding process increases significantly when using larger code blocks and interleaves that can bring better error correction performance, which causes significant processing delays and limits the channel compilation code rate [17]. When turbo coding is applied in a practical communication system, the turbo decoding algorithm needs to be improved by weighing the processing delay, system throughput, and hardware resource consumption. Common turbo decoding improvement structures include the sliding window decoding algorithm, block parallel decoding structure, and radix-8 algorithm.

An improved decoding method is to use sliding windows for decoding. The sliding window decoding algorithm divides a long block of code into multiple windows according to a preset window length and slides to each window in turn for decoding. In the processing of the sliding window algorithm, the initial values of the forward state metric and backward state metric of each window are used [18]. After analyzing the characteristics of Circular Permutation Code Words (CPC), taking the group interface of a communication private network as an example, the hardware circuit design scheme of the group CPC transceiver is introduced, and the generation of the CPC decoding table is given. The algorithm focuses on the module division, function and implementation method of the circuit, and the design and use of registers and briefly analyzes the detection performance of the circuit. The design is implemented by a field programmable logic device (FPGA), and the resource occupancy is low, has good portability, and can be used for similar interfaces after simple modification. It is a relatively common CPC transceiver circuit. By taking advantage of the convergence of the lattice recursive computation, the various unknown probabilities are in advance.

The basic structure of the Turbo code decoder is shown in Figure 7.

The hard judgment module is used to process the soft information of the final output of the second component decoder after several iterations to obtain the decoded output information bit sequence of the decoder [19].

4.3. Simulation Results for the Design of Turbo Code Coding Scheme in a High-Speed Hydroacoustic OFDM Communication. The simulation results of the Turbo code

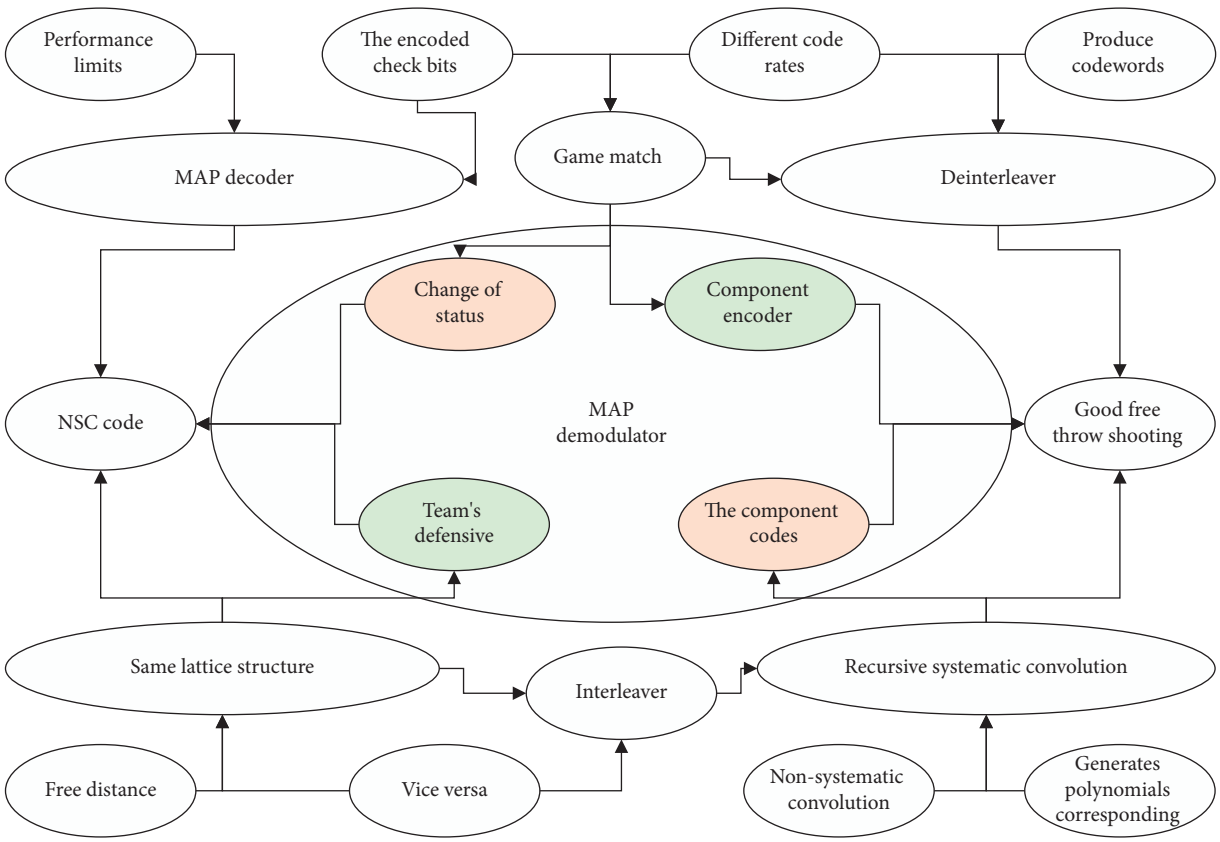


Figure 6: Turbo receiver architecture for ISI channel.

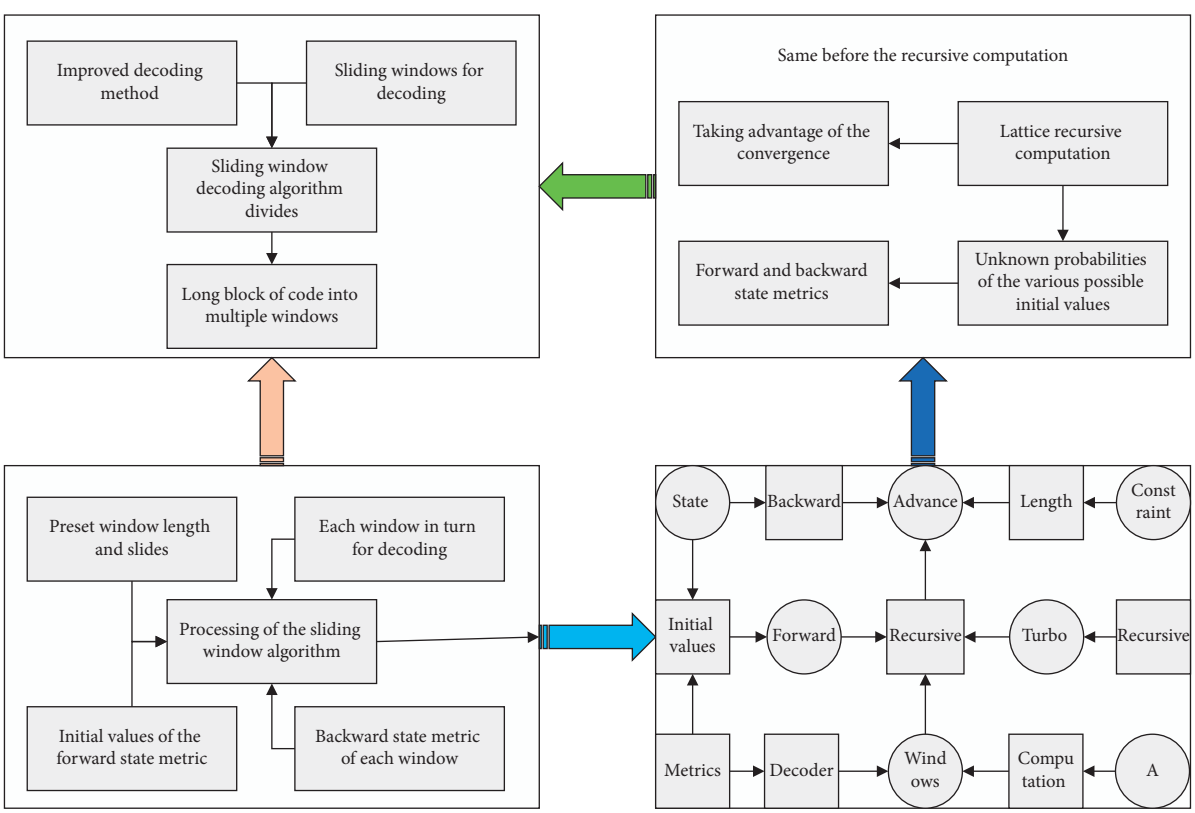


Figure 7: Block diagram of Turbo code iterative decoding structure.

coding scheme in high-speed hydroacoustic OFDM communication designed as described above are as follows. Per-packet BER, shown as a percentage packet error rate. For most of the packets, the BER is zero. See Figure 8 for simulation results of the high-speed hydroacoustic OFDM communication system based on Turbo codes, per packet BER.

When using the maximum frame length of N-6144 for turbo coding in the LTE standard, the decoding rate uses the fully parallel decoding algorithm. If shorter frame lengths are used, the decoding rate advantage of the fully parallel decoding algorithm decreases but is still much higher than that of the log-map algorithm. Although the fully parallel decoding algorithm can greatly increase the decoding rate and reduce the decoding latency, it will significantly increase the hardware resources required by the decode, and the fully parallel turbo decoder requires a large amount of computing and register resources. System sticks to LTE speed.

See Figure 9 for simulation results of high-speed hydroacoustic OFDM communication system based on Turbo codes-S/N ratio.

The fully parallel decoding algorithm with 36 iterations of approximate computation is shown in Figure 9 when a longer frame length of 4800 is used. The decoding performance of the max-log-map algorithm with 6 iterations and the fully parallel decoding algorithm with 36 iterations of approximate computation are very similar for different frame lengths.

Our system is relatively perfect, and it has almost no effect on the thermodynamics of water, so we do not need to consider the thermodynamics of water.

To avoid that the value of the state metric exceeds the range of the data bit width of the fixed-point quantization, which brings the decoder's error correction capability down, the hardware implementation adopts the method of normalizing the state metric to reduce the value of the forward and backward state metrics to ensure that they are limited to the range supported by the data quantization bit width. Within the range supported by the data quantization bit width, the backward and forward equalization of the received signal by the scatter plot gives an idea of the modulation being used by the system; see Figure 10 for simulation results of a high-speed hydroacoustic OFDM communication system based on unbalanced Turbo codes and Figure 11 for simulation results of a high-speed hydroacoustic OFDM communication system based on balanced Turbo codes.

The formation mechanism of multipath is again very different in shallow water environments, which consists of possible direct paths and reflections from the seafloor, and in deep water, which is caused by the bending of sound lines. The hydroacoustic communication in shallow sea channels is seriously affected by multipath effects and is prone to large impacts. In the shallow sea hydroacoustic communication, the sound velocity profile has a significant impact on the multipath effect, and Figure 10 gives several typical sound velocity profile cases. From the figure it can be seen that, in the sound velocity of the uniform sea, sound waves will reach the location of the receiving point through multiple surfaces

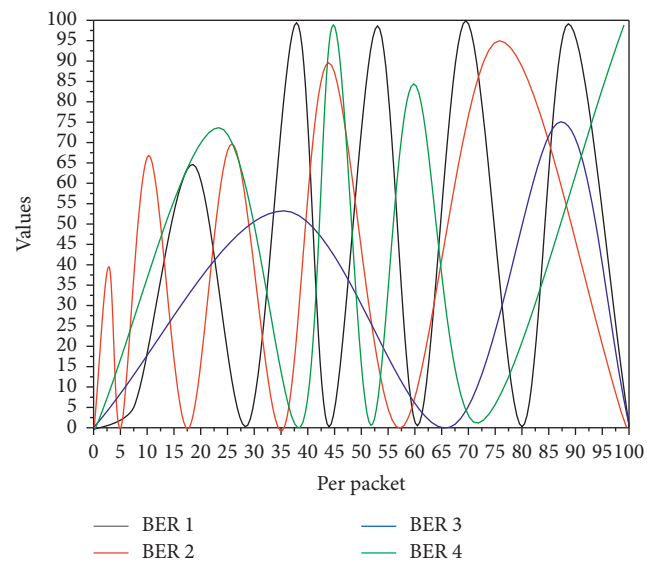


Figure 8: Simulation results of the high-speed hydroacoustic OFDM communication system based on Turbo codes, BER per packet.

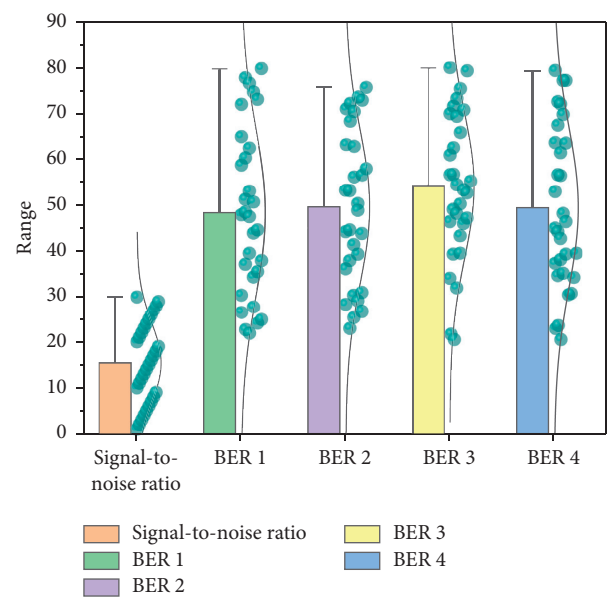


Figure 9: Simulation results of the high-speed hydroacoustic OFDM communication system based on Turbo codes-signal-to-noise ratio.

or seafloor reflections, and, in the negative gradient shallow sea, the seafloor reflection is the main mode of propagation, near the sound waves and can be reached through the straight line and surface path. In the long-distance, only part of the surface scattered signal and seafloor reflections can reach the location; for positive gradient shallow sea, sound waves reach the location of the receiving point. For positive gradient shallow sea, the acoustic wave reaches the location of the receiving point through multiple surface reflections.

The intercode interference caused by multipath seriously affects the performance of hydroacoustic communication

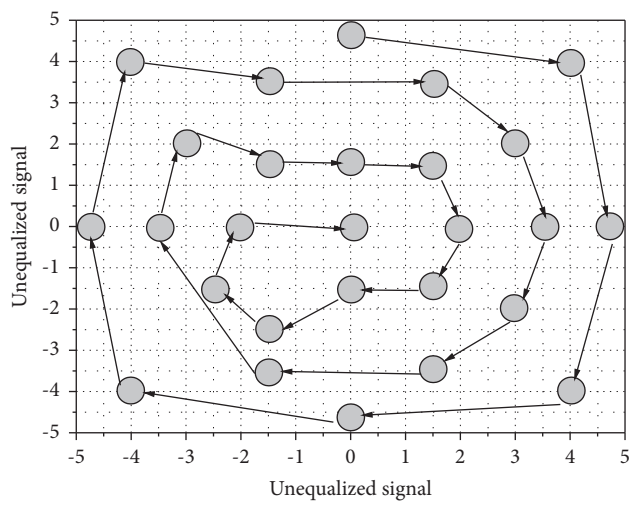


Figure 10: Simulation results of the high-speed hydroacoustic OFDM communication system based on unbalanced Turbo codes.

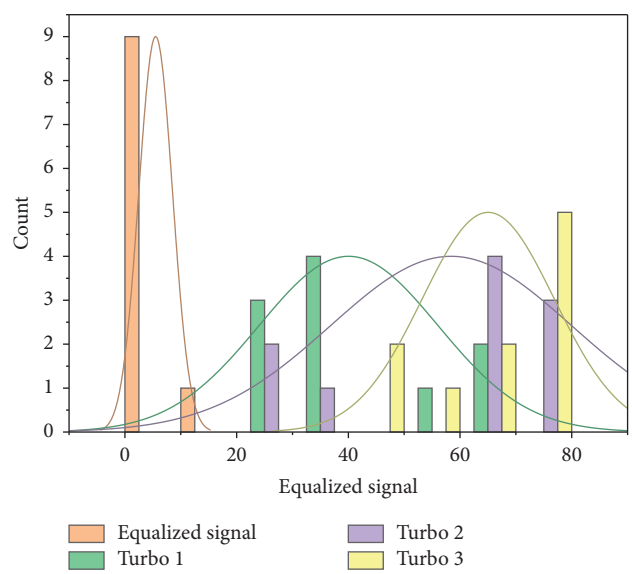


Figure 11: Simulation results of the high-speed hydroacoustic OFDM communication system based on balanced Turbo codes.

systems and causes great problems to ensure robust and high-speed hydroacoustic communication. The ISI in wireless communication is very short, often only a few code elements long. In contrast, the ISI of hydroacoustic communication is much larger, usually tens or even hundreds of code elements in length, which makes it more difficult or even impossible to distinguish the verdict at the receiving end. Therefore, to ensure robust and reliable hydroacoustic communication, the intercode interference caused by the multipath effect must be reduced or eliminated, and this is a major point and difficulty in the field of hydroacoustic communication.

For variables that are often applied throughout the algorithm, as well as input and output variables for DSP library functions used in the core algorithm, the storage location should be determined by the space they occupy. For example, allocate three segments of storage space for the operation of the DSP's built-in Fast Fourier Transform library function. The input and output arrays of the Fast Fourier Transform of length 16384 are stored in the form of alternating floating-point arrays, which occupy 13.1072 KB, so if all the input and output arrays of this function are allocated in the cache, the total space required is about 40 KB and should be placed in the L2 cache.

The SNR of the received signal can be estimated at the receiver side, and then the SNR of the received signal can be changed by adding additive Gaussian white noise, and the BER can be calculated under different SNRs, so the BER curve with SNR can be obtained, which is called the semisimulation experiment.

Figure 11 shows the BER performance comparison of the OFDM system with the precoding module, the black line shows the result of the semisimulation experiment without precoding at the transmitter, the rose line shows the result of the semisimulation experiment with ZF precoding at the transmitter, the green line shows the result of the semisimulation experiment with MMSE precoding at the transmitter, and the blue line shows the result of the semisimulation experiment with ZF-THP.

From Figure 11, after using the precoding technique at the transmitter side, only a simple detection process is required at the receiver side to achieve good performance. Moreover, the performance of the system with nonlinear precoding is better than that of the system with linear precoding.

The BER performance of the hydroacoustic STBC-SCFDE system with spatial and temporal diversity is better than that of the hydroacoustic SC-FDE system with a single transmitter and single receiver; in addition, the BER of the system decreases as the number of transmit antennas increases. The hydroacoustic STBC-SCFDE system is feasible and effective.

5. Conclusion

Through the above introduction of key technologies of high-speed hydroacoustic communication and the study of a hydroacoustic communication system with OFDM, the key technologies and advantages and disadvantages of OFDM application in a hydroacoustic communication system are analyzed, and a high-speed hydroacoustic OFDM communication coding scheme is designed based on TURBO codes. System simulation is conducted to analyze the data rate, unbalanced signal, balanced signal, received signal power spectrum, balanced signal power spectrum, the signal-to-noise ratio, bit rate, and per-packet BER results analyzed. Finally, the reliability of TURBO codes in high-speed hydroacoustic OFDM communication is analyzed with the simulation results of BER and SNR relationships under different conditions. Parametric modeling of the hydroacoustic OFDM system under impulsive noise interference is carried out. For the time domain impulse noise and large-amplitude single-frequency noise cancellation, a compression-aware technique is used, and the energy on the null

wave is used as a constraint for iterative cancellation of the time domain impulse noise and large-amplitude single-frequency noise, and the distributed cancellation method and the joint estimation cancellation method are compared. Simulation and experimental results show that the distributed cancellation method and the joint estimation cancellation method have good results for both time-domain impulse noise and large-amplitude single frequency noise cancellation, while the joint estimation cancellation method in which the large-amplitude single frequency noise cancellation is performed first has better performance.

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 College of Physics and Electrical Engineering, Kashi University.

References

[1] M. Stojanovic, "Recent advances in high-speed underwater acoustic communication," *IEEE Journal of Oceanic Engineering*, vol. 21, no. 2, pp. 125–136P, 1996.

[2] M. Stojanovic and J. Preisig, "Underwater acoustic communication channels: propagation models and statistical characterization," *IEEE Communications Magazine*, vol. 47, no. 1, pp. 84–89, 2009.

[3] H. C. Song, W. A. Kuperman, and W. S. Hodgkiss, "Basin-scale time reversal communications," *Journal of the Acoustical Society of America*, vol. 125, no. 1, pp. 212–217, 2009.

[4] G. A. Jones, D. H. Layer, and T. G. Osenkowsky, "ETSI EN 301 701 digital video broadcasting (DVB); OFDM modulation for microwave digital terrestrial television," *Language Arts & Disciplines*, vol. 13, 2013.

[5] M. Chitre, S. Shahabodeen, and M. Stojanovic, "Under water acoustic communications and networking: recent advances and future challenges," *Marine Technology Society Journal*, vol. 42, no. 1, pp. 103–116, 2008.

[6] S. Roy, T. M. Duman, V. McDonald, and J. G. Proakis, "High-rate communication for underwater acoustic channels using multiple transmitters and space-time coding: receiver structures and experimental results," *IEEE Journal Oceanic engineering*, vol. 33, 2007.

[7] B. Li, S. Zhou, M. Stojanovic, L. Freitag, and P. Willett, "Multicarrier communication over underwater acoustic channels with non-uniform Doppler shifts. Oceanic engineering," *IEEE Journal of Oceanic Engineering*, vol. 33, pp. 110–112, 2008.

[8] E. F. Sang and S. K. Xu, "Turbo codes in hydroacoustic OFDM communication," *Journal of Harbin Engineering University*, vol. 30, no. 1, 2009.

[9] J.-G. Huang, H. Wang, C.-B. He, Q.-F. Zhang, and L.-Y. Jing, "Underwater acoustic communication and the general performance evaluation criteria," *Frontiers of Information Technology & Electronic Engineering*, vol. 19, no. 8, pp. 951–971, 2018.

[10] P. Chen, Y. Rong, S. Nordholm, and Z. He, "An underwater acoustic OFDM system based on NI CompactDAQ and LabVIEW," *IEEE Systems Journal*, vol. 13, no. 4, pp. 3858–3868, 2019.

[11] M. J. Bocus, A. Doufexi, and D. Agrafiotis, "Performance of OFDM-based massive MIMO OTFS systems for underwater acoustic communication," *IET Communications*, vol. 14, no. 4, pp. 588–593, 2020.

[12] G. Qiao, Z. Babar, L. Ma, and N. Ahmed, "Channel Estimation and Equalization of Underwater Acoustic MIMO-OFDM Systems: a Review Estimation du canal et l'égalisation des systèmes MEMS-MROF acoustiques sous-marins: une revue," *Canadian Journal of Electrical and Computer Engineering*, vol. 42, no. 4, pp. 199–208, 2019.

[13] Y. X. Yun-Xiang Guo, X.-A. Song, R.-Q. Zhang, and H. Li, "Research on underwater acoustic communication system based on OFDM-OAM," *Automatic Control and Computer Sciences*, vol. 54, no. 6, pp. 541–548, 2020.

[14] C.-F. Lin, H.-H. Lai, and S.-H. Chang, "MIMO GS OVSF/OFDM based underwater acoustic multimedia communication scheme," *Wireless Personal Communications*, vol. 101, no. 2, pp. 601–617, 2018.

[15] M. R. Khan, B. Das, and B. B. Pati, "Channel estimation strategies for underwater acoustic (UWA) communication: an overview," *Journal of the Franklin Institute*, vol. 357, no. 11, pp. 7229–7265, 2020.

[16] M. Y. I. Zia, P. Otero, and J. Poncela, "Design of a low-cost modem for short-range underwater acoustic communications," *Wireless Personal Communications*, vol. 101, no. 1, pp. 375–390, 2018.

[17] G. Qiao, Z. Babar, L. Ma, and X. Li, "Cost function based soft feedback iterative channel estimation in OFDM underwater acoustic communication," *Infocommunications Journal*, vol. 11, no. 1, pp. 29–37, 2019.

[18] C.-X. Wang, J. Huang, H. Wang, X. Gao, X. You, and Y. Hao, "6G wireless channel measurements and models: trends and challenges," *IEEE Vehicular Technology Magazine*, vol. 15, no. 4, pp. 22–32, 2020.

[19] M. Y. I. Zia, J. Poncela, and P. Otero, "State-of-the-art underwater acoustic communication modems: classifications, analyses and design challenges," *Wireless Personal Communications*, vol. 116, no. 2, pp. 1325–1360, 2021.

Hindawi
Journal of Robotics
Volume 2022, Article ID 4065734, 8 pages
https://doi.org/10.1155/2022/4065734

Research Article

Target Detection of Low-Altitude UAV Based on Improved YOLOv3 Network

Haiqing Zhai [1,2] **and Yang Zhang** [1,2]

[1] *School of Computer Science and Technology, Henan Institute of Technology, Xinxiang, Henan 453003, China*
[2] *Big Data Engineering Research Center of Henan for Production and Manufacturing IoTs, Xinxiang, Henan 453003, China*

Correspondence should be addressed to Haiqing Zhai; zhaihaiqing@hait.edu.cn

Received 7 January 2022; Accepted 18 February 2022; Published 10 March 2022

Academic Editor: Shan Zhong

Most existing methods are difficult to detect low-altitude and fast-moving drones. A low-altitude unmanned aerial vehicle (UAV) target detection method based on an improved YOLOv3 network is proposed. While keeping the basic framework of the original model unchanged, the YOLOv3 model is improved. That is, multiscale prediction is added to enhance the detection ability of small-target objects. In addition, the two-axis Pan/Tilt/Zoom (PTZ) camera is controlled based on proportional integral derivative (PID), so that the target tends to the center of the field of view. It is more conducive to accurate detection. Finally, experiments are carried out using real UAV datasets. The results show that the mean average precision (mAP), AP50, and AP75 are 25.12%, 39.75%, and 26.03%, respectively, which are better than other methods. Also, the frame rate is 21 frames·s^{-1}, which meets the performance requirements.

1. Introduction

With the rise and development of UAV technology, it has been widely used in military and civil fields. However, a large number of UAVs pose a certain threat to the flight safety of aircraft and the political sensitivity of images in confidential areas. At the same time, it also brings huge challenges to urban security [1]. For the sake of public safety, local governments prohibit unauthorized UAV flight in airports, meeting places, and other areas [2]. Therefore, monitoring UAVs in specific areas is an urgent need for security. Due to the small volume and low speed of UAV, it is difficult to detect UAV by using traditional radar equipment [3]. In addition, the noise of the city is noisy, so it is difficult to detect UAVs by acoustic sensors [4]. Therefore, how to detect UAV efficiently and accurately from the environment has become an urgent problem to be solved.

So far, the identification methods of UAVs are diverse. The main methods focus on image recognition and radar data analysis [5]. However, traditional analysis methods that rely on spectrum detection and radar data are extremely susceptible to interference from external environmental factors. The method based on image recognition is limited by the bottleneck of computer technology and communication technology and has not been widely used and developed [6, 7]. At present, there has been a certain amount of research on UAV target recognition in China. For example, Thillainayagi and Senthil Kumar [8] proposed a target detection technology for UAV thermal images based on wavelet transform and singular value decomposition, using discrete wavelet transform and stationary wavelet transform to enhance image texture features and edge features. The experimental results show that the proposed method has smaller errors. However, the detection efficiency needs to be improved. Li et al. [9] proposed a fast and effective moving target detection method, which extracts cross features based on line segments. It has good detection speed and rotation accuracy, but the application scope of the algorithm is small and the practical application has great limitations. Abdulridha et al. [10] proposed a UAV hyperspectral image recognition method. Multilayer perceptual neural network and stepwise discriminant analysis are used to realize the

image detection, which effectively improves the detection accuracy. Yang et al. [11] proposed a UAV object detection model based on rotation and constant depth denoising in view of the difficulties of the multidirectional object, small pixel, and vibration interference of the UAV body in the process of aerial object detection. A selective search method is used to extract the region of interest in the aerial image, and the radial gradient of the region of interest is calculated. Combined with the deep denoising autoencoder, the original data noise is filtered out and the deep features are extracted to realize the detection of aerial image targets. However, deep-level feature extraction will increase the amount of calculation and affect the detection speed of the algorithm.

In addition, image-based target detection technology has gradually been widely used with the significant improvement of computing power and communication technology [12]. Xiaofei [13] proposes a UAV multitarget tracking and path planning method combining basic gray wolf optimizer and Gaussian distribution estimation. It overcomes the problem of real-time optimization of complex projects with traditional models and has good effectiveness and practicability. However, the algorithm focuses on target tracking and path planning, and its performance in target detection needs to be improved. Tao et al. [14] proposed a target search strategy for UAV based on reinforcement learning. Through the reinforcement learning training, the image captured by the drone is analyzed and processed to achieve target detection and tracking, but the detection effect of maneuvering targets is poor. Liu and Zhang [15] proposed an automatic vehicle detection method based on deep learning. Based on the interactive multimodel particle filter algorithm, the performance of the UAV for maneuvering target positioning has been significantly improved. However, the detection effect of small moving targets such as drones has yet to be verified.

Based on the above analysis, a target detection algorithm based on the improved YOLOv3 network is proposed for the problem of low-altitude UAV target detection. The innovations are summarized as follows:

(1) In order to make up for the inability of the YOLOv3 network to detect small targets, the proposed method adds multitarget prediction to improve the YOLOv3 network. Four different sizes of bounding boxes are provided to match the actual bounding box as much as possible, thereby improving the accuracy of target detection.

(2) Since drone monitoring is costly and difficult to implement, the proposed method uses a two-axis PTZ camera to track the drone. Among them, the PID algorithm is used to adjust the camera position to achieve efficient detection of low-altitude UAVs.

2. Target Detection Based on Improved YOLOv3

2.1. YOLOv3 Network Structure. The YOLOv3 algorithm uses a fully convolutional network composed of residual blocks as the backbone network. The network depth reaches 53 layers and is called Darknet-53. Its network structure is shown as in Figure 1. YOLOv3 draws on the idea of Feature pyramid network (FPN) and uses multiscale features for target detection. On the premise of maintaining the speed advantage, the detection accuracy is further improved, especially the detection ability for small targets is strengthened. Also, the detection effect of high-coverage images is significantly higher than that of YOLOv2 [16].

	Type	filters	size	output
	Convolutional	32	3*3	256*256
	Convolutional	64	3*3/2	128*128
1×	Convolutional	32	1*1	
	Convolutional	64	3*3	
	Residual			128*128
	Convolutional	128	3*3/2	64*64
2×	Convolutional	64	1*1	
	Convolutional	128	3*3	
	Residual			64*64
	Convolutional	32	3*3	256*256
8×	Convolutional	128	1*1	
	Convolutional	256	3*3	
	Residual			32*32
	Convolutional	512	3*3/2	16*16
8×	Convolutional	256	1*1	
	Convolutional	512	3*3	
	Residual			16*16
	Convolutional	1024	3*3/2	8*8
4×	Convolutional	512	1*1	
	Convolutional	1024	3*3	
	Residual			8*8
	Avgpool		global	
	Connected		1000	
	Softmax			

Figure 1: Network structure of YOLOv3 algorithm.

The YOLO series of algorithms constantly iteratively optimize the detection accuracy on the basis of always maintaining the advantage of high detection speed and gradually improve the detection ability of small targets. At the same time, accurate detection of high-coverage images is realized, the structure is simple, and the background false detection rate is low. It has become the most popular real-time target detection algorithm today.

2.2. Improved YOLOv3. In the detection and recognition process, when two or more types of similar objects appear, the YOLO network will often misidentify them as the same type. This shows that the YOLO network is too poor to distinguish the details. At the same time, the recognition performance of YOLO for objects with large difference in length and width ratio needs to be improved [17]. When the proportion of large and small objects in the image differs greatly, the large objects can be identified in the test results, but the small objects often cannot be detected [18, 19]. The UAV remote sensing image used in the experiment has a lower resolution, a larger image scale, rich information, and more complex details. Direct use of conventional network models for detection and recognition is often not ideal [20, 21]. Therefore, the model algorithm needs to be further improved according to the characteristics of UAV images.

For the above reasons and the analysis of the YOLO model, the YOLOv3 model is improved while maintaining the basic framework of the original model, that is, multiscale prediction is added.

In the YOLOv2 network model, in order to enhance the accuracy of small-target detection, the feature map extracted from the last layer of the network model is connected with the feature map of the previous layer through the pass-through layer. The size of the feature map of the last layer is 13×13. In YOLOv3, this method is further enhanced. YOLOv3 provides three bounding boxes of different sizes, using similar concepts to extract features of these sizes to form a pyramidal network. YOLOv3 adds several convolutional layers, and the final convolution layer is used to predict the tensor coding including boundary box, target in the box, and classification prediction [22]. YOLOv3 uses the feature fusion of multiple scales, so the number of bounding boxes is much more than before.

Adding a scale prediction to the YOLOv3 network model provides four bounding boxes of different sizes, as shown in Figure 2. The improved YOLOv3 network includes the last layer, and the feature map is 13×13. There are also 3 upsampled eltwise sums with feature maps of 26×26, 52×52, and 104×104. The largest network model uses the 104 feature map. Also, YOLOv2 takes multiscale into consideration for the training data sampling, and in the end, only the feature map of 13 is used. This should be the place that has the greatest impact on small goals [23]. At the same time, the input image size is changed to 1024×1024 to adapt to large-scale sampling. In the experiment, 12 clusters were selected. Then, divide the dimensional clusters evenly on bounding boxes of different sizes. The 12 clusters are (10×13), (16×30), (33×23), (30×61), (62×45), (59×119), (116×90), (156 × 198), (373 × 326), (312 × 536), (584 × 712), and (869 × 796).

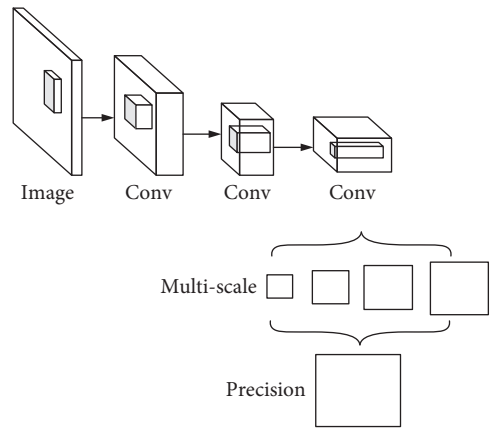


Figure 2: Multiscale prediction.

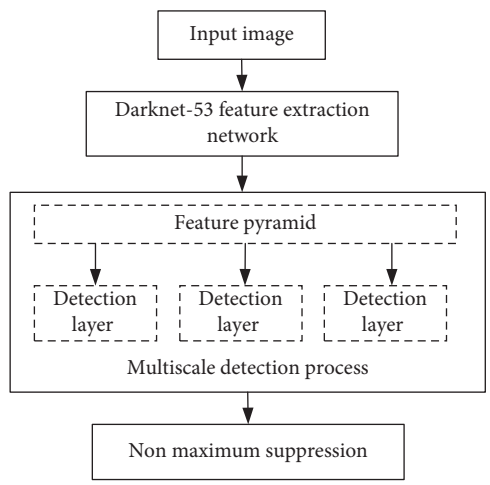


Figure 3: Detection process of YOLOv3 algorithm.

2.3. Detection Process. The improved YOLOv3 algorithm does not need to generate a region of interest (ROI) in advance but directly trains the network in a regression way. At the same time, the k-means algorithm is used to cluster the sample bounding boxes, and four groups of bounding box sizes are preset on the four scale sizes respectively, so as to make positioning prediction based on the bounding boxes of 16 sizes. The whole detection process is shown in Figure 3.

First, feature extraction is performed on the original 1024×1024 input image through the feature extraction network. Then, the feature vector is fed into the FPN structure to generate grid areas on 4 scales. They are 13×13, 26×26, 52×52, and 104×104. Each grid area predicts 4 bounding boxes, resulting in a total of (104×104 + 52×52 + 26×26 + 13×13)×4 = 57460 bounding boxes. Next, a vector Ω is predicted in each bounding box, which is expressed as follows:

$$\mathbf{\Omega} = \left(t_x + t_y + t_w + t_h\right) + l_0 + \left(l_1 + l_2 + \cdots + l_n\right). \quad (1)$$

The first 4 elements t_x, t_y, t_w, and t_h in the vector Ω are the 4 coordinates related to the bounding box, and their relationship is as follows:

$$\begin{aligned} p_x &= sigmoid\left(t_x\right) + b_x, \\ p_y &= sigmoid\left(t_y\right) + b_y, \\ p_w &= g_w \times e^{t_w}, \\ p_h &= g_h \times e^{t_h}, \end{aligned} \quad (2)$$

where b_x and b_y represent the offset of the grid to which the bounding box belongs relative to the upper left corner of the image. g_h and g_w represent the length and width of the predefined bounding box. p_x and p_y indicate the distance from the center of the bounding box of the final prediction result to the upper left corner of the image. p_w and p_h are the length and width of the predicted bounding box.

The l_0 of the vector Ω is expressed as follows:

$$l_0 = \Theta\left(o\right) \times \mathrm{IoU}_o^{gt}, \quad (3)$$

where $\Theta(o)$ represents the probability that the object is in the prediction frame. IoU_o^{gt} represents the intersection over union (IoU) of the predicted box and the true bounding box. When using logistic regression to score the prediction box the highest, the probability that the object is in the prediction box is 1, otherwise it is 0. The l_0 in vector Ω represents the

score that the predicted object belongs to one of the classes. When the prediction frame is obtained, the nonmaximum suppression is carried out to obtain the final prediction result.

3. Design of UAV Vision following Control Algorithm

The two-axis PTZ camera is shown in Figure 4. The function of the camera is to control the movement of the camera to keep the target in the center of the video. The control module is a two-axis PTZ with two steering gears. One servo is responsible for controlling the camera to move left and right, and the other controls the camera to move up and down. Each steering gear has an adjustment range of 180°.

The PID control algorithm is expressed as follows:

$$q(t) = \kappa_p \theta(t) + \kappa_i \int_0^t \theta(t')dt' + \kappa_d \frac{d\theta(t)}{dt}, \quad (4)$$

where $q(t)$ is the output of the system, which means the steering gear rotation angle, rad; $\theta(t)$ is the deviation angle between the image center and the UAV center, rad; κ_p, κ_i, and κ_d are all constant coefficients, corresponding to proportional gain, integral gain, and differential gain [24, 25].

Equation (4) is composed of three parts: proportional, integral, and differential. The first part makes the camera rotate with the movement of the drone. The integral part is used to eliminate the stabilization error and prevent the drone from being out of the video center. The derivative part is used to control the rate of change of the deviation [26].

The PID control process is shown in Figure 5. Using OpenCV (Computer Vision Library) to process the camera video stream, improve YOLOv3 detection for each frame of the video stream. After obtaining the position of the drone in the picture, calculate the distance between its center and the center of the picture. The distance parameter is passed to the PID process for calculation, so as to control the rotation of the steering gear.

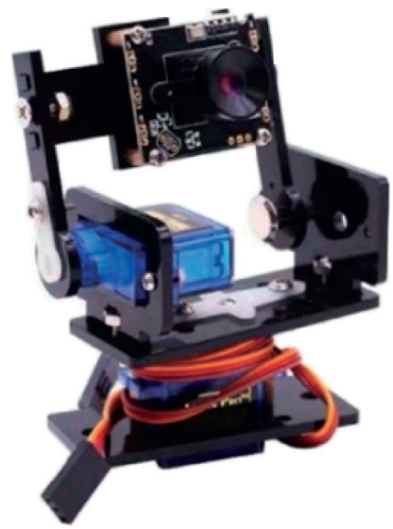


Figure 4: Two-axis PTZ camera.

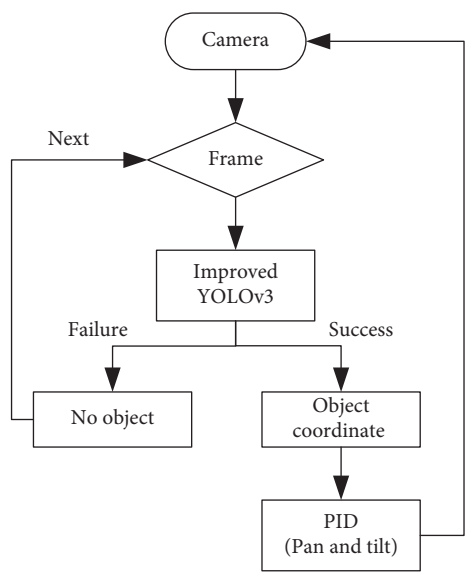


Figure 5: PID control process.

4. Experiment and Analysis

The experimental platform is a computer with Intel core i7-7700HQ2.8 GHz CPU and GeForce 1050 ti 2 GB GPU. The minibatch during training is set to 5, and the learning rate is 0.001.

4.1. Dataset. There are few low-altitude UAV detection and recognition methods based on deep learning, and there is no public dataset or standard dataset. Therefore, firstly, the dataset is collected and constructed.

4.1.1. Data Collection. Using visible light detectors and two-axis pan-tilt cameras to take images of 4 types of civilian drones at different times and in different backgrounds, the UAV models are DJI-Elf 3 (DJ-3), DJI-Yu Pro (DJ-Pro), DJI-Yu Mavic 2 zoom version (DJ-M2 Z), and DJI-Yu Air (DJ-Air). In order to ensure the diversity of data, the various flight attitudes of low-altitude drones, including hovering, rapid ascent and descent, and smooth flight, were fully considered during the shooting process. In the end, 3864 visible light images were obtained.

4.1.2. Data Annotation. In order to ensure the validity of the data, the manual labeling method is adopted to label the samples whose target occluded area is greater than or equal to 50%. Finally, the bounding boxes of the low-altitude drone targets in the 3258 visible light images were labeled, and data with labeled information were obtained. According to the ratio of 5 : 1, it is divided into the training set and test set.

4.1.3. Image Enhancement. In order to improve the detection and recognition accuracy of the method, the general image enhancement method in the field of target detection is

adopted to enhance the training set, including the operation processing of brightness and contrast.

4.1.4. Data Expansion. Taking into account that the attitude of the UAV during flight is not completely horizontal and inclined, etc., the training set is flipped and rotated at ±10° and ±20°. If the target near the edge in the image is damaged or completely lost after the rotation processing, the sample data are discarded.

4.1.5. Dataset Construction. Through the enhancement and expansion of the images in the training set, three types of datasets are obtained. As shown in Table 1, the training set of UAV-A is composed of low-altitude UAV targets in the original image. The training set of UAV-B is composed of the original image (UAV-A) and the image after image enhancement processing. The training set of UAV-C is composed of UAV-A and UAV-B and their expanded samples (including data sets UAV-A and UAV-B). Among them, the test set uses the same sample for method verification.

4.2. Visual Control Field Test Experiment. The target UAV hovers at the center of the field of view of the cooperative UAV at a relatively long initial distance to make $\tilde{x}(t) = \tilde{x}', \Delta(t) < \Delta'$. In order to quantitatively measure the size error between the actual bounding box and the expected bounding box, the bounding box size error is defined as $\vartheta_{\Delta}(t) = (\Delta(t) - \overline{\Delta})/\overline{\Delta}$. The experimental result of step response is shown in Figure 6.

It can be seen from Figure 6 that when $t = 2.5$s, the $v(t)$ step response curve quickly stabilizes. However, there is still a certain steady-state error between the actual bounding box and the expected bounding box.

4.3. Comparison with YOLOv3 Algorithm Classification Effect. In order to describe the classification ability of the improved YOLOv3 algorithm for the dataset, the classification confusion matrix on the data set is calculated. The YOLOv3 classification confusion matrix and the improved YOLOv3 classification confusion matrix are shown in Tables 2 and 3, respectively.

It can be seen from Tables 2 and 3 that DJ-Air has the best detection effect. The accuracy obtained by YOLOv3 and improved YOLOv3 algorithm reached 92.74% and 93.26%, respectively. Compared with other types of UAVs, DJ-Air has obvious characteristics, irregular shape, and easy to distinguish. The detection accuracy of YOLOv3 algorithm for DJ-3, DJ-Pro, and DJ-M2 Z is 88.69%%, 92.41%, and 89.91%, respectively. The colors of DJ-3, DJ-Pro, and DJ-M2 Z are not obvious, and the characteristics have certain similarities. In addition, the DJ-Pro and DJ-Air detection results are better in the modified YOLOv3 detection results, with the accuracy reaching 92.51% and 93.26%. Mainly because these categories are quite different from categories other than themselves. Compared with other types of UAVs, the target features are obvious and easier to distinguish. The detection accuracy of DJ-3 and DJ-M2 Z are 91.74% and 91.98%, respectively. The image performance characteristics of DJ-3 and DJ-M2 Z are similar, so they are easy to confuse.

According to Tables 2 and 3, the classification effect of the YOLOv3 algorithm and the improved YOLOv3 algorithm on the dataset are obtained, as shown in Table 4.

It can be seen from Table 4 that the improved YOLOv3 algorithm has a certain improvement in the recognition and detection effect compared with the classic YOLOv3 algorithm. Compared with the YOLOv3 algorithm, the average detection accuracy rate is increased by about 1.5%. The improved YOLOv3 algorithm adds multiscale prediction, which can strengthen the detection of small targets. Therefore, DJ-Pro, DJ-M2 Z, etc., can be better distinguished.

4.4. Performance Comparison with Comparison Algorithm. In order to demonstrate the performance of the proposed method, comparison with [8, 11], and [14] was performed. Among them, the comparison experiment stage evaluates the models obtained by different mainstream target detection methods through the same data training. In the experiment, the average accuracy mAP, AP^{50}, and AP^{75}, and frame rate evaluation indicators were used to perform quantitative analysis of detection accuracy and detection speed. The results are shown in Table 5. Among them, AP^{50} is an effective index to evaluate the classification ability of the algorithm. The AP^{75} can reflect the ability of the detection frame to return to the position of the bounding box.

It can be seen from Table 5 that the proposed method has been greatly improved, and its mAP, AP^{50}, and AP^{75} are 25.12%, 39.75%, and 26.03%, respectively. This shows that the proposed improved YOLOv3 target detection framework shows better classification ability and higher frame regression accuracy. The main reason for the obvious improvement of detection accuracy is the use of multilevel feature maps with different scales for target prediction. This greatly improves the detection effect of various targets that change with the drone's viewing angle and flying height. In addition, multiscale target prediction can predict the position and shape of the candidate frame based on image features and generates sparse and arbitrary-shaped candidate frames, which more closely match the real target frame. In addition, the two-axis PTZ camera is controlled based on PID, so that the target tends to the center of the field of view, which is more conducive to target recognition. Thillainayagi and Senthil Kumar [8] realize target detection in UAV thermal image based on wavelet transform and singular value decomposition. The detection model is simple and easy to implement, but the detection accuracy in a complex environment is not high. Its mAP is only 16.14%. Yang et al. [11] proposed a UAV object detection model based on rotation invariant depth denoising. A deep denoising autoencoder is used to filter out the noise of the original data and extract the deep features to realize the target detection of aerial images. Compared with reference [8], its detection accuracy has been improved. However, it is difficult to accurately detect small target objects, and it is easy to cause

Table 1: Processing methods and sample numbers of different training sets.

Train dataset	Processing method	Number of samples
UAV-A	Original	3200
UAV-B	Original + image enhancement	6400
UAV-C	Original + image enhancement + data augmentation	65000

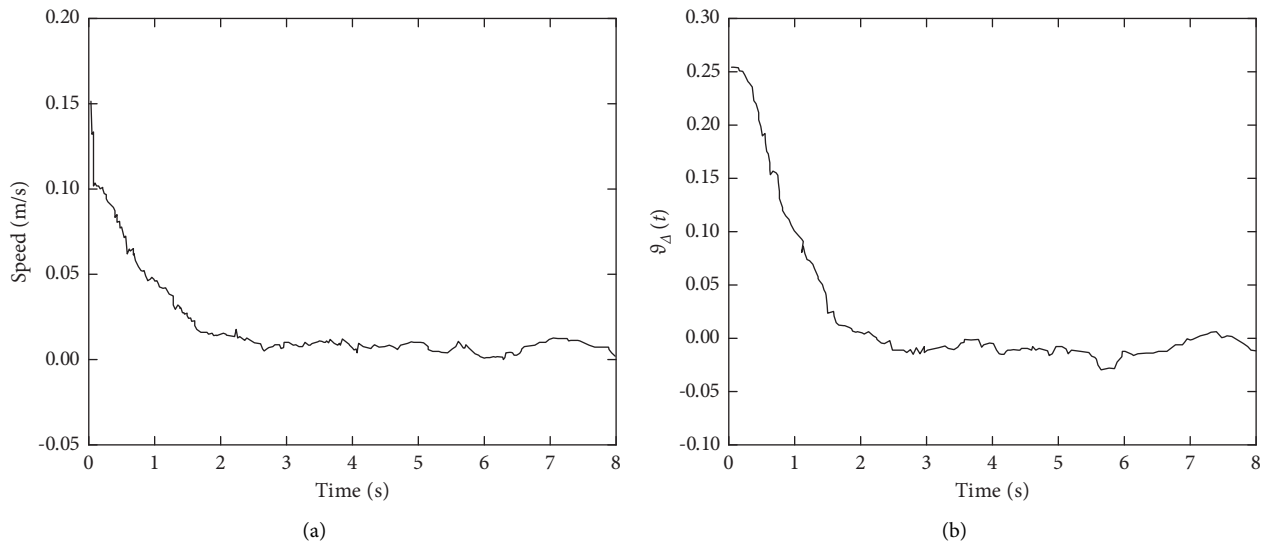


Figure 6: Experimental results of step response. (a) Step response curve of flight speed. (b) Boundary frame size error variation curve.

Table 2: Confusion matrix of YOLOv3 classification results.

Category	DJ-3	DJ-Pro	DJ-M2 Z	DJ-Air
DJ-3	6382	49	187	21
DJ-pro	274	2336	141	39
DJ-M2 Z	401	108	3625	67
DJ-air	139	35	79	1239
Accuracy rate, %	88.69	92.41	89.91	92.74

Table 3: Confusion matrix of improved YOLOv3 classification results.

Category	DJ-3	DJ-Pro	DJ-M2 Z	DJ-Air
DJ-3	6701	43	102	10
DJ-Pro	236	2286	117	26
DJ-M2 Z	295	107	3306	59
DJ-Air	72	35	69	1314
Accuracy rate, %	91.74	92.51	91.98	93.26

Table 4: Classification accuracy of two algorithms on datasets.

Category	YOLOv3	Improved YOLOv3
DJ-3	88.69	91.74
DJ-Pro	92.41	92.51
DJ-M2 Z	89.91	91.98
DJ-Air	92.74	93.26
Mean value of accuracy, %	90.96	92.37

Table 5: Comparative experimental results of different detection methods on UAV aerial photography data.

Method	Ref. [8]	Ref. [11]	Ref. [14]	Proposed method
mAP, %	16.14	22.37	23.64	25.12
AP^{50}, %	26.60	34.29	38.08	39.75
AP^{75}, %	13.62	19.65	24.71	26.03
Frame rate, frames·s^{-1}	12	19	36	21

confusion. Its AP^{50} is 34.29%. Tao et al. [14] use reinforcement learning to search for drones, and the detection process is easy to implement. Therefore, the detection speed is higher than the 21 frames·s^{-1} of the proposed method. Its frame rate is as high as 36 frames·s^{-1}. However, due to the lack of PTZ camera assistance, it is not effective in detecting fast-moving targets. Its AP^{75} is 1.32% lower than the proposed method.

5. Conclusion

The rapid development and application of UAV not only brings convenience to the society but also poses a serious threat to public security, personal privacy, and military security. It is becoming more and more important to find unknown UAVs quickly and accurately. Therefore, a target detection method of low-altitude UAV based on improved YOLOv3 network is proposed. The improved YOLOv3 network is used to extract the boundary box of the target UAV. The coordinate error and size error between the boundary box and the desired boundary box are calculated and input into the PID algorithm. Thus, the center moving angle of the two-axis PTZ camera is controlled to realize the accurate detection of the target UAV. The experimental results based on the constructed dataset show that:

(1) The improved YOLOv3 network has better detection capabilities for UAV targets and is easy to distinguish similar objects. Compared with the YOLOv3 algorithm, the average detection accuracy rate is increased by about 1.5%.

(2) The proposed method uses the improved YOLOv3 network to obtain the bounding box, and combines the PID algorithm to control the two-axis PTZ camera, which can better obtain the position of the drone. Its mAP, AP50, and AP75 are 25.12%, 39.75%, and 26.03%, respectively, and the frame rate is 21 frames·s^{-1}. The overall detection performance is the most ideal.

In the next work, we will obtain more drone data sets in scenarios for model training. A model with a faster speed and a smaller footprint will be selected for UAV identification. It will allow the model to get rid of the dependence on the graphics card, making the cost of the system more affordable.

Data Availability

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

Conflicts of Interest

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

References

[1] D. Chen, S. Sun, Z. Lei, H. Shao, and Y. Wang, "Ship target detection algorithm based on improved YOLOv3 for maritime image," *Journal of Advanced Transportation*, vol. 2021, p. 11, Article ID 9440212, 2021.

[2] J. Abdulridha, Y. Ampatzidis, and S. C. Kakarla, "Detection of target spot and bacterial spot diseases in tomato using UAV-based and benchtop-based hyperspectral imaging techniques," *Precision Agriculture*, vol. 21, no. 5, pp. 1–24, 2020.

[3] Z. Geng, R. Xu, and H. Deng, "LTE-based multistatic passive radar system for UAV detection," *IET Radar, Sonar & Navigation*, vol. 14, no. 7, pp. 1088–1097, 2020.

[4] J. Yu, Y. Liu, Y. Bai, and F. Liu, "A double-threshold target detection method in detecting low slow small target," *Procedia Computer Science*, vol. 174, no. 5, pp. 616–624, 2020.

[5] X. Chen, J. Guan, G. Wang, H. Ding, and Y. Huang, "Fast and refined processing of radar maneuvering target based on hierarchical detection via sparse fractional representation," *IEEE Access*, vol. 7, no. 7, pp. 149878–149889, 2019.

[6] T. Shi, H. Wang, W. Cui, and L. Ren, "Indoor space target searching based on EEG and EOG for UAV," *Soft Computing*, vol. 23, no. 21, pp. 11199–11215, 2019.

[7] L. Xiaoping, L. Songze, and Z. Boxing, "Fast aerial UAV detection using improved inter-frame difference and SVM," *Journal of Physics: Conference Series*, vol. 1187, no. 2, pp. 032082–032093, 2019.

[8] R. Thillainayagi and K. Senthil Kumar, "Combination of wavelet transform and singular value decomposition-based contrast enhancement technique for target detection in UAV reconnaissance thermal images," *Journal of Modern Optics*, vol. 66, no. 6, pp. 606–617, 2019.

[9] Z. Li, C. Meng, F. Zhou et al., "Fast vision-based autonomous detection of moving cooperative target for unmanned aerial vehicle landing," *Journal of Field Robotics*, vol. 36, no. 1, pp. 34–48, 2019.

[10] J. Abdulridha, Y. Ampatzidis, S. C. Kakarla, and P. Roberts, "Detection of target spot and bacterial spot diseases in tomato using UAV-based and benchtop-based hyperspectral imaging techniques," *Precision Agriculture*, vol. 21, no. 5, pp. 955–978, 2020.

[11] F. Yang, B. Ma, J. Wang, H. Gao, and Z. Liu, "Target detection of UAV aerial image based on rotational invariant depth denoising automatic encoder," *Xibei Gongye Daxue Xuebao/Journal of Northwestern Polytechnical University*, vol. 38, no. 6, pp. 1345–1351, 2020.

[12] B. Jla, J. A. Rong, and L. A. Wei, "High precision detection algorithm based on improved RetinaNet for defect recognition of transmission lines - ScienceDirect," *Energy Reports*, vol. 6, no. 2, pp. 2430–2440, 2020.

[13] W. Xiaofei, "A grey wolf optimizer using Gaussian estimation of distribution and its application in the multi-UAV multi-target urban tracking problem - ScienceDirect," *Applied Soft Computing*, vol. 78, no. 8, pp. 240–260, 2019.

[14] X. Tao, S. Hou, and Y. Wang, "Feasibility of reinforcement learning for UAV-based target searching in a simulated communication denied environment," *Scientia Sinica Informationis*, vol. 50, no. 3, pp. 375–395, 2020.

[15] X. Liu and Z. Zhang, "A vision-based target detection, tracking, and positioning algorithm for unmanned aerial vehicle," *Wireless Communications and Mobile Computing*, vol. 2021, no. 7, pp. 1–12, 2021.

[16] O. Doukhi, S. Hossain, and D.-J. Lee, "Real-time deep learning for moving target detection and tracking using unmanned aerial vehicle," *Journal of Institute of Control, Robotics and Systems*, vol. 26, no. 5, pp. 295–301, 2020.

[17] H.-J. Lim, S.-H. Choi, J. Oh, B. S. Kim, S. Kim, and J. H. Yang, "Adaptive ground control system of multiple-UAV operators in a simulated environment," *Aerospace Medicine and Human Performance*, vol. 90, no. 10, pp. 841–850, 2019.

[18] S. Pérez-Carabaza and J. Sch Erer, "UAV trajectory optimization for Minimum Time Search with communication constraints and collision avoidance," *Engineering Applications of Artificial Intelligence*, vol. 85, no. 11, pp. 357–371, 2019.

[19] O. Kechagias-Stamatis and N. Aouf, "A new passive 3-D automatic target recognition architecture for aerial platforms," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 57, no. 1, pp. 406–415, 2019.

[20] D. Cavaliere, V. Loia, A. Saggese, S. Senatore, and M. Vento, "A human-like description of scene events for a proper UAV-based video content analysis," *Knowledge-Based Systems*, vol. 178, no. 8, pp. 163–175, 2019.

[21] A. A. d. Santos, J. Marcato Junior, M. S. Araújo et al., "Assessment of CNN-based methods for individual tree detection on images captured by RGB cameras attached to UAVs," *Sensors*, vol. 19, no. 16, pp. 3595–3603, 2019.

[22] M. Rosamilia, A. D. Maio, and A. Aubry, "Single-pulse simultaneous target detection and angle estimation in a multichannel phased array radar," *IEEE Transactions on Signal Processing*, vol. 68, no. 7, pp. 6649–6664, 2020.

[23] Z. Li, Q. Zhang, and T. Long, "Ship target detection and recognition method on sea surface based on multi-level hybrid network," *Journal of Beijing Institute of Technology (Social Sciences Edition)*, vol. 30, no. 2, pp. 1–10, 2021.

[24] E. Dong, B. Han, and H. Jian, "Moving target detection based on improved Gaussian mixture model considering camera motion," *Multimedia Tools and Applications*, vol. 79, no. 11, pp. 7005–7020, 2020.

[25] M. Li, G. Sun, J. Tong, and Z. He, "Training-free moving target detection with uncertain a priori knowledge for airborne radar," *IET Radar, Sonar & Navigation*, vol. 14, no. 3, pp. 372–380, 2020.

[26] A. Tz, A. Zp, and W. A. Hao, "Infrared small target detection via self-regularized weighted sparse model - ScienceDirect," *Neurocomputing*, vol. 420, no. 3, pp. 124–148, 2021.

Hindawi
Journal of Robotics
Volume 2023, Article ID 9860830, 1 page
https://doi.org/10.1155/2023/9860830

Retraction

Retracted: Research on the Detection Countermeasures of Telecommunication Network Fraud Based on Big Data for Killing Pigs and Plates

Journal of Robotics

Received 19 September 2023; Accepted 19 September 2023; Published 20 September 2023

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

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

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

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

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

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

References

[1] G. Li and Y. Wen, "Research on the Detection Countermeasures of Telecommunication Network Fraud Based on Big Data for Killing Pigs and Plates," *Journal of Robotics*, vol. 2022, Article ID 4761230, 11 pages, 2022.

Hindawi
Journal of Robotics
Volume 2022, Article ID 4761230, 11 pages
https://doi.org/10.1155/2022/4761230

Research Article

Research on the Detection Countermeasures of Telecommunication Network Fraud Based on Big Data for Killing Pigs and Plates

Gang Li[1] and Yong Wen[2]

[1]*Guangxi Police College, Nanning 530028, Guangxi, China*
[2]*School of Artificial Intelligence, Guangxi University for Nationalities, Nanning 530006, Guangxi, China*

Correspondence should be addressed to Yong Wen; 20190304001@stu.gxun.edu.cn

Received 2 December 2021; Revised 18 January 2022; Accepted 12 February 2022; Published 7 March 2022

Academic Editor: Shan Zhong

This paper uses big data analysis to conduct in-depth analysis and research on the investigation countermeasures of the pig-killing dish telecommunication network fraud. This paper takes the case of "piggy bank" as a perspective, starting from analyzing the concept and characteristics of "piggy bank" type network fraud, analyzing the current phenomenon of this type of network fraud crime, and analyzing the current situation of the relevant laws, to explore the judicial situation regarding this type of network fraud crime, and to explore the problems and dilemmas in the judicial determination of such network fraud crimes, the determination of evidence, the determination of principal and accessory in joint crimes, the determination of crime amount, the determination of attempt and attempt form, and the determination of related crimes. Based on the analysis of the problems, the judicial determination process of "piggyback" network fraud is improved in terms of perfecting the evidence determination process, combining the facts of the whole case to determine the joint crime, accurately characterizing the associated criminal acts, regulating the judicial determination rules of the crime amount, and strictly applying the criteria of the circumstances of attempt and completion. This paper explores the forward shift of investigation work, anticipating telecommunication network fraud crimes in advance or the middle of the matter through active investigation, taking technical countermeasures, especially early warning and dissuasion to avoid the occurrence of telecommunication network fraud crimes, and reversing the unfavorable situation of passive investigation in the past. Finally, it focuses on the strategy and innovation points of the use of network investigation methods in telecom fraud crimes and proposes the investigation method of tracing the money chain and information chain along the line and cross-complementing the two lines at the same time according to the characteristics contained in telecom fraud crimes, relying on multiple groups such as public security, hackers, and network volunteers to jointly combat in a variety of network environments.

1. Introduction

With the information technology revolution sweeping up, the digital economy is growing at a geometric rate, and science and technology have penetrated all human production life and social interaction, and people are not bound to the traditional face-to-face communication at all, but more often choose noncontact information exchange [1]. At the same time, following the development of communication and network media, telecommunication network fraud is rapidly spreading and expanding, not only infringing on the most basic economic interests of the people but also endangering the sustainable development of the economy and society, increasing the risk and difficulty of social management, and the governance of telecommunication network fraud, a hot issue of people's livelihood, is gradually highlighting its importance and urgency in the field of social management. How to effectively manage telecommunication network fraud has become a social issue that needs to be urgently addressed by government departments at all levels [2].

The first half of 2020 was affected by the epidemic and the economic downturn in all sectors. The pressure on employment increased, the epidemic made it impossible for people to leave their homes, and people shifted their productive lives to the Internet, making telecommunication network fraud crimes frequent [3]. While the Internet is enriching people's lives, Internet technology is also being used by criminals, which makes traditional contact crimes decline and noncontact crimes represented by telecommunication network fraud crimes occur frequently. Nowadays, telecommunication network fraud, as a highly prevalent crime, is also making renovations and changes in line with the development of society. "Impersonating public prosecutors and law enforcement officers" is the primary version of telecom network fraud. Criminals continue to upgrade the criminal techniques on this basis; "loan lending" has become an upgraded version of telecom network fraud [4]. When "lending" became the key target of public security authorities, criminals designed new fraudulent ways according to the new needs of society; thus the "piggy bank" telecom network fraud came into being. The "piggy bank" type network fraud refers to the perpetrator to "love" "dating" and other means to deceive the victim's trust, with a variety of reasons to induce the victim into the precontrolled false. The victim is then lured into gambling and investing on a false website or platform for various reasons and then cheats the victim out of money. Due to the rapid spread and continuous renovation of online and noncontact fraud, telecommunications network fraud is no longer just an illegal and criminal act but has become a hot spot of social concern and difficulty in governance. As an upgraded version of telecommunication network fraud, "piggy bank" is more concealed and more socially harmful. In recent years, the "piggy bank" type of network fraud cases has been a high trend and has become one of the main ways of committing telecommunication network fraud. However, because the crime is characterized by a variety of criminal acts, concealed criminal means, and a long-lasting crime, the effectiveness of the public security authorities in combating and preventing it is not obvious [5].

In such an environment, to effectively curb and prevent the continued high incidence of telecommunication network fraud cases, society has made more efforts to strengthen institutional supervision, enhance technical support, and optimize combat effectiveness. However, due to the rapid spread and continuous renovation of network-based and noncontact fraud forms, telecommunication network fraud is no longer just an illegal criminal act but has developed into a hot spot of social concern and a difficult point of governance [6]. There is still a large gap between the actual effectiveness of telecom network fraud governance and the current situation of crime, economic losses, and the expectations of the public. How to break the current predicament faced by governance from the perspective of actual combat is a topic worth studying.

2. Related Works

Internet fraud has been around since when the knowledge of contemporary telephone communication became widespread in the developed countries of the West, and with the advancement of telephone communication, frauds can be committed without the risk of "face-to-face contact," by impersonating identities over the telephone and setting up various frauds to make illegal profits [7]. The high incidence of such cases has led to serious damage to citizens' money, so most Western countries have actively engaged in research and study of this issue in the fight against Internet fraud. One of the effective ways they have adopted is to increase the management of the real-name system so that the privacy of each citizen can be protected to the maximum extent and at the same time, establish relevant databases to reduce the crime of Internet fraud at the source [8]. Not only that, but foreign countries have also made more specific elucidation around the legal punishment of network fraud. For example, the US federal law on such cases is specified, according to the actual amount of fraud to develop the corresponding sentences, to prevent the emergence of judicial corruption in the sentencing process. In addition, the FBI has also cracked down on cyber fraud, drawing on the relevant statistical reports of the US FCC, and has implemented monitoring and data analysis of ex-convicts in many countries and regions around the world and has created cyber fraud complaint centers, relying on huge data for comparison and analysis, so that criminals have nowhere to hide. It has also prepared severe legal penalties. In addition, severe legal penalties have been prepared to suppress fraudulent behavior [9]. The United States is the first and most advanced country in terms of network construction, and it has a wealth of practical experience in the management of cyber fraud, and many countries such as Japan have also actively borrowed from the United States in their legal management of cyber fraud.

The "mobile phone in the hand, the world I have" is a true picture of people's travel life, with a variety of clothing, food, housing, and transportation APP online, going out without a wallet with a mobile phone has become a habit of life for most people; increased suspects aim at this new direction, depending on the popularity of mobile payment and online shopping [10]. They impersonate online shopping platform customer service, to "after-sales service," "inferior recall," and other reasons, to lure the victim by remodifying the payment order and other methods to transfer the money in the bound bank card, which is the most typical crime of telecom fraud crime, one of the modus operandi. The investigation will always lag the crime, and from the practical side, as the modus operandi of telecom fraud crimes continues to be renovated, the number of its cases also continues to climb, but the detection rate is generally low [11]. At this stage, suspects use increasingly developed communication technology and fast payment means to create a variety of frauds to implement fraud in a noncontact state, and telecom fraud crime cases are still in a high incidence and low detection situation. The earliest academic research on network fraud should have originated in Taiwan when the research was conducted on the spread of network fraud cases from Taiwan to the mainland coast [12]. The continuous development of network technology has led to a high incidence of network fraud as well. By studying and finding the similarities between different incidents of cyber

fraud crimes, many researchers aim to start with the theory and explore the underlying causes of the composition of this crime type and the ways to deal with it later, to propose strategies and recommendations to combat cyber fraud crimes [13].

3. Big Data-Based Model Designs for the Detection of Telecommunication Network Fraud in Piggy Banks

3.1. Building a Big Data Analytics Model. To deal with different cases of Internet fraud, government departments have continuously adjusted their governance programs, and public security authorities have gradually made improvements to their detection methods, grasping special remedies and making such cases a key target of their crackdown. The use of governance crackdowns across the country and extreme publicity has effectively curbed the spread of fraudulent activities [14]. The media, along with financial institutions in general, have used prominent signs to also inform the public about the various types of frauds and the dangers they pose to society. The coverage and publicity of telecom network fraud have substantially increased the public's awareness of prevention. However, each case is still a high-risk factor for cyber security. To solve and get out of the immediate dilemma of fraud, typical cases around the country need to be used as a vehicle to analyze the shortcomings and deficiencies of the country's public governance and bring reference to the relevant departments to improve the efficiency of governance. Reviewing the typical cases that have occurred around the country, we find that there are certain commonalities behind these online frauds. From the analysis of the above fraud cases, their modus operandi is all relatively close. The fraudsters all impersonate their identities and use a kind of virtual scenario to commit fraud. Although the education and identity of the victims were different, ranging from high school graduates to finance and accounting staff of famous enterprises to university graduates, they all came across frauds of a bad nature. In addition, two of the victims paid with their lives for this. From this, it is enough to recognize that frauds are very harmful to this society. This is since telecommunication network frauds are now pervasive and they will always find ways to find their targets.

$$\max h_{j0} = \frac{u_r y_{rjo}}{v_i x_{ijo}}. \tag{1}$$

The New Crown epidemic in the early part of the 20th century broke the routine of the entire society. "With the outbreak of New Crown pneumonia across the country and even the world, fraudsters took advantage of the global shortage of antiepidemic supplies to perpetrate a new wave of frauds, taking advantage of the nervous urgency to procure masks, alcohol, and other supplies. The fraud during the "New Pneumonia" epidemic is a fraud that is "tailored" to the uniqueness of the events taking place during a specific period [15]. It has its uniqueness compared to the frauds of normal life, and during this period, this fraud was significantly more numerous, significantly more successful, and significantly more socially dangerous. In terms of the main target, these fraud cases are becoming increasingly precise in terms of their implementation. Long before committing a crime, criminal groups have basic information about the victim from different sources, including occupation, finance, etc. The inadequacy of the personal information confidentiality system has allowed these fraudsters to find opportunities. From the follow-up review, most of the fraudsters use the Internet to collect the information of the victims and combine their identity and occupation to implement the fraud. In this regard, this paper designs a crime case analysis model based on the big data of previous telecommunication frauds, as detailed in Figure 1.

To effectively carry out the prevention and combating of telecommunication network fraud, the Ministry of Public Security has summarized and announced 48 common telecommunication network frauds to society. In practice, the Ministry of Public Security's platform for investigating telecom network fraud cases divides telecom network fraud into 19 categories and 53 subcategories. However, there are so many different types of frauds that the above figures cannot cover all types of telecom network frauds. In addition, under the vigorous crackdown by public security organs, criminals continue to renovate and upgrade their modus operandi, and many new types of frauds such as "piggy banks" appear and grow, and the types of telecommunication network frauds that are highly prevalent in different periods may vary. This paper intends to analyze the current high incidence of telecommunications network fraud types of modus operandi to solve the case to indicate the direction of the investigation and provides reference for targeted early warning, dissuasion, and accurate prevention. The "killing pig" type of online fraud means that after the perpetrator defrauds the victim's trust under the guise of "love" and "friendship," he induces the victim to gamble and invest in precontrolled fake websites and platforms for various reasons, and defraud the victim of money—a new type of cybercrime.

From the number of criminal cases of telecommunication network fraud in Texas in 2019, the top five cases were 718 brush-off frauds, 430 loan frauds, 342 investment and finance frauds, 162 refund frauds posing as shopping customer service, and 102 online dating induced gambling and investment frauds (kill pans). As shown in Figure 2, single-scaling scams topped the list.

From the 2019 Texas telecommunication network fraud cases loss amount, the loss occupying the top five is an investment and financial fraud 30,959,700 yuan, brush single type fraud 15,261,400 yuan, online dating induced gambling, investment fraud (kill the piggy bank) 12,986,800 yuan, loan fraud 9,357,200 yuan, and impersonating shopping customer service refund fraud 4,519,700 yuan. A comprehensive analysis of the above shows that the current high incidence of common types of telecommunication network fraud is swipe type fraud, loan type fraud, investment, and finance type fraud, online dating induced gambling, investment fraud (piggy bank), and impersonation of shopping customer service refund fraud.

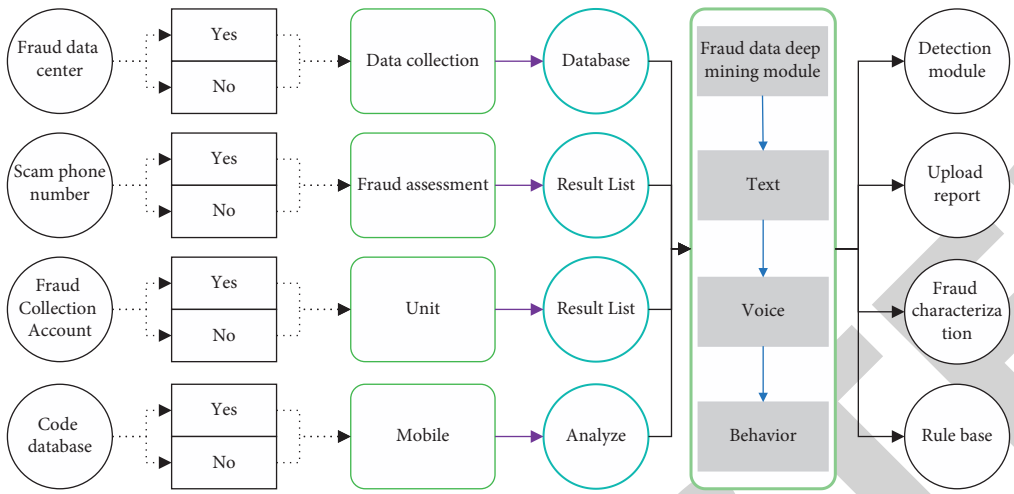


Figure 1: Flow chart of big data analysis of telecom fraud.

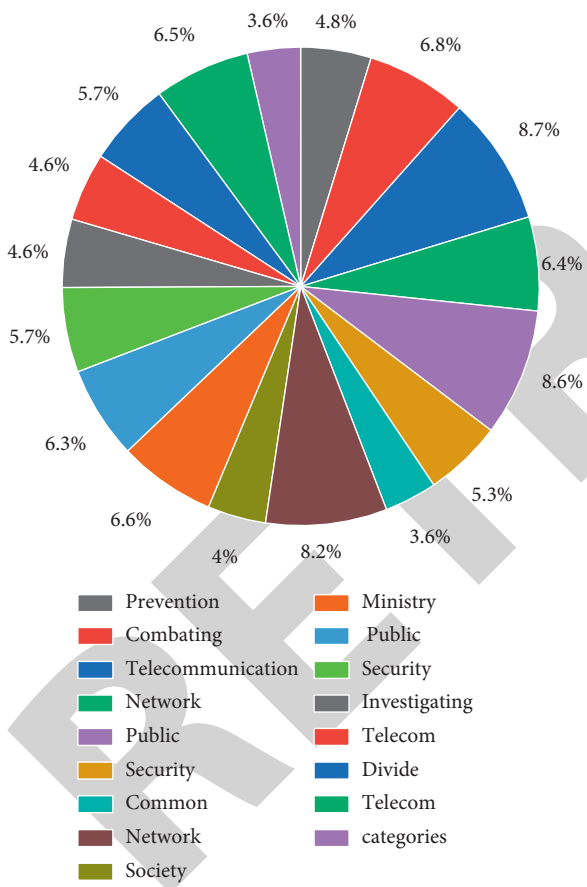


Figure 2: Statistical analysis of telecommunication network fraud cases.

$$\alpha = \frac{1}{v^t x_1^0},$$
$$\sum_{i=1}^{n} Y_{rj} - S_r^- = Y_{rj0}. \tag{2}$$

With the rapid development of the Internet, all kinds of e-commerce platforms, and rapid development and growth, online shopping is becoming increasingly popular and has gradually become an essential part of people's lives. When shopping online, the important reference for people to choose online stores is the sales volume and good reviews; some online, stores to improve sales volume and good reviews, they will hire "water army" to brush single; to enhance competitiveness, the "part-time job" of online brushing was born, many of which are professional white-collar workers and full-time earners with a certain amount of free time. The "part-time job" has emerged. Many have some free time; eager college students, professional white-collar workers, full-time mothers, etc., have joined the army of the brush single. At the same time, criminals also aimed at this opportunity, taking advantage of people being anxious to seek part-time jobs and greedy for cheap psychology, under the guise of recruiting network single personnel to implement fraud, brush single fraud has developed into the most frequent type of telecom network fraud crime.

$$X = 0.9 + 0.1\frac{X_i}{X_{\max} + X_{\min}}. \tag{3}$$

The victims of this type of fraud not only lose money, but also their feelings are cheated, and some people even have the idea of living lightly after being cheated. The suspects use the fake identity information purchased to register on the dating sites such as Lily.com, Century Jayquan, Zhenai.com, or add the victims as friends through short video social platforms such as Jitterbug, Raptor, Stranger, and other instant messengers such as QQ, WeChat and other nearby people functions, disguise themselves, pretend to be rich and handsome, rich and beautiful, successful people and other identities, in the name of love and dating contact with the victims, through a period of interaction to determine the relationship between men and women [16]. The fact is that the actual person who is in the position to get a good deal more than just a few of these is a lot more than just a few of these. At the same time, the suspects will tamper with the data in the background, creating the illusion that the victim is making money through investment, deceiving the victim

into using the website loophole to make money and then asking the victim to invest; once the victim falls for it, many people will take out all their funds to invest, or even go to borrow money or loans, with great losses.

With the convenience of express coordination and convenient mobile payment, online shopping has penetrated all aspects of people's lives, such as clothing, food, housing, and transportation. The suspects obtained the victim's online shopping or express information through illegal channels, posing as website customer service, saying that the quality of the victim's purchase has problems, can give the corresponding compensation, or posing as customer service, saying that the loss of express can give financial compensation, and then induce the victim to scan the QR code or click on the link to log into the so-called claims website, and wait until they get the victim's bank card account, password, mobile phone verification code, and other information; after the money will be transferred away, the victim's PayPal reputation points cannot be refunded for the reason of guiding the victim online loans, and induce the victim to transfer the loan to the designated account.

3.2. Analyzing Past Cases with the Help of a Big Data Model to Obtain the Characteristics of "Piggy Bank" Telecom Network Fraud. The characteristics of telecommunication fraud crime are determined by its attributes and development trends, which are summarized in this paper as the following five points.

3.2.1. Indirect Contact and Strong Anonymity. The biggest difference between telecom fraud, as one of the variants of fraud crimes, and traditional fraud crimes is that the suspect does not have direct contact with the victim, but indirect contact through telecommunication channels such as telephone, SMS, WeChat and VoIP, and the victim often does not have the suspect's identity information, nor does he or she meet with the suspect, making it difficult to confirm the identity. The economic downturn in all occupations, the increased employment pressure, and the epidemic have made people unable to go out, and people have transferred their production and life to the Internet, resulting in frequent telecom and network fraud crimes. This is the most significant feature of telecom fraud crimes, which is why telecom fraud is also known as remote, indirect contact fraud.

3.2.2. Gang Collaboration and Cross-Territory Operations. Telecom fraud crimes are mostly group fraud. From the origin of telecom fraud crimes, most frauds require multiple people to complete the fraud. The division of labor is clear and the grouping is fine in the process of fraud implementation. For example, fraud gangs posing as public security, prosecutors, and law are divided into script groups, online chat groups, talk groups, money laundering groups, withdrawal groups, etc. At the same time, telecommunications fraud crimes present ecological, familial, and geographical characteristics. A certain type of telecommunications fraud only exists in one place. The gang-like and family-style crimes have undoubtedly increased the difficulty and risk of hunting down the public security organs. It is difficult for the public security organs to wipe out a certain gang. This is also one of the direct reasons for the low detection rate of the current telecommunication fraud crime cases.

3.2.3. Low Starting Point, "A Book of Profit". The original wire fraud crime required little investment, and a mobile phone, a few mobile cards, and a good script were all that was needed to perpetrate a complete fraud. The "lucrative" nature of wire fraud has surpassed drug crimes, while the sentencing range is much lower than that of drug crimes, and the current penalties are not effective in deterring suspected wire fraud offenders.

3.2.4. Clear and Fine Division of Labor, High Degree of Organization of Criminal Gangs, and Strong Technology. Like the evolution of a virus, telecom fraud continues to evolve and change as the battle between good and evil continues, "progressing" with technological advances. While traditional telecom fraud only relied on phone calls, today's telecom fraud can rely on high-tech equipment such as cat pooling devices (sending and receiving mass SMS messages), SMS sniffing devices (forcing the collection of signals to create fake base stations), GOIP devices (hiding real numbers), and high-tech systems such as group control systems (one person controlling multiple mobile phone terminals), and Trojan horse theft systems [17]. The public has never even heard of these devices and systems, which directly leads to the fact that such frauds cannot be prevented and many times, they are cheated without realizing it. In telecom network fraud crimes, the suspects work alone rarely, and most of them work in the form of gangs with obvious professional characteristics. The internal organization of the gang is tight and the division of labor is clear, mainly divided into organizers, material dealers, car dealers, technical service groups, telephone and promotion groups, money laundering groups ("water room"), withdrawal groups ("riders"); each group is closely coordinated, but also independent of each other. To avoid crackdowns, some gangs adopt enterprise-based management, formulate strict rules and regulations, and require the separate use of crime cell phones and life cell phones, with the criminal chain becoming more and more elaborate. The organizational structure of a complete criminal gang is rough as follows (Figure 3).

3.2.5. Broad Victim Base. Criminal gangs will not only target a specific region or be limited to certain specific people but are based on fraud dens, using mobile phone group calling systems and SMS groupers to implement fraud on multiple people at the same time, with a large and scattered number of victims across the country. Regardless of gender, age, education, or work, any mobile phone or Internet user may be a potential victim of telecommunication network fraud

Figure 3: Organizational chart of the criminal group.

crimes. The base of mobile phone and Internet users in China is huge. As of June 2019, the scale of China's Internet users reached 854 million, the Internet penetration rate reached 61.2%, and the scale of online shopping users reached 639 million, accounting for 74.8% of the overall Internet users. Faced with a wide variety of fraudulent information, people are highly likely to fall into the trap of fraudsters and suffer property losses if they cannot accurately identify them.

$$R_p = \frac{N}{Q_{PL}}. \tag{4}$$

The high-tech features of telecom fraud crimes are currently reflected in the proficiency and use of sophisticated means of communication and electronic technology, and the suspects are more sophisticatedly equipped. Not only that, their knowledge base of psychology and other skills are also increasing, and the suspects are constantly updating their existing scripts by reading, analyzing, and deducing the possible psychological changes of their victims, and even marking even the tone aids and punctuation marks in their communication, which is amazing [18]. Due to the inconsistent level of access and proficiency in relevant information between the two parties to the transaction, where the dominant party may have more information, or exclusive and blocked information, while the disadvantaged party lacks sufficient necessary information, resulting in a situation where it is unable to make a full and effective decision, resulting in adverse consequences. In the prevention of bank telecommunication network fraud, banks are less autonomous in accepting government supervision for their interests, and the information available to the relevant regulatory authorities is not comprehensive enough to understand and confirm whether the corresponding regulatory policy requirements are implemented and put in place by banks, thus making it more difficult to assess the effectiveness of the relevant role of regulatory measures after implementation.

The leakage of citizens' personal information is serious due to the inadequate risk control measures of enterprises and institutions, lax internal supervision, and citizens' lack of awareness of information protection. Criminal gangs obtain victims' personal information through a variety of illegal means such as implanting Trojan horses, phishing websites, hacking, and black-market purchases, and then classify and develop different types of fraudulent scripts to implement precise frauds. In the types of frauds such as impersonating public prosecutors, impersonating customer service refunds, subsidized tax refunds, ticket changes, impersonating triads, etc., when the criminals accurately say the victim's name, ID number, home address, shopping orders, express delivery numbers, records of buying houses and cars, airline ticket orders and other information, some victims will relax their vigilance, reduce their ability to recognize, and are extremely vulnerable to being cheated.

Unlike other crimes, telecommunication network fraud requires the cooperation of the victim to be completed. To get the victim's cooperation, the suspect will prepare the script carefully in advance, master the fraudulent words, and take advantage of people's tendency to avoid harm, adopt bullying, and to lure the victim into believing their fabricated lies, and once the victim fails to recognize the fraud, he will

fall into the trap and transfer money as requested by the suspect or provide his bank card number, password, verification code, and other information to the other party. In addition, anyone has cognitive limitations; some victims lack basic legal knowledge and unfamiliarity with relevant national policies, airline ticket change process, shopping platform claims, etc., and do not understand the online loan process and the procedures required, resulting in serious information asymmetry between the victim and the suspect, resulting in the victim easily believing the suspect's fictitious identity and fictitious facts. The possibility of being cheated is extremely high. In recent years, to enhance people's awareness of prevention and prevention ability, the relevant functional departments have been committed to fraud prevention publicity, but there are still problems such as single form, uniform content, insufficient coverage, and poor targeting, which cannot arouse the interest of the propagandized persons; there are blind spots for publicity; even though a lot of human, material, and financial resources have been invested, they cannot achieve the expected effect, coupled with the blind self-confidence of some victims. Even if they receive fraud prevention publicity, they do not pay attention to it.

$$Q = \sum_{i=1}^{n} LB_i. \tag{5}$$

4. Analysis of Results

4.1. Big Data Analytics Model Test Results. "Due to the popularity of online shopping and online payment, transaction fraud is the most prevalent type of fraud." The report also points out that consumers are becoming increasingly rational, in the face of "gifts" such as "pie in the sky" frauds; vigilance has increased significantl; pure profit frauds have decreased significantly, which is inseparable from the long-term unremitting publicity and guidance and targeted special treatment. Data show that the proportion of victims between the ages of 18 and 28 is as high as 54%, and the post-90s group has the highest probability of being cheated [19]. In this regard, the report analyzes that young people who are studying or have just entered society are prone to become the target of fraudulent techniques such as payment rebates, free shipping, low-cost lures, and part-time recruitment. Young people who have some work experience and a certain economic base have a higher proportion of frauds in pornography, online dating, and financial credit. For middle-aged and elderly people, the report points out that the amount of fraud is often higher, with the per capita number of victims over 45 years old being about 7,000 yuan, far more than other age groups. They are easily exploited and controlled when they encounter frauds about financial investment, health care, and network technology and invest large sums of money, which is also more difficult to persuade. The age distribution of victims of different fraud types is shown in Figure 4.

Fraud brings great harm and has a wide impact, especially the use of telecommunications networks to engage in

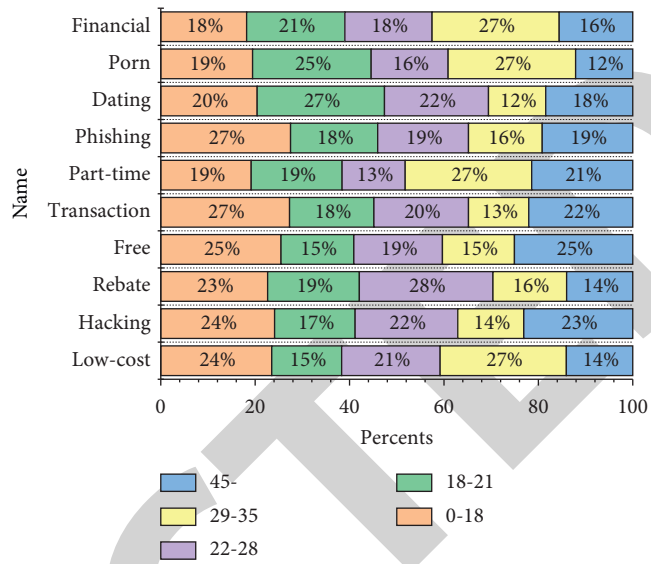


Figure 4: Distribution of age groups of victims of different fraud types in the last four years in County X.

illegal and criminal activities that are highly concealed; so, in the specific implementation process, it is necessary not only to strengthen the fight but also to do a good job of prevention, to achieve comprehensive governance. Governance units include banking institutions, payment organizations, the Internet, and so on, in addition to relevant government departments. But in the specific governance process, there is an obvious phenomenon of division, especially when part of the work belongs to the crossover of the functions of multiple departments, it is easy to produce gray management, there is no lead unit, and some work is not covered or paid attention to, resulting in a regulatory vacuum. With the rapid change of science and technology, following the development trend of communication and network media, the rapid spread and expansion of telecommunication and network fraud has not only infringed the most basic economic interests of the people, but also endangered the sustainable development of the economy and society. At present, there are 25 departments in X County that deal with telecommunications network fraud. The units are more accustomed to independently undertaking their affairs. In the process of forming a joint effort, they are unable to achieve systematic unity for various reasons, such as distribution of interests, monopolization of resources, and competition for technical cores, etc. They are more based on interests and only undertake various tasks within the division of labor and pursue the maximization of departmental interests.

The manner of telecommunication network fraud has evolved with the changing times, and the scale of the syndicate, its technical means, and its concealment and confusion have expanded and progressed with the development of science and technology. The government's concept of governing telecom network fraud has not yet formed a relatively mature model, and the means and methods of combating, preventing, and controlling telecom network fraud lag significantly relative to the changes in such

behavior. For the ever-changing forms of telecom network fraud entrapment, the full chain of crimes, cross-regional cross-border operations, and in information exclusion, fund interception, money destination, and propaganda prevention, all show a certain sense of powerlessness; the lawless elements in the front of the run, government departments in the back of the chase, and even the emergence of governance while flooding, this side of the governance of the situation over the flooding, the timeliness of the governance of illegal criminal behavior has also been discounted, invariably expanding the damaged side of the public [20]. In recent years, in the governance of illegal criminal activities arising from network telecommunications, the relevant government departments have invested a lot of resources and exerted much effort in terms of facilities and financial resources; in the establishment and operation of systems, mechanisms, and institutions, the industry has put a lot of effort, made a lot of attempts and practice. Still, from the actual governance effect, the spread of telecommunications network fraud has not been fundamentally transformed every day. There are still a large number of such policies and cases occurring, and the property security of the public is violated and damaged all the time. This phenomenon is widespread throughout the country. In terms of cost-benefit, the benefits and effects achieved by governance are far less than the costs and efforts invested, and a great deal of effort has been spent, but a good result has not been achieved, not to mention enhancing the public's sense of security and satisfaction. As shown in Figure 5, the fraud gang used a virus to implant electronic devices to obtain the suspect's personal privacy information, such as identity information, mobile phone number, payment password, bank card account password, etc.

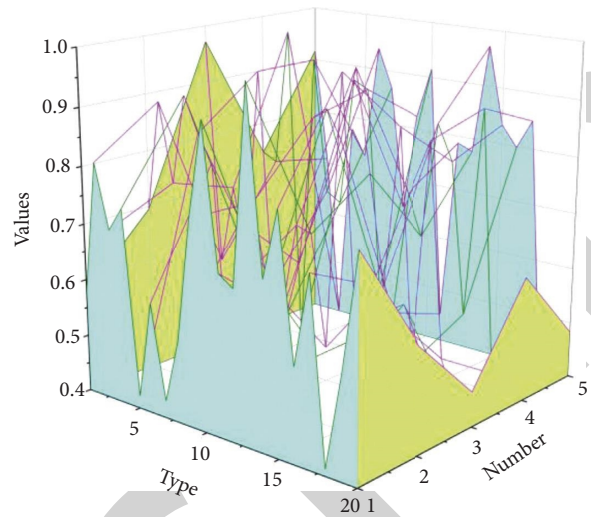


Figure 5: Distribution of telecommunications fraud viruses in household electronic products.

4.2. "Pig-Killing Tray" Telecom Network Fraud Detection Simulation Test. In the judicial practice of pig killing plate network fraud crime, because this kind of fraud crime involves very complicated personnel, evidence, resulting in the public security organs investigation process, is more difficult because it is a network fraud, so most of the evidence involved is electronic evidence; in the process of crime, because the criminal facts occurred or the place of the criminal act is not accurate, it will lead to evidence that can be easily hidden or destroyed, thus increasing the resistance of the judiciary to evidence and factual determination. In the investigation of network fraud crime, the integrity and validity of evidence must be protected; to strictly punish the crime, accurate determination of the facts of the case must be ensured to achieve fairness and justice and to ensure the probative power of evidence. Access to electronic evidence to identify the network fraud crime is a very important link. Complete and effective electronic evidence for the conviction and sentencing of the crime is vital to accurately identify the behavior of criminals, fraudulent means and sentencing circumstances, accurate and standardized application of the law. In the investigation process, to ensure the effectiveness and integrity of electronic evidence, the investigating authorities in the collection and preservation of electronic evidence must be carried out in strict accordance with legal procedures and standardized means of investigation. This paper takes X city as an example to start a quarterly simulation test. A quarter of antitelecom fraud data statistics are as follows (Figure 6).

In the investigation process of the case, to do step-by-step division of labor, not a cluster, in the collection, preservation of electronic evidence involved in the case should be assigned to the public security organs special technical personnel to implement following the relevant legal provisions of lawful forensics, both to protect the integrity of the evidence and to improve the efficiency of the evidence acquisition process. Secondly, network fraud technology is becoming more and more novel, and the means are becoming more and more difficult to figure out; just like the pig-killing dish network fraud crime, criminals in the name of dating network fraud, the use of Internet technology to manipulate the software backstage, and criminal gangs work together, with a clear division of labor; in the process of combating crime to the forensic process of legality, effectiveness, the public security organs must improve the forensic, investigative capabilities. In addition, other network fraud crimes also exist using pseudobase stations, Trojan horse links, and other ways to obtain the victim's information, to use the victim's information to operate to obtain a large amount of property. With the derivation of intelligent crime tools, the suspect can use any number-changing software to fictitious facts to carry out fraudulent behavior. It is also due to the continuous updating of high-tech means of crime, resulting in more and more people being unsuspecting. It is also due to the continuous updating of high-tech modus operandi that more and more people are falling for the scam without any precaution, and the cases are frequent. In response to this situation, the government's means must also keep pace with the times, with the help of big data analysis model (as shown in Figure 7), after a quarter of the public security organs have significantly improved technical means, "kill plate" telecom network fraud has been significantly curbed.

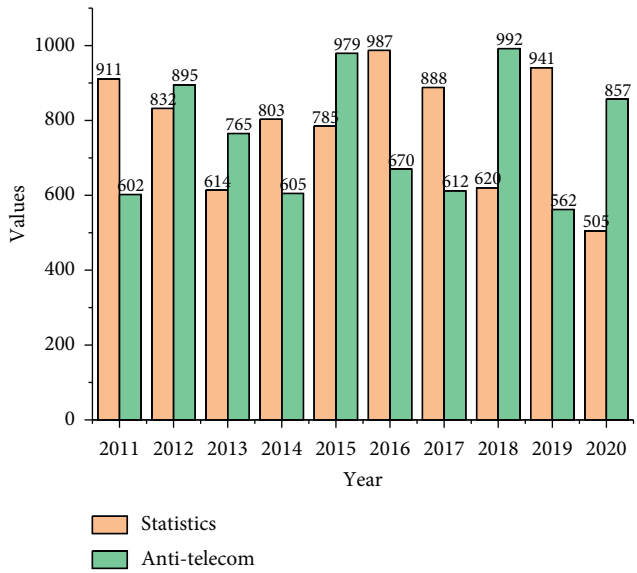


Figure 6: Q1 antitelco fraud big data report.

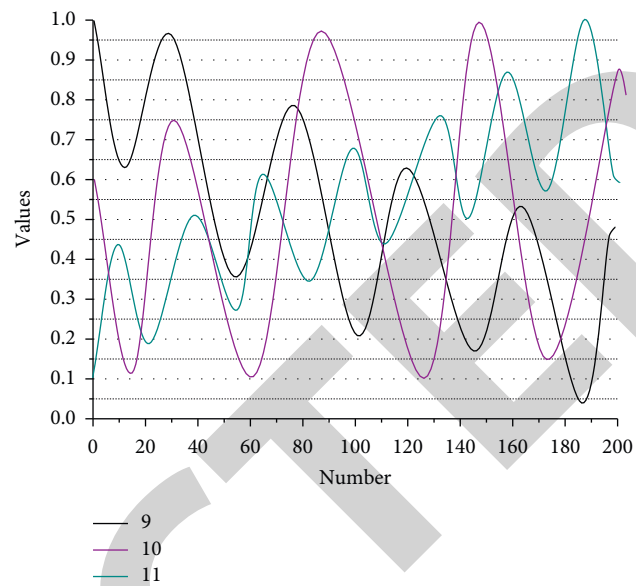


Figure 7: The decline of "piggy bank" telecom network fraud data in the first quarter of X city.

Strengthening regulation of the industries involved in telecommunications network fraud cases is an important means of improving governance capacity. Banks and telecommunications departments must implement the "real-name system" in strict accordance with the regulations and eliminate the existence of false, impostor, and shell bank cards and real-name card users. They should strengthen the cleanup of "black cards" and take effective measures to effectively solve the problem of "real names but not real people" on mobile phone cards, especially for new businesses such as Internet cards and Internet of Things cards, and carry out comprehensive risk assessments to prevent them from becoming tools for criminals to commit crimes. Banks and other financial departments should standardize the process of opening bank accounts, strengthen the audit and supervision of personal accounts and public accounts, judge the risk of fraud for unreasonable behavior that exceeds the set target, and observe the flow of funds and bank card account transactions to provide a basis for decision-making in the next step. Relevant functional departments should change the work system process, further limit the number of personal bank cards and telephone cards, clarify the legal responsibilities of individuals and units, and incorporate the real-name system of the telecommunications and financial sectors into personal social credit for the management, to solve the problem of unrealistic real names.

Telecommunications network fraud involves multiple sectors and aspects of society, not a matter for public security alone, the length of its illegal and criminal chain, the complexity and mobility of the people involved, must rely on the strength of all parties, and pool their efforts to govern. As the author, we should also actively contribute our part in the academic library, based on previous data, to give active literature research for society and public security organs, banks, and other relevant departments to borrow to study. From the statistics in Figure 8, the literature on "telecommunication network fraud" was published: the total number of articles was small in 2014, reached the peak in 2017, slightly decreased in 2018, warmed up in 2019, and always maintained a high degree of heat in recent years.

One of the important aspects of the governance of telecommunications network fraud is the game and confrontation between the scientific and technological capabilities of government departments and the renovation methods of lawless elements; only by continuously strengthening their technical countermeasure capabilities can they more efficiently research and detect lawbreakers and timely stop the implementation of interrupted telecommunications network frauds. What changes have taken place in the communication methods between people?

With the revolution of information technology swept up, the digital economy is growing at a geometric rate, and science and technology have penetrated all human production, life, and social communication methods. We must vigorously implement the protective technical measures of the Internet, telecommunications, financial, and other network enterprises to trace the perpetrators of telecom network fraud in the first instance and provide timely and effective assistance to public security departments in the process of apprehending suspects. Banks should further enhance the standard requirements for smart machine identification to avoid the impersonation of head cards through face recognition. Public security authorities, as the main department, to crack down telecom network fraud, have to strengthen team building to improve the business quality capabilities of public security authorities, which is the core of improving the effectiveness of combating telecom network fraud. Telecommunication network fraud seizes the psychology that the masses are greedy for small bargains, weak-minded and afraid of trouble, etc. Once the masses improve their self-defense ability, there will be no opportunity for fraudsters to take advantage of it. Therefore, it is

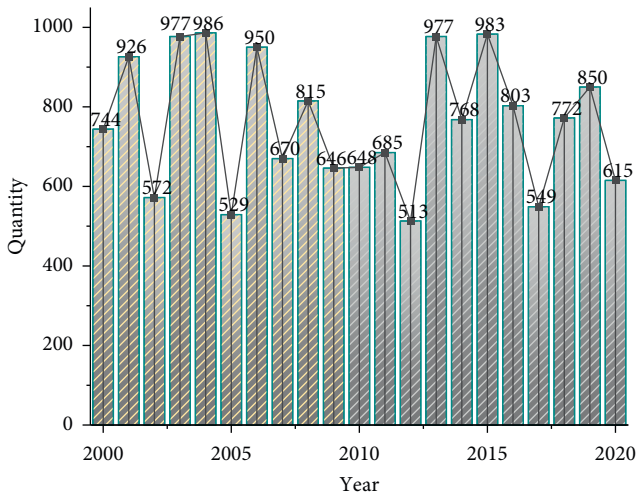


Figure 8: Annual trends in the publication of literature on "telecommunications network fraud".

necessary to constantly carry out publicity work on telecommunication network fraud to citizens so that the public's awareness of safety and prevention can be enhanced.

5. Conclusion

With the development of the Internet, telecommunications network fraud is frequent, and "piggybacking" as a new type of telecommunications network fraud crime, on the one hand, brings trauma to people's emotions, and on the other hand, also makes people's property suffer huge losses. The new era of network crime means constantly iterative updates; timely curbing the trend of network crime is imperative; public security organs should update the concept and implement diversified, multisectoral coordination of social security governance. At present, the effectiveness of the public security organs in combating "piggy-backing" network fraud is not obvious, which urgently requires an updated concept. In the era of big data, public security organs should address the characteristics of "piggy-back" telecom network fraud, expand intelligence sources, integrate and analyze intelligence information, strengthen the interconnection and mutual sharing of information resources, improve the investigative cooperation mechanism, establish a data-oriented prevention and early warning mechanism, and at the same time, pay attention to strengthening the protection of citizens' personal information when retrieving clues. At the same time, attention is paid to strengthening the protection of citizens' personal information when investigating clues, establishing a long-term mechanism for investigation, achieving "fast, accurate and fierce" combat against telecommunication network fraud crimes, protecting citizens' property, and maintaining social stability.

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 paper was supported by the Project Research and Practice on Teaching Reform of Public Security Based on Mobile Network Security Information Platform, supported by special project of 2019 Guangxi higher education undergraduate teaching reform foundation. Project Number: 2019JGB402.

References

[1] V. Jain, "Perspective analysis of telecommunication fraud detection using data stream analytics and neural network classification based data mining," *International Journal of Information Technology*, vol. 9, no. 3, pp. 303–310, 2017.

[2] X. Liu and X. G. Wang, "Probabilistic graphical model based approach for bank telecommunication fraud detection," *Computer Science*, vol. 45, no. 7, pp. 122–128, 2018.

[3] A. H. Ahmad, R. Masri, and C. M. Zeh, "The impact of digitalization on occupational fraud opportunity in telecommunication industry: a strategic review," *PalArch's Journal of Archaeology of Egypt/Egyptology*, vol. 17, no. 9, pp. 1308–1326, 2020.

[4] H. H. Kilinc, "A case study on fraudulent user behaviors in the telecommunication network," *Electrica*, vol. 21, no. 1, pp. 74–84, 2021.

[5] H. Lin, G. Liu, and J. Wu, "Fraud detection in dynamic interaction network," *IEEE Transactions on Knowledge and Data Engineering*, vol. 32, no. 10, pp. 1936–1950, 2019.

[6] G. Praveen, V. Chamola, V. Hassija, and N. Kumar, "Blockchain for 5G: a prelude to future telecommunication," *IEEE Network*, vol. 34, no. 6, pp. 106–113, 2020.

[7] N. Zhang and Y. Xu, "Environmental study on cooperation system of cross-border tracking economic crimes based on block chain—take telecommunication fraud as an example," *Ekoloji*, vol. 28, no. 107, pp. 4437–4446, 2019.

[8] M. Qiu and Y. Yang, "Analysis of the current situation and characteristics of college student "online fraud cases"," *International Journal of Mobile Computing and Multimedia Communications*, vol. 12, no. 2, pp. 56–73, 2021.

[9] A. Battal and R. Samli, "An action management system design and case study on its usage for cyber fraud prevention and risk analysis," *Journal of Innovative Science and Engineering*, vol. 5, no. 2, pp. 143–161, 2021.

[10] N. Duha, "Short message services (SMS) fraud against mobile telephone provider consumer review from law number 8 of 1999 concerning consumer protection," *Journal of Law Science*, vol. 3, no. 1, pp. 36–43, 2021.

[11] N. Kala, "A study on internet bypass fraud: national security threat," *Forensic Research and Criminology International Journal*, vol. 7, no. 1, pp. 31–35, 2019.

[12] R. A. Leite, T. Gschwandtner, S. Miksch, E. Gstrein, and J. Kuntner, "Visual analytics for event detection: Focusing on fraud," *Visual Informatics*, vol. 2, no. 4, pp. 198–212, 2018.

[13] O. S. Yee, S. Sagadevan, and N. H. A. H. Malim, "Credit card fraud detection using machine learning as data mining technique," *Journal of Telecommunication, Electronic and*

Hindawi
Journal of Robotics
Volume 2022, Article ID 6178917, 12 pages
https://doi.org/10.1155/2022/6178917

Research Article

A Brain-Computer Interface for Teleoperation of a Semiautonomous Mobile Robotic Assistive System Using SLAM

Vidya Nandikolla, Bryan Ghoslin, Kevin Matsuno, and Daniel A. Medina Portilla

Department of Mechanical Engineering, College of Engineering and Computer Science, California State University, Northridge, Los Angeles, CA, USA

Correspondence should be addressed to Vidya Nandikolla; vidya.nandikolla@csun.edu

Received 18 November 2021; Revised 7 February 2022; Accepted 9 February 2022; Published 26 February 2022

Academic Editor: Kaijian Xia

The proposed assistive hybrid brain-computer interface (BCI) semiautonomous mobile robotic arm demonstrates a design that is (1) adaptable by observing environmental changes with sensors and deploying alternate solutions and (2) versatile by receiving commands from the user's brainwave signals through a noninvasive electroencephalogram cap. Composed of three integrated subsystems, a hybrid BCI controller, an omnidirectional mobile base, and a robotic arm, the proposed robot has commands mapped to the user's brainwaves related to a set of specific physical or mental tasks. The implementation of sensors and the camera systems enable both the mobile base and the arm to be semiautonomous. The mobile base's SLAM algorithm has obstacle avoidance capability and path planning to assist the robot maneuver safely. The robot arm calculates and deploys the necessary joint movement to pick up or drop off a desired object selected by the user via a brainwave controlled cursor on a camera feed. Validation, testing, and implementation of the subsystems were conducted using Gazebo. Communication between the BCI controller and the subsystems is tested independently. A loop of prerecorded brainwave data related to each specific task is used to ensure that the mobile base command is executed; the same prerecorded file is used to move the robot arm cursor and initiate a pick-up or drop-off action. A final system test is conducted where the BCI controller input moves the cursor and selects a goal point. Successful virtual demonstrations of the assistive robotic arm show the feasibility of restoring movement capability and autonomy for a disabled user.

1. Introduction

In the fields of robotics and biomedical engineering, there is a growing trend in designing assistive robots to aid users in demanding tasks. The design parameters for these robots involve a detailed understanding of the operating environment, intended task, and user-control inputs. Sensors, cameras, and other instrumentation feedbacks have propelled engineers to create "smart" controllers that allow assistive robots to detect changes in the working environment, plan a work around, and execute sensitivity and accuracy. In addition to improving the assistive robot's robustness and resilience, researchers are developing versatile controller inputs that allow users, regardless of their severe physical or neurological limitations, to interface with the assistive robot. In most cases, the severity of the user's condition increases the controller's complexity and the robot's level of autonomy.

This paper focuses on the implementation of an assistive hybrid brain-computer interface (BCI) semiautonomous mobile robotic arm for users with debilitating paralysis or progressive nervous system diseases. While these users have impaired brain lobes that prohibit motor movement and/or speech, their cognitive and sensory functions are often uncompromised. Brain signals from these functional lobes are obtained from a noninvasive electroencephalogram (EEG) cap. As a result, a set of specific physical or mental tasks (i.e., jaw clench and imagined left/right hand squeeze) are mapped to robot commands with some user training. Receiving these commands, a mobile robotic arm is a

solution for everyday object manipulation and transportation in household and occupation environments. To keep the number of inputs low and prevent users from getting fatigued, the proposed robot is semiautonomous and goal position oriented. The omnidirectional mobile robot base has obstacle avoidance capability and can plan paths through known free spaces. Likewise, the robot arm uses a vision system where the user specifies a goal point and the controller calculates the necessary joint movement to grab or release the recognized object.

The three integrated subsystems, BCI controller, input, mobile base, and arm, are developed separately to ensure that the individual criteria are met before the system level testing. For the BCI controller input, the active electrodes for the specified tasks are identified, a proper machine learning algorithm is selected, and the final controller configuration is created. For the robot base, rapid prototyping is used to generate the SLAM algorithm that localizes the robot within its environment while building a map of obstacles and free space from sensor data. For the robot arm, the incoming camera stream cloud data is calibrated to link kinematic algorithms and physical markers. To verify the subsystems, the software communicates and operates cohesively; virtual testing is performed in a simulation environment called Gazebo. The final version of the BCI controller input is paired with the robot base and arm. A simulated brainwave stream is used to run through every command and demonstrate the successful action. The following sections of the paper summarize the development of the subsystems: modeling of the proposed robot and the simulated environment, simulation and testing verification, results, and conclusions.

2. Motivation and Literature Review for Semiautonomous Mobile Robotic Assistive Systems

2.1. Motivation of Semiautonomous Mobile Robotic Assistive System. In 2011, the World Health Organization (WHO) reported that over a billion people (or about 15% of the world's population) were living with a disability [1]. This significant number of people worldwide shows that access to education and work opportunities, even performing activities of daily living (ADLs) at home, is unreachable by themselves and require caregiving personnel [2]. The additional impact of COVID-19 and epidemiological measures (i.e., stay-at-home orders) has added to the burden of caregivers and care recipients; one journal points that about one out of five care recipients had difficulty in obtaining care from outside the household and states the need for intervention [3]. While health policymakers and social organizations can certainly improve aid to the disabled, the deployment of semiautonomous mobile robotic assistive systems can empower one of the more severe handicap conditions, a quadriplegic, to perform daily tasks at home that normally require personnel in stay-at-home scenarios. In the journal, improving the autonomy for people suffering from paralysis in all four limbs is the primary focus.

2.2. Literature Review for Semiautonomous Mobile Robotic Assistive Systems. Semiautonomous mobile robotic assistive systems are in development in the forms of smart wheelchairs and robotic arms. For quadriplegics, or people suffering from amyotrophic lateral sclerosis (ALS), controlling the advance devices requires input from the user generating specific signals from the brain. Some researchers utilized the user's brain signal that mimicked the frequency of a specific flashing light on a screen [4, 5]. However, this approach is not optimal for operating a mobile robotic device due to the following: (1) driving a robot or operating a robotic hand requires the user's undivided attention of the surroundings and (2) waiting on the brain to mimic one of the flashing frequencies may take longer time to implement and become ineffective if the environment suddenly changes. As a result, the control chosen for this study is an asynchronous approach, where the user has to perform mental tasks and clench his/her jaw. Studies that favor this approach show that the action implemented by the user is quicker and accurate as if the training is conducted in real time [6–8]. In terms of hardware, combined smart wheelchairs and robotic arms are developed by universities and companies [9, 10]. While these have greatly increased the mobility and functional ability of the disabled, the end user is still required to operate either the wheelchair or the arm by either the joystick or the eye tracker. The assistive robotic system presented in this paper is designed to operate solely on the user's brain signals. In addition, since this proposed mobile robotic arm is separate from a wheelchair, the end user can retrieve objects without having the hassle of leaving his/her current position.

3. BCI Input Signal

3.1. 6-Class BCI Mobile Arm Controller: Schematic. Similar to a truck crane, the user operates the 6-class BCI mobile arm's base and arm separately (Figure 1). First, the robot base is moved to a desired location; then, the user switches to operating the arm's focus cursor to select an object that will be picked up. Likewise, to drop off the held object, the user navigates using the robot base to the new location and then switches to the robot arm once again to select the location to place the item. Since these two operations are isolated, the same mental tasks can be assigned multiple commands.

In this case, the asynchronous signals of the imagined right foot, left foot, right hand, and left hand correspond to the robot base moving forwards, backwards, right, and left; the robot arm cursor moves up, down, right, and left on a camera screen with the same mental tasks. An electromyogram signal of the user's jaw clench is used for transitioning between the robot base and arm and initiating a pick-up or drop-off action. When switching from the robot arm to the base, the jaw clench first confirms the desired object selected by the cursor. A built-in software calculates the distance between the robot and the object and moves the proper arm motors to pick up. Likewise, when dropping off the object, the user first guides the cursor and clenches his/her jaw to confirm the location point; the smart robot arm

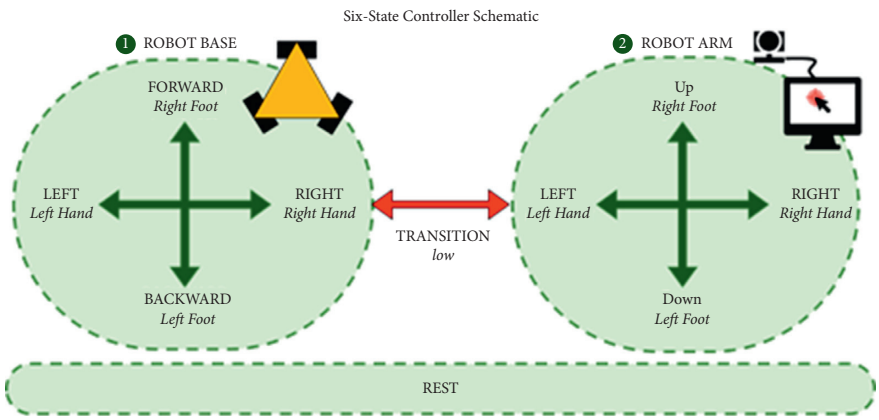


Figure 1: A 6-class BCI mobile arm controller with commands.

software then performs the action, freeing the user from the complexity of controlling every mechanical joint like the fingers and wrist. Additional sensors on the mobile base help detect obstacles or walls and avoid collisions while the user is giving movement commands.

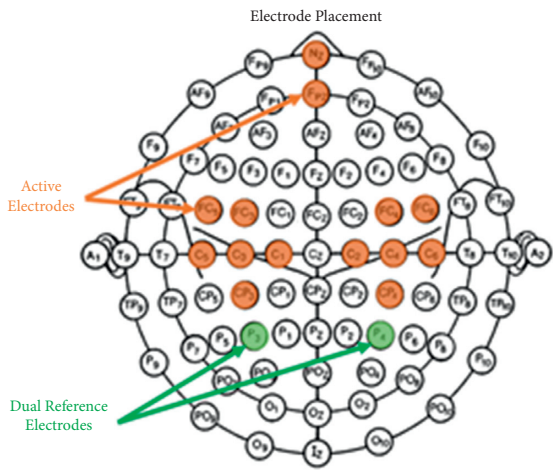
3.2. 6-Class BCI Mobile Arm Controller: Development. Development of the 6-class BCI controller involved identifying responsive electrodes on the scalp and selecting the appropriate machine learning algorithm for the targeted mental tasks. Three male subjects participated in a series of experiments where their brainwaves were recorded in response to physical and imaginary tasks (i.e., hand squeezes, foot taps, and jaw clenches). These recordings were preprocessed and analyzed using event-related potential plots and topographical maps [11, 12]. The results for each subject showed that the twelve electrode locations with respect to the 10–20 international system, FC3, FC4, FC5, FC6, C1, C2, C3, C4, C5, C6, CP4, and CP4, were capable of recognizing the unique brain signal characteristics of each task; as a result, this electrode layout is finalized and incorporated into the BCI controller headset (Figure 2(a)). To overcome the nonstationary nature and incoming clusters of other brainwaves, the next step is to downselect the appropriate machine learning algorithm between linear discriminant analysis (LDA) and relevance vector machines (RVM) (Figure 2(b)). Using the same preprocessed recordings, the error rate percentages for each algorithm were calculated and compared for two-, three-, five-, and six-class controllers. The results showed that RVM had a higher accuracy than LDA on multiclass controllers among all three subjects. As a result, RVM is selected for the final 6-class BCI controller configuration.

3.3. 6-Class BCI Mobile Arm Controller: Virtual Testing Architecture. With the 6-class BCI controller fully functional, the software architecture needed to support virtual testing is developed. Referencing to Figure 3, the block diagram starts with a compiled set of brainwave recordings that are looped to simulate a twelve-electrode channel EEG input stream from a user. The combined brainwave file of one male subject contains fifteen consecutive examples of each mental task and is placed in nearly sequential order: right foot (RF), left foot (LF), right hand (RH), left hand (LH), and jaw. By systematically running through the sequence of tasks, any discrepancy or deviation from this order is ruled as an error.

The "rest" mental task did not follow the order since it was embedded in the five other examples. Having more examples of the "rest" task is strategically done to bias the controller to keep the robot idle if it is unable to classify the user's brainwaves; this increases the controller's overall safety and reliability. With a complete runtime of 6 minutes and 25 seconds, the simulated brainwave stream is fed to the lab streaming layer (LSL). LSL is an open source multisignal acquisition system that collects unified measurement time series data and allows compatible programs to retrieve or add to the data [13]. This system's built-in code has libraries and wrappers that allow the transaction of data to occur across platforms with different languages (i.e., C/C++, MATLAB, Java, and Python). BCILAB, an open source MATLAB plug-in that contains the 6-class RVM BCI controller, is linked to LSL and retrieves the incoming stream of brainwaves. The controller actively compares the sequenced brainwaves to its trained examples and outputs a stream of integers that are uniquely associated with each mental task (Table 1); note that this integer represents the controller's highest calculated probability out of the six mental tasks. BCILAB then adds the controller's command stream of integers back to LSL for the virtual robot code to retrieve. Since both the robot mobile base and the arm subsystem's control center use Python, LSL converts the stream of integers into a compatible data type in Python. These control centers receive additional feedback from its semiautonomous simulated sensors before performing the intended action. In summary, the virtual testing block diagram starts with a raw stream of brain signals and ends with a robotic action dictated by a command stream.

4. Robot System

The robot system consists of 3 degrees of freedom (DoF) omnidirectional mobile base, to which a 3 DoF robotic arm



(a)

Machine Learning Comparison

Error Rate Percentage		subject 501		subject 502		subject 503	
		LDA	RVM	LDA	RVM	LDA	RVM
Hands	2-Class	26.7	26.7	30	43.3	10	10
Feet	2-Class	30	35	26.7	16.7	23.3	30
jaw/roof	2-Class	-	-	-	-	3.3	33
Hands	3-Class	41.7	40	28.3	20	16.7	20
Feet	3-Class	43.3	40	38.3	31.7	23.3	25
jaw/roof	3-Class	-	-	-	-	11.7	10
	5-Class	52.5	45	35.8	30.8	33.3	29.2
	6-Class	-	-	-	-	34.5	34.5

(b)

Figure 2: (a) Finalized electrode layout [1] and (b) LDA vs. RVM performance comparison.

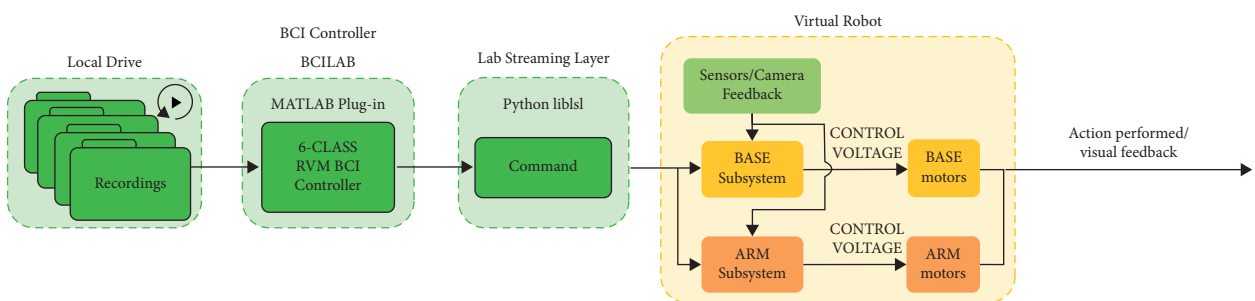


Figure 3: A detailed block diagram showing the virtual testing of the 6-class BCI controller.

Table 1: 6-class BCI controller reference showing the mental task input and associated robot action.

Input brainwave mental tasks	BCILAB output integers	Robotic base actions	Robotic arm cursor actions
RF	1	FORWARD	DOWN
LF	2	BACKWARD	U
RH	3	RIGHT	RIGHT
LH	4	LEFT	LEFT
JAW	5	N/A	GOAL POINT
REST	6	IDLE	IDLE

with a 4 DoF end-effector is attached. The total system together consists of 10 DoF which is controlled together. The principle idea of an omnidirectional wheel is that the central wheel is surrounded by rollers which are placed at an angle of either 45° or 90° around the periphery of the wheel. The rollers introduce a linear slip perpendicular to the face of the wheel, which does not exist in a fixed standard wheel [14]. Combined with the rotational motion of the wheel and the rotational twist at the contact point of the ground, omnidirectional wheels can achieve 3 DoF. As such, a mobile system which utilizes omnidirectional wheels is also capable of 3 DoF and can travel through an environment under any orientation. To achieve this, each omnidirectional wheel requires its own motor and controller. Referencing to Figure 4, the overall motion of the mobile robot is dictated by the rate of rotation of each wheel [15].

The wheels of a mobile robot must rotate at appropriate speeds to achieve a desired chassis speed, which includes both the linear and the rotational velocities. To determine the required speeds of each wheel, the kinematics of the mobile robot are studied. The kinematic calculations were derived from the works of [16] and were modified to the specific notations as shown in Figure 5. The work was also simplified by using the body twist V_b, which does not rely on the robot's orientation relatively to the work environment but rather the angular velocity of the chassis:

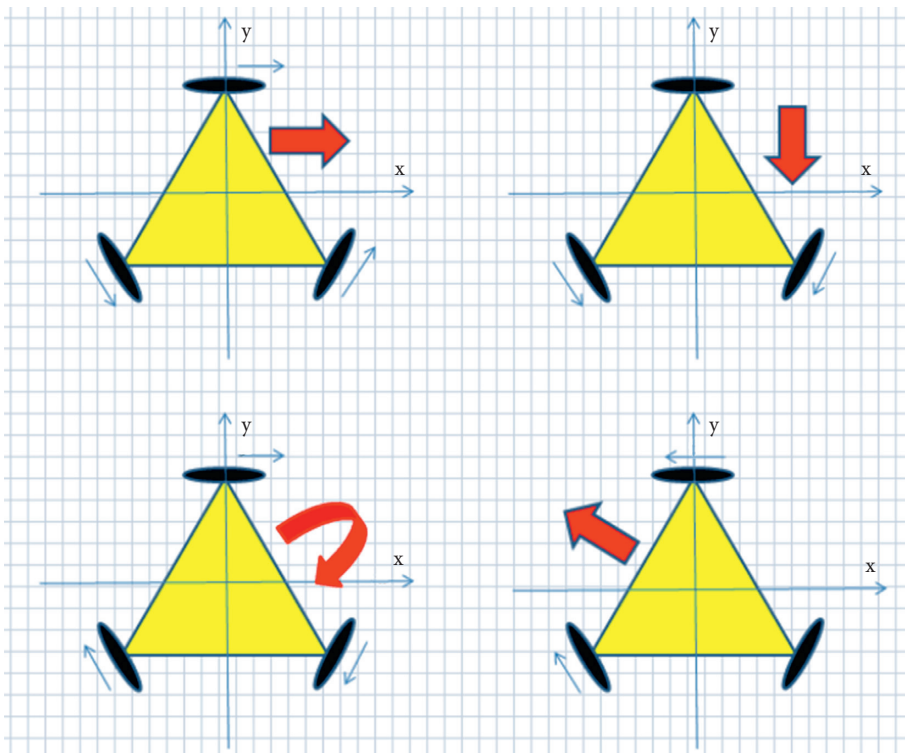


Figure 4: Robot body motion and required wheel rotation.

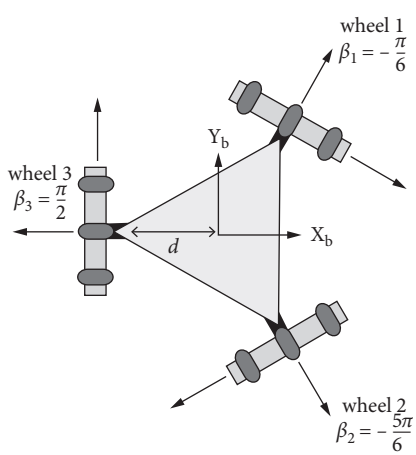


Figure 5: Kinematic model of the robot [6].

$$V_b = \begin{bmatrix} \omega_b \\ v_x \\ v_y \end{bmatrix}. \tag{1}$$

$$u = \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} = H(0)V_b = \frac{1}{r}\begin{bmatrix} -d & \cos\left(-\frac{\pi}{6}\right) & \sin\left(-\frac{\pi}{6}\right) \\ -d & \cos\left(-\frac{5\pi}{6}\right) & \sin\left(-\frac{5\pi}{6}\right) \\ -d & 0 & 1 \end{bmatrix}\begin{bmatrix} \omega_b \\ v_x \\ v_y \end{bmatrix}. \tag{2}$$

As referenced in Figure 5, d is the distance from the center of the robot to the contact point of each wheel and β_i is the driving angle of each omnidirectional wheel relative to the X_b-axis of the body reference frame. Thefollowing equation is the solution of the kinematic model and expresses the wheel rotational velocities with respect to the body twist:

Letting $d = 0.1554$ m and the wheel radius $r = 0.05$ m, the wheel velocity can be expressed as

$$\begin{aligned} u_1 &= \frac{1}{0.05}\left(-0.1554 * \omega_b + 0.866 * v_x - 0.5 * v_y\right), \\ u_2 &= \frac{1}{0.05}\left(-0.1554 * \omega_b - 0.866 * v_x - 0.5 * v_y\right), \\ u_3 &= \frac{1}{0.05}\left(-0.1554 * \omega_b + v_y\right). \end{aligned} \tag{3}$$

The 3 Dof robot arm consists of an elbow, shoulder, and wrist joint. The arm is rigidly attached and has a planar workspace parallel to wheel 2 as shown in Figure 6. The end-effector used is the Brunel Hand 2.0 and is attached to the wrist joint of the arm. The kinematics of the manipulator is solved using the mobile base as the fixed reference. For the end-effector to reach a goal target, the inverse kinematics first solves for a goal location of the mobile base that places the goal target in the arms workspace. This goal location is solved based on a desired arm pose when grabbing an object.

The goal target is the input from a user interface that uses BCI commands to select a point from a camera feed. The camera feed provides the point cloud data used to measure

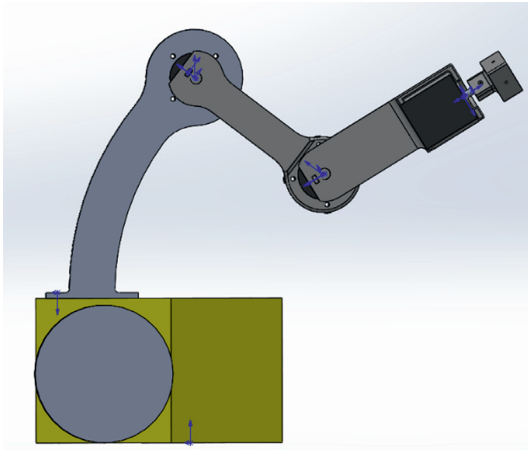


Figure 6: Robot arm 3D model.

distance from the robot. Using end-effector control allows the robot to automatically solve for the movement of the entire system [17]. This semiautonomous mode uses the camera system to facilitate the user input while providing the more complex robot movements required for object manipulation [18]. Once the robot arm achieves the desired pose, the Brunel Hand initiates a grabbing command which has been physically tested on objects [19].

5. Virtual Software Interface

Due to the restrictions caused by COVID-19, a physical build of the robot was not possible, and the project was moved to a virtual environment. Robot Operating System (ROS) and Gazebo are used to visualize and simulate the entire assistive robot system. The user interface is created using RViz, a visualization tool available in ROS, and is shown in Figure 7.

The user interface consists of a visualization display on the left and a camera feed for user input on the right. During operation, a cursor is bounded within the right camera feed. The cursor moves and can select points using BCI commands. The camera cannot move since it is not connected to the mobile robot; therefore, visualization markers are created to show the intended motion of the mobile base and appear in the visualization display. The motion is also simulated in the Gazebo environment, which offers the use of joint torque limits and collision physics. The simulation environment uses the robot's collision and inertial models alongside joint torque and velocity limits to simulate the results of input commands. PID controllers are used for each joint and the outputs of these controllers are sent to the Gazebo environment.

Figure 8 shows a portion of the block diagram for the robot system. After including the torque and velocity limits of the selected motors to the robot definition, the PID controller of each joint is tuned. The arm uses position controllers and is tuned to reach different grabbing poses with no steady state error. The controllers for the omnidirectional wheels are designed with the assistance of the Joint Effort Interface. This ROS controller converts the desired joint speed into a torque or force required for the joint to achieve the goal velocity. The resulting speed is recorded by the sensors and compared to the desired speed to determine the error of the system.

6. SLAM

The ability of accurately estimating the position of a mobile robot as it travels through an environment is a computationally challenging problem, as it is difficult to calculate the instantaneous speed of a robot. Without an accurate measure of the speed, it is hard to estimate how far it has traveled after a period. To produce an estimate, a robot must perform odometry by identifying obstacles within its immediate vicinity. Tracking obstacles and calculating the change of the robot's distance between them can help estimate the speed at which the robot is traveling. The collected odometry data can be used to achieve any high level of autonomous control. The inclusion of a map of the environment can assist in localizing the robot around known obstacles and completing path planning calculations. A robot which can also produce or update a map of the environment while localizing itself to the map can achieve even higher levels of autonomous control as it can account for known obstacles. This computational capability is what is known as Simultaneous Localization and Mapping (SLAM). There are many different SLAM algorithms which can achieve the same purposes [19–22]; most just rely on differing sensors employed in the system. The SLAM algorithm employed in this project is developed with the use of ROS nodes which rely on the sensor data of the entire system.

The localization node runs a function called Adaptive Monte Carlo Localization (AMCL) [23]. AMCL parses the data from laser scans and cameras to determine the distance from the robot to obstacles. The obstacles are compared to those identified on the map to create an approximation of the robot's current location. The AMCL algorithm compares to the robot odometry which is calculated by the data recorded by encoders on the motor which estimates how fast the robot is moving and therefore how far it traveled. This cross-reference is used to verify the measurements from the sensor data and improve the accuracy in the localization approximation.

The mapping node, known as Move Base [24], is responsible for updating costmaps by marking obstacles and free space as well as running the trajectory planning algorithm. This project is developed with the use of a prebuilt map of the environment, which is provided to the robot upon deployment. Move Base is responsible for creating two separate costmaps which are built on top of the existing map. In this setup, the boundaries of the environment remain consistent, but the changes of obstacles within the environment are updated as they change. One of the costmaps is a global costmap, which will mark and record obstacles even once they are out of range of the robot's sensors and no longer an immediate obstacle. The other costmap is a local costmap, which only keeps track of obstacles observed within a defined distance from the robot. Both costmaps are

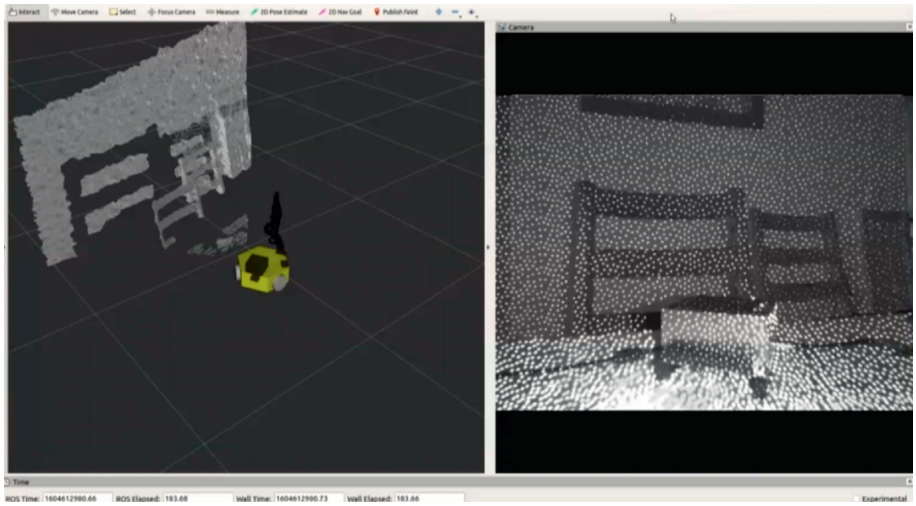


Figure 7: User interface using RViz.

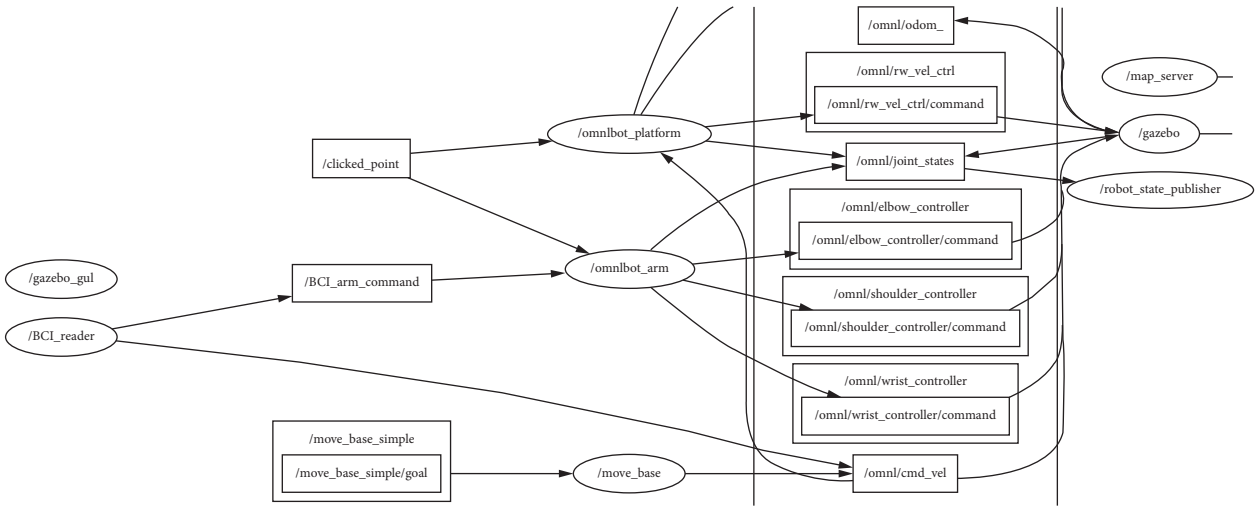


Figure 8: Block diagram of the robotic system.

updated at the same time; the only difference is how long each one records the obstacles.

Move Base handles trajectory planning with the Dynamic Window Approach (DWA) planner [25]. Given a goal destination, a global path is set within the global costmap. It follows the free space and avoids as many obstacles as possible, while still following the shortest path. The DWA planner then performs a sampling of forward simulations, by calculating the outcomes of running the robot a short distance along the global plan at varying speeds. Each forward simulation is scored by weighted criteria which determine the safest and shortest route possible for the robot. Forward simulations which result in collisions are immediately discredited. The trajectory of the best forward simulation is sent to the robot system and the DWA planner performs another round of simulations based on the change of location of the robot since the last simulation. The deployment of this SLAM algorithm permits the robot to safely navigate from one goal to the next. The routine checks accurately account for the robot and allow for precise remote BCI control. They also keep the robot from crashing into any obstacles which may have been moved and also account for any delays in the data stream.

7. Methods

7.1. Experimental Setup. To ensure that the BCI controller, the robotic base, and the arm subsystems all work cohesively and a virtual demonstration is built to run through every command, the virtual version of the robot assembly is designed with coded sensors and cameras to closely represent the real hardware. In addition, naturally occurring phenomena, like coefficients of friction between the wheels and floor, and moments of inertia for the motors driving the arm, hand, and wheels, are included. The virtual environment in Figure 9 is a single-floor space with typical objects (i.e., chairs, tables, and trash cans) and walls. This is used as a sandbox for testing the entire assistive robot system.

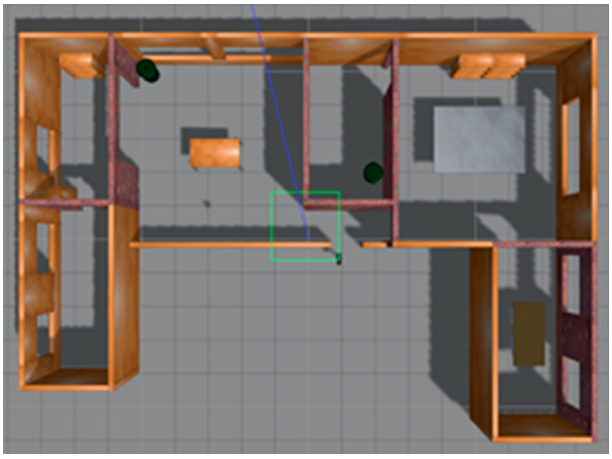


Figure 9: Environment in Gazebo.

7.2. Experimental Procedure. To generate the command input stream used for the tests, the BCILAB MATLAB plug-in is launched. The incoming EEG signal stream is configured to read the recorded brainwave data with the following parameters: input stream name "laststream," an update frequency of 20 Hz, an internal buffering length of 10, a conversion to double precision data type, and a looped playback. With the incoming stream being processed by the 6-class RVM BCI controller, the controller's output stream of integers is then configured to be written to LSL with the following parameters: command stream name "bci," output form "mode," and an update frequency of 10 Hz. A verbose output is checked to ensure that the command stream generated in MATLAB is being sent correctly to LSL. The robotic base and arm have a shared node which reads the stream as it comes in over LSL and collects data over the span of five seconds. The BCILAB "modes" form outputs what the BCI controller identifies as the most likely user input command, which can fluctuate slightly with artifacts and other oddities in the EEG signals. By recording all output during a five-second interval, the intended command output as determined by the controller is guaranteed to fall into that window; that is, how the EEG data is epoch. The ROS node reading the LSL stream then takes the mode of input signals to determine the appropriate action that the robotic base or arm should perform as referenced in Table 1.

Because both robotic subsystems operate independently, the virtual demonstration of controlling the mobile robotic arm is done in a modular approach. First, the recorded brainwaves and controller are used to move the robot base in the simulated environment based on the actions in Table 1. Next, the 6-class BCI controller goes through the string of commands in Table 1 to control a cursor on the user input camera feed and creates a goal point to initiate a pick-up or drop-off action. The entire system is tested using the goal point control. The processed BCI input signals move a cursor over a video feed and selects an object for the arm to grab with its manipulator. A grabbing pose is selected, and the corresponding inverse kinematics solves for the goal location of the mobile base to place the goal object on the arms grabbing point. The user selects an arm pose that dictates the set of inverse kinematic equations which needs to be solved. An application may require the manipulator to keep an object in its original orientation or to fully reach out to the end of the workspace. These conditions allow for a unique solution of joint angles with minimal user input. A sample of the inverse equations to solve an arm pose to keep the end manipulator horizontal is shown in Figure 10. With a selected goal point TP, the shoulder and elbow joints are solved to satisfy the final arm pose ends at TG after rotating and moving the mobile base a distance R_Y.

The shoulder angle θ_s is calculated using the first following equation, a horizontal arm pose desired θ_e is calculated using the second following equation, and the last equation is used to find the mobile base goal location R_Y. The vector space approach is used to derive the equations of different arm poses, including a fully extended pose that reaches the end of the workspace and is used in testing to verify max load conditions:

$$L_s C\theta_s = \left({}^T P_z + d_T\right), \tag{4}$$

$$\theta_s + \theta_e = 90^\circ, \tag{5}$$

$${}^T P_Y = L_s S\theta_s + \left(L_e + h_{\text{offset}}\right) + R_Y. \tag{6}$$

With the assistance of the implemented SLAM algorithm, the mobile base maneuvers until it reaches the goal location. Once the goal location is reached, the arm moves to the grabbing point and the grabbing/releasing operation begins. For these tests, the system did not use a live feed of BCI commands but rather a loop of prerecorded signals.

8. Results

The developed 6-class RVM BCI controller had an error rate percentage of 34.5%. When applying the recorded brainwave file to the controller, BCILAB recalculated the controller accuracy on the simulated data to 18.5%. The increase in accuracy is most likely due to the combination of reused brain recording segments used to train the controller. While this is not entirely representative of in-person testing, the overall objective is to demonstrate that the 6-class BCI controller can use brainwaves as a command input to control the robot. The controller designed for the three wheels of the omnidirectional mobile system only utilized the proportional and integral gains to achieve the desired results. Issues within the physics simulation prevented the controllers from removing all errors from the wheel velocity goal, but, as Table 2 shows, the steady state error was reduced as much as possible to produce acceptable results.

Figures 11–14 show the results of the experimental procedure for the entire system. Figure 11 shows the robot in the Gazebo environment waiting for a command. A point is selected through the user interface and the goal position visualization markers are created, as shown in Figure 12. The

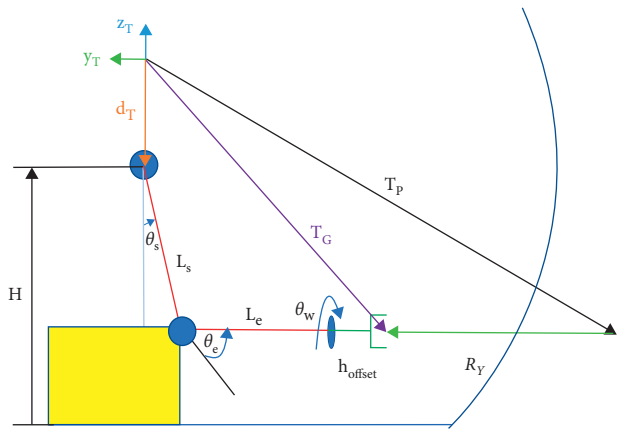


Figure 10: Workspace plane view of horizontal grabbing pose.

Table 2: PI gains and error for each omnidirectional wheel.

Controller	P	I	Error %
Wheel 1	0.060375	0.04	4.5
Wheel 2	0.06035	0.04	1.26
Wheel 3	0.052	0.04	3.93

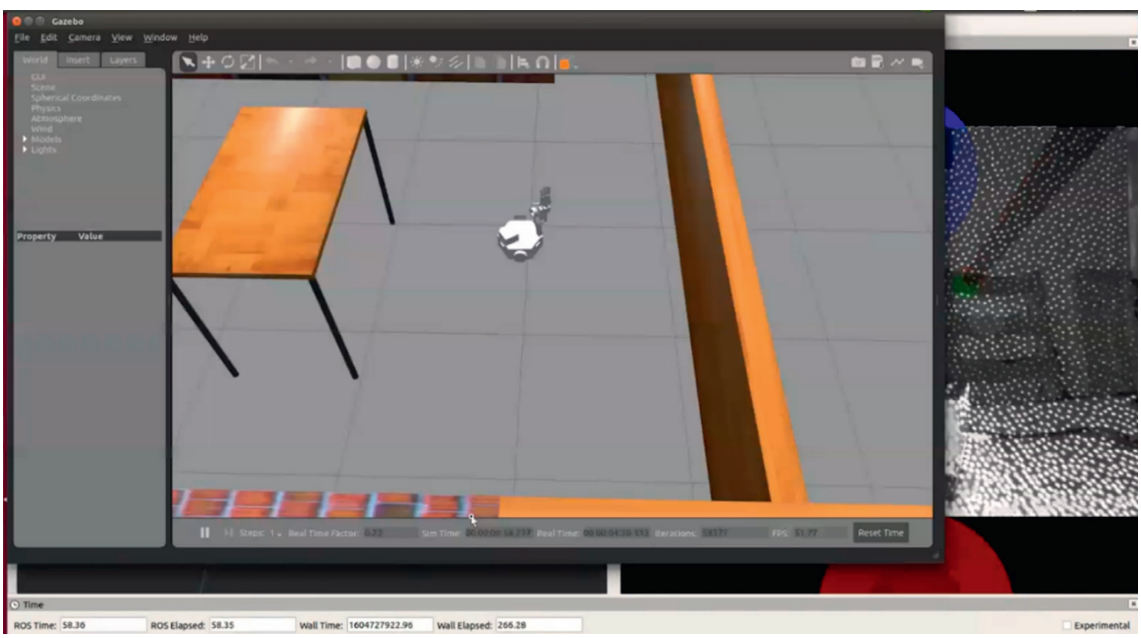


Figure 11: Starting robot position in Gazebo.

green vector is the user input and points to the goal point from the camera. The red vector points to the goal point from the manipulator frame and is used to solve for the movement of the mobile base. The blue vector is the location of the manipulator after it moves to its grabbing pose, in this case extending to the edge of the arm workspace.

The white vector is the difference between the goal point and the grabbing point, which defines the goal destination

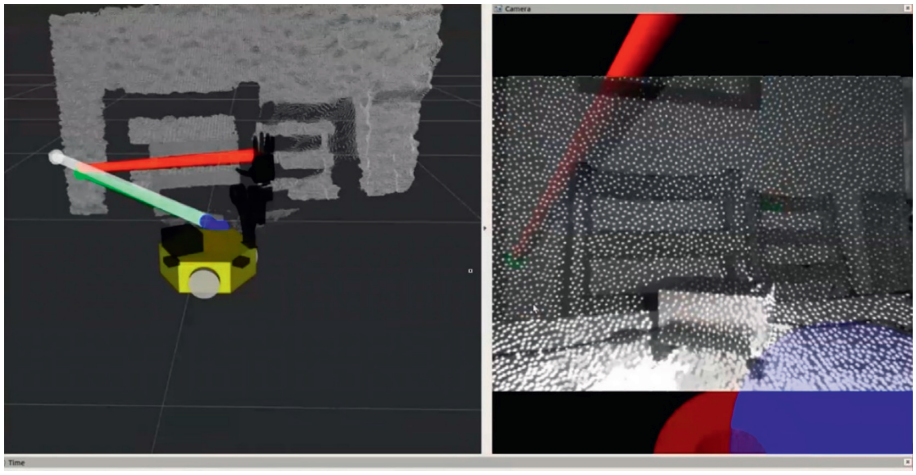


Figure 12: BCI cursor control to select point from camera feed in RViz.

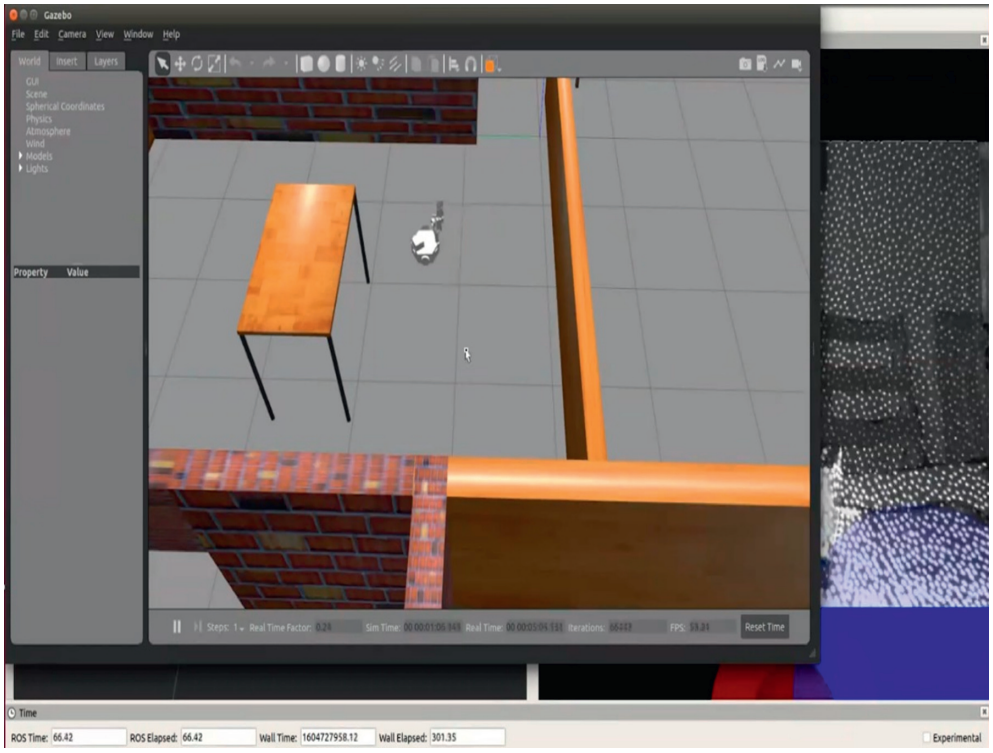


Figure 13: Robot moving in Gazebo after BCI input is processed.

for the mobile base. With the solved mobile base goal location, a path is planned for the robot to move in the Gazebo environment, as shown in Figure 13, using the SLAM algorithm. Figure 14 shows a top view of the robot arm pose to grab an object after the mobile base reaches its goal destination.

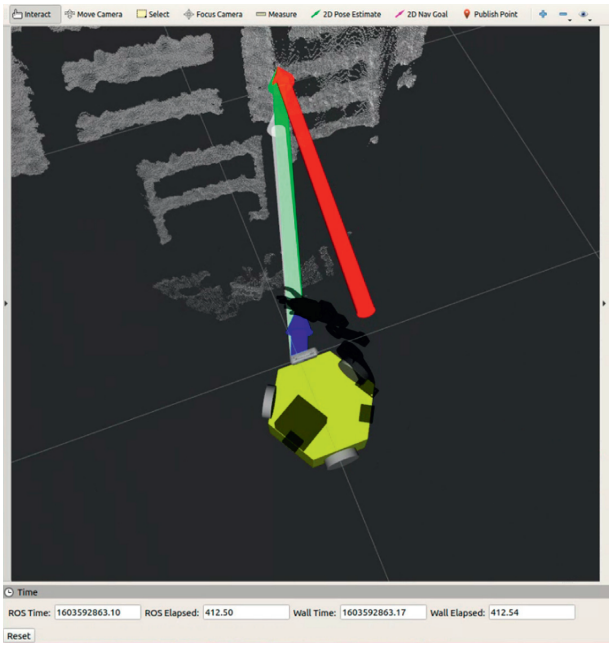


Figure 14: Robot arm movement to grabbing point.

9. Conclusion

This study implemented three independently developed subsystems: a hybrid BCI controller, an omnidirectional mobile base, and a robotic arm into one semiautonomous assistive robot. Each subsystem of the robot was tested independently and allows for changes to be easily made based on the desired application. In these tests, goal location steady state errors were minimized to ensure successful grabbing operations. The errors associated with the velocity controller performance of the wheels are addressed to have more realistic simulations, but the goal location of the mobile base using SLAM places the robot on the desired location. The process to tune the controllers of a physical robot follows the same procedure and can be done in the ROS environment, allowing for quick deployment based on simulation results and errors. Future work is to use the live data to test robot usability with added delays in the system. This test will accurately highlight the aspects of the system which need improvement to design a responsive, intuitive BCI controlled robot. The camera feed comes from an Intel RealSense Depth Camera D435 to verify its use in the physical build of the robot. To develop more robust tests using live data in this environment, virtual point cloud data should be used to interact with the virtual environment directly before physical testing. The manipulator was tested independently in a lab environment, but, due to pandemic access restrictions, different end-effectors were imported to this virtual environment to test the grabbing and placing command loops and confirmation using BCI. The robot control environment developed during this research allows a user to interact with a simple interface to achieve complex robot movements.

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] World Health Organization (WHO), *World Report on Disability*, WHO Press, Geneva, Switzerland, 2011.

[2] P. F. Edemekong, D. L. Bomgaars, S. Sukumaran, and S. B. Levy, *Activities of Daily Living*, StatPearls Publishing, Treasure Island, FL, USA, 2022.

[3] M. Bergmann and M. Wagner, "The impact of COVID-19 on informal caregiving and care receiving across Europe during the first Phase of the pandemic," *Frontiers in Public Health*, vol. 9, pp. 1–17, 2021.

[4] R. Spataro, A. Chella, B. Allison et al., "Reaching and grasping a glass of water by locked-in ALS patients through a BCI-controlled humanoid robot," *Frontiers in Human Neuroscience*, vol. 11, no. 68, pp. 68–10, 2017.

[5] J. Tang, Y. Liu, D. Hu, and Z. Zhou, "Towards BCI-actuated smart wheelchair system," *BioMedical Engineering Online*, vol. 17, no. 1, p. 111, 2018.

[6] D. Huang, K. Qian, D.-Y. Fei, W. Jia, X. Chen, and O. Bai, "Electroencephalography (EEG)-based brain-computer interface (BCI): a 2-D virtual wheelchair control based on event-related desynchronization/synchronization and state control," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 20, no. 3, pp. 379–388, 2012.

[7] R. Leeb, D. Friedman, R. G. Muller-Putz, R. Scherer, M. Slater, and G. Pfurtscheller, "Self-paced (asynchronous) BCI control of a wheelchair in virtual environments: a case study with a tetraplegic," *Computational Intelligence and Neuroscience*, vol. 2007, Article ID 79642, 8 pages, 2007.

[8] R. Bousseta, I. El Ouakouak, M. Gharbi, and F. Regragui, "EEG based brain computer interface for controlling a robot arm movement through thought," *IRBM*, vol. 39, no. 2, pp. 129–135, 2018.

[9] A. Campeau-Lecours, V. Maheu, S. Lepage et al., "JACO assistive robotic device: empowering people with disabilities through innovative algorithms," in *Proceedings of the Rehabilitation Engineering and Assistive Technology Society of North America (RESNA)*, Alrington, TX, USA, 2016.

[10] S. H. Sunny, I. I. Zarif, I. Rulik et al., "Eye-gaze control of a wheelchair mounted 6DOF assistive robot for activities of daily living," *Journal of NeuroEngineering and Rehabilitation*, vol. 18, no. 173, pp. 1–12, 2021.

[11] M. Seeck, L. Koessler, T. Bast et al., "The standardized EEG electrode array of the IFCN," *Clinical Neurophysiology*, vol. 128, no. 10, pp. 1–8, 2017.

[12] K. Matsuno and V. K. Nandikolla, "Machine learning using brain computer interface system," in *Proceedings of the ASME 2020 International Mechanical Engineering Congress and Exposition*, Portland, OR, USA, 2020.

[13] C. A. Kothe, D. Medine, C. Boulay, M. Grivich, and T. Stenner, "LabStreamingLayer's documentation," 2019, https://labstreaminglayer.readthedocs.io/index.html.

[14] A. Florentina and D. Ioan, "Practical applications for mobile robots based on mecanum wheels—a systematic survey," *Romanian Review Precision Mechanics, Optics and Mechatronics*, vol. 40, pp. 21–29, 2011.

[15] V. Kálmán, *On Modeling and Control of Omnidirectional Wheels*, Budapest University of Technology and Economics, Budapest, Hungary, 2013.

[16] K. M. Lynch and F. C. Park, *Modern Robotics: Mechanics, Planning, and Control*, Cambridge University Press, Cambridge, UK, 2017.

[17] J. Tang, Z. Zhou, and Y. Yu, "A hybrid brain computer interface for robot arm control," in *Proceedings of the 2016 8th International Conference on Information Technology in Medicine and Education*, pp. 365–369, Fuzhou, China, 2016.

[18] J. Kofrnan, X. Wu, T. J. Luu, and S. Verma, "Teleoperation of a robot manipulator using a vision-based human-robot interface," *IEEE Transactions on Industrial Electronics*, vol. 52, no. 5, pp. 1206–1219, 2005.

[19] B. Landavazo and V. K. Nandikolla, "Brain-computer interface application in robotic gripper control," in *Proceedings of the ASME 2018 International Mechanical Engineering Congress and Exposition*, Pittsburgh, PA, USA, 2018.

[20] P. Kim, J. Chen, and Y. K. Cho, "SLAM-driven robotic mapping and registration of 3D point clouds," *Automation in Construction*, vol. 89, pp. 38–48, 2018.

[21] K. Yousif, A. Bab-Hadiashar, and R. Hoseinnezhad, "An overview to visual odometry and visual SLAM: applications to mobile robotics," *Intelligent Industrial Systems*, vol. 1, no. 4, pp. 289–311, 2015.

[22] R. Martinez-Cantin, N. D. Freitas, and J. A. Castellanos, "Analysis of particle methods for simultaneous robot localization and mapping and a new algorithm: marginal-SLAM," in *Proceedings of the 2007 IEEE International Conference on Robotics and Automation*, Rome, Italy, 2007.

[23] ROS Basics for AMCL Documentation, http://wiki.ros.org/amcl.

[24] ROS Basics for Move Base Documentation, http://wiki.ros.org/move_base.

[25] ROS Basics for DWA Local Planner Description, http://wiki.ros.org/dwa_local_planner?distro=melodic.

Hindawi
Journal of Robotics
Volume 2023, Article ID 9785409, 1 page
https://doi.org/10.1155/2023/9785409

Retraction

Retracted: Research on Clue Mining in Criminal Cases of Smart Phone Trojan Horse under the Background of Information Security

Journal of Robotics

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 investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

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

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

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

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

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

References

[1] L. Gang and Y. Wen, "Research on Clue Mining in Criminal Cases of Smart Phone Trojan Horse under the Background of Information Security," *Journal of Robotics*, vol. 2022, Article ID 7568110, 11 pages, 2022.

Hindawi
Journal of Robotics
Volume 2022, Article ID 7568110, 11 pages
https://doi.org/10.1155/2022/7568110

Research Article

Research on Clue Mining in Criminal Cases of Smart Phone Trojan Horse under the Background of Information Security

Li Gang[1] and Yong Wen[2]

[1]*Guangxi Police College, Nanning, Guangxi 530028, China*
[2]*School of Artificial Intelligence, Guangxi University for Nationalities, Nanning, Guangxi 530006, China*

Correspondence should be addressed to Yong Wen; 20190304001@stu.gxun.edu.cn

Received 1 December 2021; Accepted 6 January 2022; Published 25 February 2022

Academic Editor: Shan Zhong

With the massive popularity and wide application of Android smartphones, there are more and more malware targeting Android smartphones. Research and analysis Android smart phone Trojan horses can provide corresponding technical support for malware detection on Android smart phones, which has good scientific research significance and broad market value. This work studies and analyzes the existing implantation technology of Android smart phone Trojans and analyzes the basic principles and implementation methods of obtaining root permissions on mobile phones. At the same time, the basic principles and implementation methods of mobile phone Trojan horse hiding are also studied. Through the research of the broadcast receiver model of the Android platform, the background monitoring principle and implementation technology of the mobile phone Trojan horse are analyzed, and the theoretical foundation and technical support are provided for the implementation of the Trojan horse background monitoring program in this article. Aiming at the problem of insufficient training corpus in the event relationship classification task, this work proposes an event relationship classification method based on tritraining. This method first trains three different classifiers based on the labeled dataset. In the collaborative training process, the new labeled event pairs used to expand each classifier are provided by the other two classifiers. For the same unlabeled event pair, the relationship prediction results are consistent; then, the event pair is considered to have a higher classification confidence and is placed in the labeled set of the third classifier after labeling. Finally, a well-trained classifier is used to determine the relationship between the pair of events to be tested by voting. This study constructs a weighted network structure model called the conceptual network and determines its upper weight based on the structural information and text data of the knowledge network. Aiming at the problem of the lack of means for mining-related forms between things, the pheromone strategy of absorbing the ant colony algorithm is proposed, and random walks are performed on the conceptual network. By analyzing the pheromone distribution information in the convergent state, the calculation of the semantic relevance is completed. At the same time, the method of semantic clue discovery is realized. The experimental results show that the human cognitive information contained in the knowledge network can meet the needs of the mining of the related forms of things, and the performance of the semantic correlation calculation method based on the convergence pheromone is close to other random walk methods based on the knowledge network.

1. Introduction

With the development and prosperity of the mobile Internet, smart phones have gradually become an indispensable tool in our daily lives due to their powerful computing power and portability [1]. According to data from the China Internet Network Information Center, there were more than 700 million mobile Internet users in China in 2019, surpassing the number of computer Internet users. Due to the proliferation of smart phones and the privacy of data in smart phones, we also pay more attention to the security of smart phones. Although the open source, open, free, and other features of the Android platform have brought a lot of market share, it also brings many security risks to smartphone users. The makers of smartphone Trojan horses can implant malicious code in normal mobile phone

applications and release applications bundled with mobile phone Trojan horses on the Android platform arbitrarily [2, 3]. In 2020, the number of Android smartphone users infected by malware exceeded 10 million, making it the hardest hit area for mobile Trojan horses. The total number of new malware for Android mobile phones in China exceeds 3,000. According to the 2020 security threat report released by the NetQin Security Center, Android smartphones have become the smartphones most concerned about mobile malware.

Network information is closely related to social life, so the security of network information is essential to the normal operation of social life [4, 5]. For individuals, the interference or leakage of network information may cause the infringement of personal privacy information; for enterprises, network information security is related to the normal production and operation of enterprises. Therefore, network information security has become an important issue in social life, and ensuring network information security has become one of the most important tasks of the information supervision department [6]. There are many existing network information security control systems, which can monitor and rectify various network information security issues that may have occurred in the past. However, these methods have different degrees of limitations and deficiencies for today's huge amount of data [7]. The huge amount of big data, faster transmission speed requirements, and more effective value extraction methods make the previous management methods no longer fully suitable for network information security control in the era of big data [8]. Therefore, analyzing the characteristics of big data and establishing a more complete network information security control mechanism based on the characteristics of network information in the big data era is to ensure greater efficiency of big data traffic transmission and realize more secure network information functions. Using the network information security evaluation system to improve the reliability of data security during the entire life cycle of network information data can provide more complete working standards for the high-performance realization of the systematic and scientific development of network security [9].

This work studies the implantation technology of Android smart phone Trojan horses, analyzes the basic methods of mobile phone Trojan horses using the principles of social engineering to achieve deceptive implantation of Trojan horses, and analyzes the basic principles and implementation methods of obtaining mobile phone root permissions. At the same time, it studies the hiding technology of mobile phone Trojans, analyzes the basic principles and implementation methods of file hiding, process hiding, network connection hiding, and kernel module hiding, and analyzes the basic methods of communication hiding technology and collaborative hiding. Then, by analyzing the broadcast receiver model of the Android platform, the background monitoring principle and technology of the mobile phone Trojan horse are studied and analyzed. This work studies the application of the semisupervised learning method based on tritraining in the task of event relationship classification. Aiming at the lack of training corpus in the event relationship classification task, this method uses the tri-training method to train three different classifiers on a small amount of existing manually labeled event relationship datasets and then extracts from the mined ones that contain connectives. In a large number of unlabeled datasets, the training corpus is expanded through simple voting, repeated iterations, and continuous optimization of the classification model, ultimately achieving the goal of improving the performance of event relationship classification. This study tests the execution performance of the semantic relevance calculation model and the semantic clue mining model and designs an experimental program. In the experiment, the calculation results of the semantic relevance calculation model are compared with similar algorithms, and the results show that the model has better calculation accuracy; in addition, the stability of the semantic clue mining model is tested, and the results show that the effective path distribution of the model's feedback is in line with expectations and has basic stability.

2. Related Work

Regarding the current research status of network information security in the context of big data, relevant scholars have made research reports on related issues of big data [10]. The reports all propose that in the era of big data, data security needs to be more effectively protected [11]. It is an important means to maintain information security, and technological advancements play an important role in promoting the progress of information security work. Research believes that in the era of big data, the processing and analysis of information security data have become relatively difficult [12]. In the face of information security challenges under new circumstances, it is necessary to combine multiple methods and means to comprehensively view information security issues. In terms of information security evaluation research, relevant scholars have put forward specific implementation methods for information security evaluation through research and built an enterprise information security evaluation model based on the information security management framework [13].

In the area of network information security strategy research under the background of big data, scholars discussed the necessity of building a national competitive intelligence system based on network information security and proposed measures to build a national competitive intelligence system based on network information security [14]. Researchers proposed that intelligence literacy should be regarded as the core element of information security theory and explored solutions to information security theory in the context of big data [15]. Relevant scholars believe that due to the complexity of network information security, whether it is theoretically or technically, network information security problems cannot be completely solved [16]. Therefore, network information security prevention technologies should be combined with other technologies to use existing technologies. Relevant scholars analyzed the problems facing my country's network information security from the "Prism Gate" incident and put forward the viewpoint of

constructing my country's network information security strategy from the perspective of the rule of law [17]. Based on the analysis of the importance of network information security in the era of big data, scholars proposed that network information security must be promoted to the height of national security strategy, unified deployment of the top-level design of network security work, and mobilized the entire society to implement network information security work [18]. Relevant scholars analyzed information security risks that may occur in the context of big data from 9 perspectives, including infrastructure and data processing, and constructed a big data information security risk framework and proposed corresponding solutions [19].

Related scholars use a weakly supervised machine learning method to classify the temporal relationship between events [20]. First, they learn a general classification template from the labeled corpus; then, based on the assumption that "each document contains only one temporal relationship subtype," the documents are clustered according to the temporal relationship subtype, that is, the subtypes of the temporal relationship are consistent. The documents are clustered together.

Related scholars have proposed an extraction method for "event pairs" with causal relationships [21]. The article combines the "relation-oriented template-based" method and the "slot-oriented attribute-based" method to mine "event pairs" resources with sequential relationships. First, it mines the causal "predicate pairs" according to the predefined template; then, it uses the two predefined "type slots" to filter the extracted "predicate pairs" and finally obtains the causal "event pair" resources.

Researchers first search for "concept pairs" with relationships in the text, then use the relationship types defined in WordNet to identify explicit causal "event pairs" in the text, and use the extracted results in the answering system [22]. It uses machine learning methods to classify these two types of event relationships (chronological relationship and causal relationship) on manually labeled corpus with temporal relationship and causal relationship [23]. At the same time, a fully supervised classification method can be used to improve the accuracy of event relationship detection. Related scholars use a fully supervised classifier based on the support vector machines (SVMs) algorithm to perform "relational event pairs" on the four causal relationships (i.e., "causes," "effects," "preconditions," and "postconditions") [24].

The event relationship detection method based on pattern matching can identify more fine-grained event types. Relevant scholars have used pattern matching methods to identify six types of event relationships [25]. Relevant scholars use the method of pattern matching to classify and identify the subtypes of event causality [26]. However, the accuracy of the event relationship detection method that uses pattern matching alone is low. Therefore, researchers propose an event relationship detection method that combines pattern matching and event-oriented element filtering to filter the "relational event pairs" obtained from template matching [27]. Experimental results show that an event relationship detection system that combines pattern matching and event element filtering can effectively improve the accuracy of recognition. Relevant scholars also adopted a fusion of template matching and rule-based methods, and the experimental results show that the performance of the system is significantly better than template matching and rule-based performance [28].

3. Methods

3.1. The Implantation Technology of Android Smart Phone Trojan Horse. Due to the open source and openness of Android, the makers of mobile phone Trojan horses can bundle the Trojan horse programs in some normal applications. When users install these legitimate applications, they also install the Trojan horse programs. Although all Android applications must have digital signatures, in theory, applications with the same digital signature can safely share data resources. Applications with different signatures can also access signature-based APIs by granting permissions to each other. The digital certificate does not require the certification of an authority, so the Trojan horse can complete the signature certification by himself and release the application bundled with the Trojan horse at will to bind a Trojan horse to a legal file or application, and it is necessary to study the source code and vulnerabilities of the application. Therefore, the installation of the Trojan horse also needs to study the vulnerabilities of the application used by the target mobile phone and develop the corresponding binding based on its vulnerabilities. Some Trojan horse programs use popular QR codes to bind malicious codes into legal files (for example, implant the source code of Trojan horses into popular theme files) and use technology to pretend to trick mobile phone users into scanning the QR code. There are also some Trojan horse programs that have been implanted in the mobile phone, using the mobile phone user's contact information and using the user's trustworthy social relationship to mass transit network links containing Trojan horse programs to trick other mobile phone users into opening the link to achieve mobile phone Trojan horses. Therefore, the implantation of Trojan horses requires the use of social engineering methods to target the implantation and spread of Trojan horses.

On the Android platform because root permissions can access all applications and data resources, the makers of mobile phone Trojan horses work hard to study the Trojan horses that obtain root permissions. Although the Android platform only allows the system kernel and very few core programs to run with root permissions, the Android platform allows users or applications to obtain root permissions by modifying the source code of the system kernel or core programs.

Generally, mobile phone users do not have root permissions on Android smartphones, so the makers of mobile Trojan horses can take advantage of vulnerabilities in the Android system to obtain root permissions on mobile phones by flashing the phone, using rootkit technology, and modifying the Android source code.

3.2. Background Monitoring Technology of Android Smartphone Trojan Horse. In the Android system, broadcast is an event generated by the operating system and a mechanism for transferring messages between applications. The broadcast receiver is a component provided for the

realization of system broadcasts, which can receive and respond to broadcasts. Intent is the medium used to store broadcast messages. The Android platform can either use Intent to start a component or use send Broadcast() to initiate a system-level event broadcast to transmit messages. The Android system allows programmers to develop a broadcast receiver in the application, and then, register the broadcast receiver to the Android system and notify the operating system that there is such a broadcast receiver waiting to receive the broadcast of the Android system, so that the broadcast receiver can monitor and respond to the broadcast. Each broadcast receiver can receive Intent triggered by one or more events. When an event occurs, the system will send a message to the broadcast receiver.

When the broadcast receiver receives a broadcast message, it creates an extended class that inherits the broadcast receiver and executes the on Receive() function. After executing the on Receive() function, the extended class that inherits the broadcast receiver will be destroyed. If the on Receive() function is not executed within a short period of time, the Android system will consider the program to be unresponsive. Therefore, some time-consuming operations cannot be done in the broadcast receiver. If you need to complete some time-consuming operations, you should send an Intent to the service, and the service will complete it.

As shown in Figure 1, the processing principle for the broadcast receiver is the same regardless of whether it is sending a normal broadcast or an ordered broadcast. The Android system needs to manage the registered Receiver, pass the filtered system message to the Receiver implementation class, trigger the on Receive() function, and pass the Intent object passed by the sender to the on Receive() function. The sender passes the message through the Intent object. Android compares the description in the Intent object with the description in the configuration file Android Manifest.xml, finds out the component (Receiver) that matches the description of the Intent object, and then passes the Intent object as a parameter to the on Receive () in Receiver function. When the broadcast receiver receives the relevant message, it executes the corresponding processing program, such as starting an activity for interaction or opening a service.

In this study, the problem of determining event relations is transformed into a binary classification problem, that is, to classify various event relations separately. The classification results are "this relationship" and "not this relationship." In the test process, for the test event pair p, the classifiers C1, C2, and C3, respectively, give the classification results R1, R2, and R3. Then, we count the results that appear most in R1, R2, and R3 as the final classification result of p. For example, for binary classification of "causal relationship," the three classifiers C1, C2, and C3 will give the results of "it is the relationship," "it is the relationship" and "not the relationship," and the final relationship of the test event pair is judged as "this relationship," that is, "causal relationship."

3.3. Design Ideas and System Framework. Aiming at the problem of insufficient training corpus in the event relationship classification task, this study uses a collaborative training algorithm to propose an event relationship classification method based on tritraining. This method first trains three different classifiers based on the labeled corpus. On this basis, the training set is expanded by selecting samples with higher confidence in the classification from the unlabeled set by majority voting. Then, they use the expanded training set to retrain the classifier, loop and iterate repeatedly, and continuously improve the classification model until the termination condition is met. Finally, three well-trained classifiers are used to determine the event relationship, and the final event relationship type is also output by voting. The system framework of the event relationship classification method based on tritraining is shown in Figure 2.

3.4. Cotraining Method Based on Tritraining. Aiming at the problem of insufficient event relationship training corpus, this study proposes an event relationship classification method based on tritraining. This method uses Gigaword linguistic resources as external data resources, mines event relation pairs containing connectives from it, and uses the relational categories corresponding to connectives as prior knowledge.

On this basis, four features of frame semantics, event trigger word, trigger word part of speech, and event category are extracted to generate an unlabeled event relationship dataset. The labeled set is generated after extracting the same features from a small number of manually labeled event relationship datasets.

In this study, half of the labeled set is used as the training set, and the other half is used as the test set. Then, we use the tritraining method to select higher confidence event pairs from the unlabeled set and add them to the training set and iteratively train the model until the unlabeled set is empty or the size of the unlabeled set no longer changes. Finally, the test set is classified according to the learned model, and the final event relationship classification result is generated.

The tritraining algorithm is designed to repeatedly sample a small number of labeled datasets, train three different classifiers X, Y, Z, and classify and label a large number of unlabeled sample data through the consistency judgment between the three classifiers. In detail, during the training process, the newly labeled samples obtained by any one classifier (for example, classifier X) are jointly determined by the other two classifiers (classifier Y and classifier Z). If two classifiers give the same unlabeled sample x the same classification label L, that is, $Y(x) = Z(x)$, then the unlabeled sample is added to the current classifier X.

It is worth noting that for an unlabeled event relationship pair, when the classification results given by the two classifiers are the same, the same conditions as the prior relationship category must also be met before the event pair will be added. For example, if the prior relationship category of an unlabeled event pair is expansion and the categories given by classifiers C1 and C2 are expansion, then the event pair will be added to the labeled set L, and the category will be marked expansion.

When the cooperative training stop condition is met, that is, when the size of the unlabeled dataset no longer

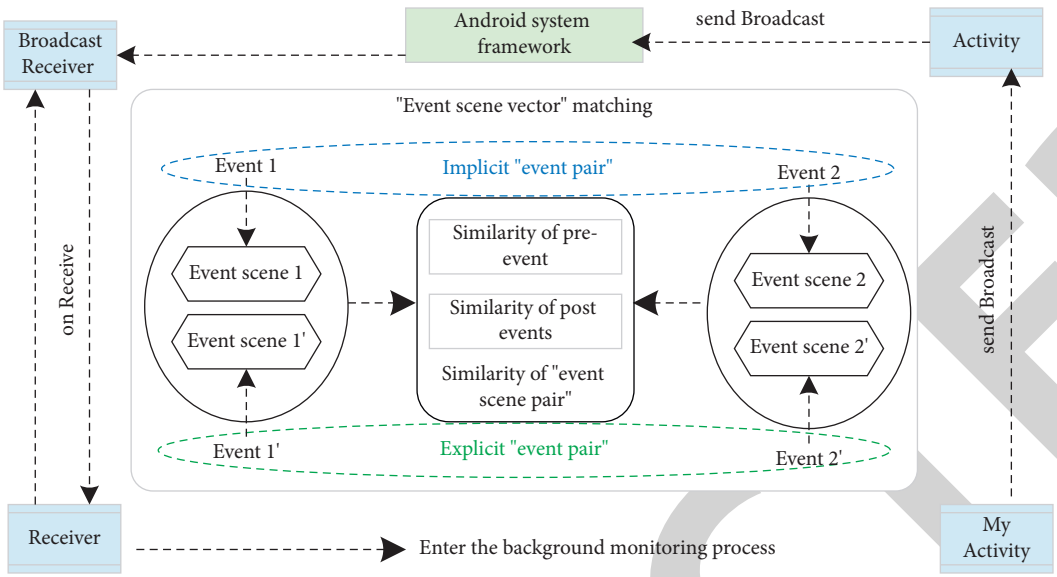


Figure 1: Broadcast receiver model.

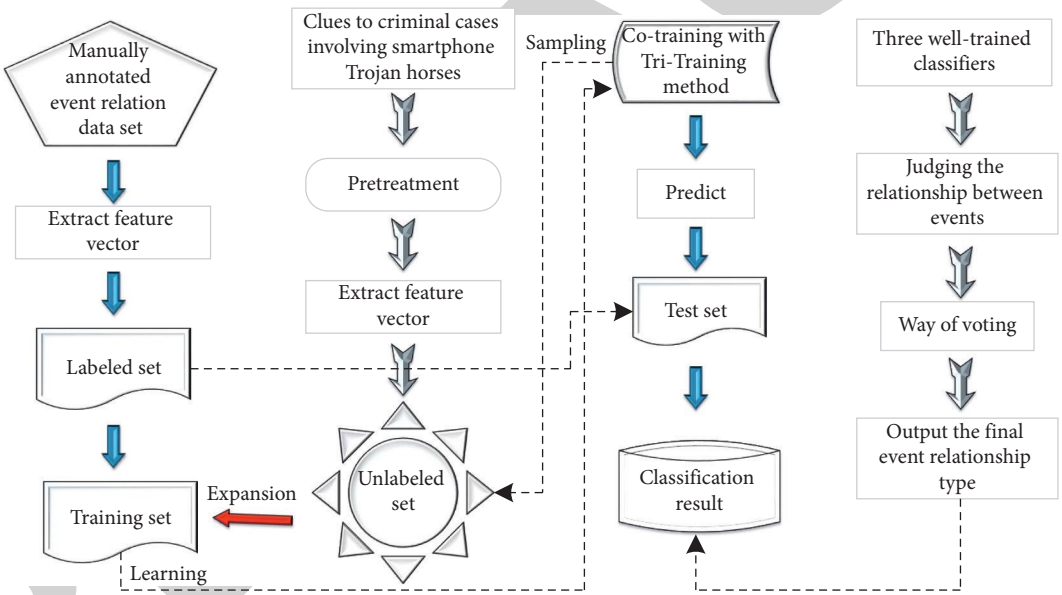


Figure 2: Flowchart of event relationship classification method based on tritraining.

changes or is empty, the training is stopped. At this point, three well-trained classifiers C1, C2, and C3 are obtained, and these three classifiers are used to classify and determine the relationship between the event pairs in the test set.

3.5. Implicit Relationship Detection Based on the Conceptual Model. Figure 3 shows the implicit relation detection framework based on the conceptual model of parallel arguments. This study extracts key information from the implicit arguments to be tested, mines parallel arguments containing connection clue words (explicit connectives and functional connectives) based on a large-scale local corpus, and classifies parallel arguments related to the same clue words. The reason is that similar knowledge is involved in similar arguments, which can form semantically targeted conceptual descriptions. For the parallel reference argument set and the implicit argument to be tested after classification, feature extraction and attribute description are performed, respectively, to construct conceptual models A and B (knowledge units formed by unique combination of argument features). The two conceptual models measure parallelism through similarity matching. Through the use of feature vector similarity and conceptual submodel similarity measurement methods, statistical methods are used to obtain conceptual model A with the highest similarity to conceptual model B, and the largest possible text relationship is output through the mapping system as the final result.

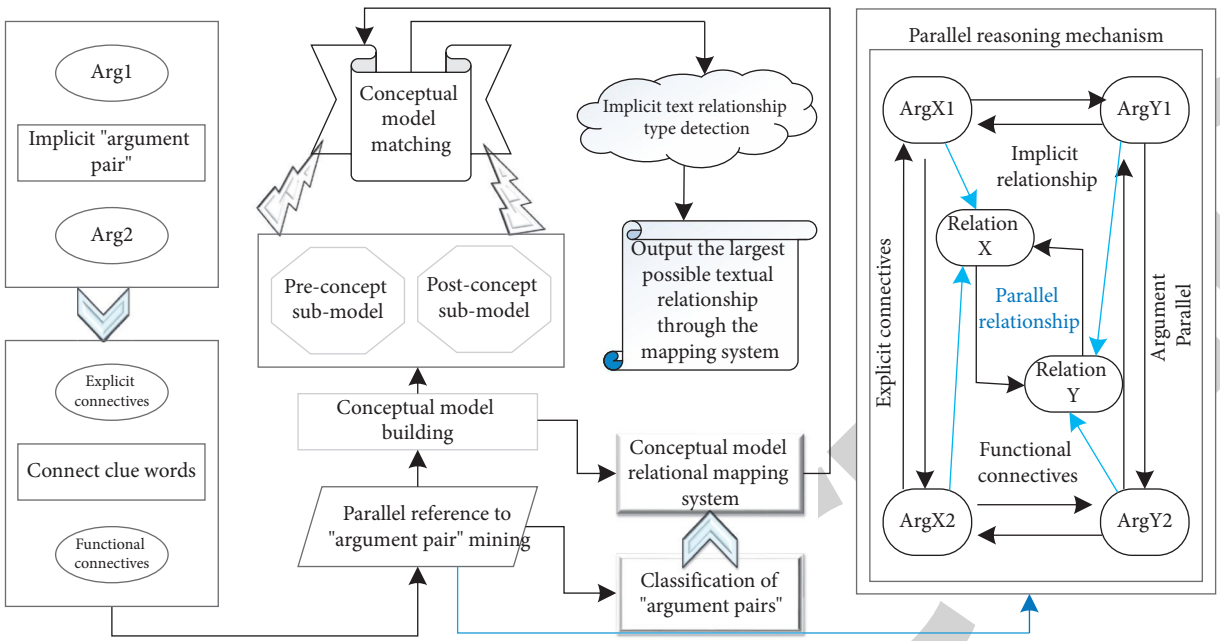


Figure 3: Implicit relationship detection framework based on the parallel argument concept model.

Semantic parallelism calculation consists of two parts: one is the balance calculation model and the other is the parallel calculation model. Among them, the balance calculation model attempts to solve the problem of balance between the preargument Arg1 and the postargument Arg2 in the calculation of semantic parallelism. The parallelism of the preargument Arg1 of the two types of "argument pairs" is PreSim, and the parallelism of the postargument Arg2 is PosSim. After the numerical balance of PreSim and PosSim, the overall "argument pair" is formed. The calculation method is as follows. Among them, the exponential operation of the denominator plays a normalizing role:

$$\mathrm{Sim}(a_t, a_c) = \frac{\mathrm{PreSim} - \mathrm{PosSim}}{4} e^{-2(\mathrm{PreSim} - \mathrm{PosSim})}. \quad (1)$$

The parallel calculation model is used to calculate the parallelism PreSim and PosSim. This method uses the vector space model (VSM) as the semantic description structure of the argument and uses the Jaccard algorithm to calculate the parallelism. The Jaccard algorithm is

$$J = \frac{a_t \bullet 2a_c}{a_t \bullet a_c}(a_t - a_c). \quad (2)$$

Among them, J is the Jaccard metric value of PreSim and PosSim. a_t and a_c, respectively, represent the feature vector of Arg1 or Arg2 in the "argument pair" to be tested and the candidate parallel "argument pair." When PosSim is calculated, a_t and a_c are the feature vectors of Arg2 in the "argument pair" to be tested and the candidate parallel "argument pair," respectively. The numerator is the inner product operation, that is, the weights of the corresponding dimensions of at and ac are, respectively, multiplied and then summed; the absolute value of the denominator part represents the length of the feature vector. The weight of each dimension of the feature vector is calculated by TFIDF, where TF is the frequency of the word feature where the argument is located (obtained after filtering by stop words). The TF of the word features in "pair" is counted, respectively, on the PDTB where the "argument pair" is located and the GIGAWORD local static corpus; the IDF of the antidocument frequency is counted by Gigaword.

An important part of the method framework of this study is the construction of a mapping system that maps conceptual model B to conceptual model A. The mapping of the two conceptual models involves the similarity matching of the two, and the overall similarity measurement method is

$$\mathrm{Similarity} = \frac{\mathrm{PosCSMSim} - \mathrm{PreCSMSim}}{4} e^{-(\mathrm{PosCSMSim} - 2\mathrm{PreCSMSim})}. \quad (3)$$

Among them, PreCSMSim and PosCSMSim, respectively, represent the similarity between the preconcept submodels and the similarity between the postconcept submodels. The denominator in the formula plays a normalizing role.

The calculation process of PreCSMSim and PosCSMSim involves calculation of the similarity of three pairs of conceptual submodels. Each concept submodel is composed of several clusters, and the elements in each cluster are vectors (such as KeyWordi) that characterize the attributes of this

type of cluster. Therefore, the similarity of a pair of preentity concept submodels is the similarity of the two clusters, that is, the final refinement is the similarity of the elements in the clusters. The measurement method is

$$\mathrm{Sim\_}FV(X,Y) = \prod_{k=0}^{N_1} Set_{XY}(k) - \prod_{i=0}^{N_2} OP_{XY}(i) - \prod_{j=0}^{N_3} Val_{XY}(j). \quad (4)$$

Among them, X and Y, respectively, represent the elements in the two clusters (such as KeyWordi and KeyWordj) that need to be similarly calculated, and their similarity is the sum of the similarity weights of the 9-dimensional features. Since each dimension feature has both a numerical form and a word set form, it cannot be directly calculated using the vector space model (VSM).

$$\mathrm{Val}_{XY}(j) = \frac{F_j(X) + F_j(Y)}{\mathrm{Min}\left[S_j(X), S_j(Y)\right]} - \frac{F_j(X) - F_j(Y)}{\mathrm{Max}\left[S_j(X), S_j(Y)\right]}, \quad (5)$$

$$DF = \log_2 n - \log_2 (N-1).$$

When $j = 1$, $F_j(X)$ and $F_j(Y)$ are the position eigenvalues in X and Y; $S_j(X)$ and $S_j(Y)$ are the lengths of arguments to construct X and Y, which are obtained after normalization. When $j = 2$, $F_j(X)$ and $F_j(Y)$ are the DF values in X and Y (n is the number of argument categories containing characteristic words, and N is the total number of argument categories); $S_j(X)$ and S_j (Y) are the respective n values. After normalization, the similarity weight of the DF value of X and Y is obtained.

$$\mathrm{Set}_{XY}(k) = 1 - \frac{G\left[S_k(X), S_k(Y)\right]}{Min\left[N_k(X), N_k(Y)\right]} + \frac{G\left[S_k(X), S_k(Y)\right]}{Max\left[N_k(X), N_k(Y)\right]}. \quad (6)$$

$S_k(X)$ and $S_k(Y)$ represent the single-sentence dependent word set, cross-sentence dependent word set, synonym set, upper word set, and lower word set (according to the value of k) of the respective characteristic words in X and Y; $G(S_k(X), S_k(Y))$ represents the number of word co-occurrences of the feature word set corresponding to X and Y (the number of cross words in the word set); $\mathrm{Max}(N_k(X), N_k(Y))$ represents the respective feature words of X and Y.

The similarity calculation between the preentity concept submodels (consisting of multiple clusters) in the two conceptual models is based on the similarity calculation of feature vectors. The calculation of the similarity between the submodels is the calculation of the similarity between the two clusters. Similarity matching needs to be classified and matched according to the "attribution" of the concept. For example, the pre-subconcept of the "argument pair" to be tested must match the pre-sub-subconcept of the parallel reference "argument pair," but the post-subconcept or the sub-subconcept of the parallel reference "argument pair" cannot be selected. Through the "concept model-text relationship" mapping system, that is, the text relationship to which the connected clue words corresponding to the statistically optimally matched conceptual model belongs, the textual semantic relationship of the "argument pair" to be tested is inferred.

4. Results and Analysis

4.1. Model Parameter Setting. In order to determine the influence of different types of subnetworks in the calculation of semantic relevance, three groups of conceptual networks with different components were prepared in the experiment: a conceptual network based entirely on hyperlinks and a concept with 50% similarity between hyperlinks and texts. The network is completely based on the conceptual network of text similarity. Then, on these three groups of conceptual networks, using the data of illegal and criminal cases as the dataset, we perform semantic walks and sort out the path data obtained from the walks to obtain the distribution of the number of paths between conceptual nodes, as shown in Figure 4. The use of the hyperlink structure increases the connectivity between conceptual nodes, increases the effective path between nodes, and reduces the average length of the path at the same time.

The hyperlink structure is added to the conceptual network based on text similarity, and the calculation method of semantic relevance in this study is used to calculate based on the data of illegal and criminal cases. With the difference in the ratio of the hyperlink structure and the text similarity structure, the performance change of the algorithm is shown in Figure 5. It can be seen that when the weight ratio of the concept subnetwork based on text similarity and the concept subnetwork based on hyperlink is close to 0.5, the Pearson correlation coefficient reaches the highest value of 0.78. The following experiment is implemented based on a weight ratio of 0.5.

In this study, the pheromone residual ratio is set to 0.85, and the value of the termination probability is determined through experiments. In order to determine a reasonable value, a test program is set up in this article, and K in TOP-K is set to 50. The length distribution of the obtained path is shown in Figure 6. It can be seen from the figure that most of the paths are paths less than or equal to 240 in length. A small termination probability value will make the wandering agent explore a longer path. Taking into account that the termination probability should be set as far as possible to take into account the path coverage and execution cost, the termination probability is set to 0.3.

4.2. Semantic Relevance Calculation. The semantic relevance calculation model proposed in this study is used to perform semantic calculation on the word pairs, and the calculation results are compared with the average value of the expert evaluation results of the word pairs in the criminal case data, and the Spearman rank correlation coefficient is used as the comparison result measurement index. At the same time, the calculation performance of the algorithm in this study is compared with similar algorithms in the data of crimes.

Figure 7(a) shows the value calculated by the calculation model of semantic relevance in this study for the vocabulary pairs in the data of crimes. Figure 7(b) shows the expert evaluation values. It can be seen from the figure that the two have a certain linear relationship.

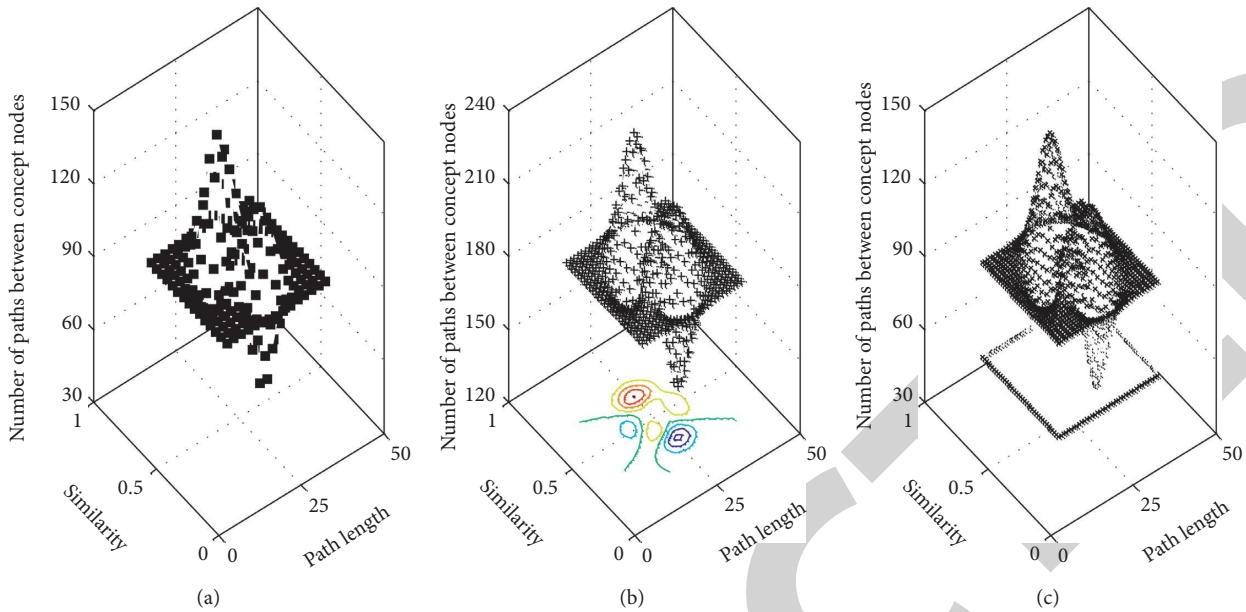


Figure 4: Distribution of the number of paths between conceptual nodes. (a) A conceptual network based entirely on hyperlinks. (b) A conceptual network with 50% similarity between hyperlinks and text. (c) A conceptual network based entirely on text similarity.

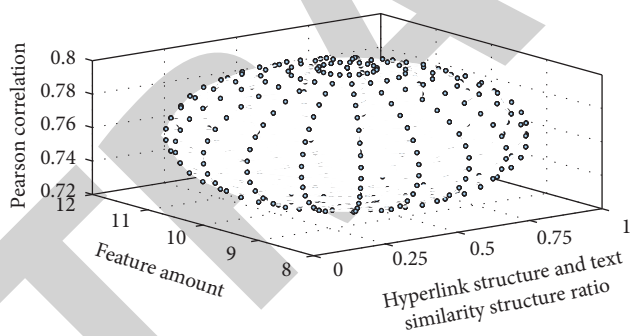


Figure 5: The influence of different ratios of concept subnetworks on algorithm performance (Pearson correlation coefficient).

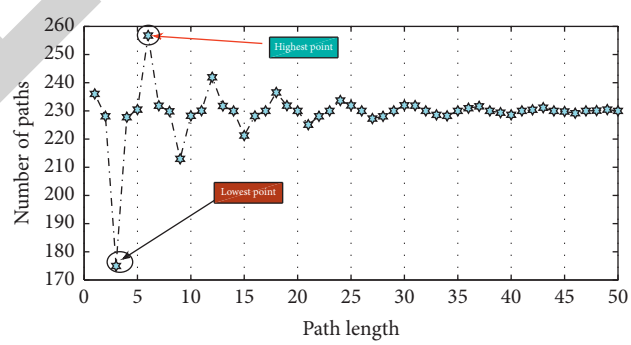


Figure 6: The length distribution of the sampling connection path.

In order to quantify the relationship between the settlement results of the semantic relevance calculation model in this study and the expert judgment standard dataset, to test the calculation performance of the algorithm in this study, the experiment uses Spearman's rank correlation coefficient as a measure, and the correlation coefficient is 0.74.

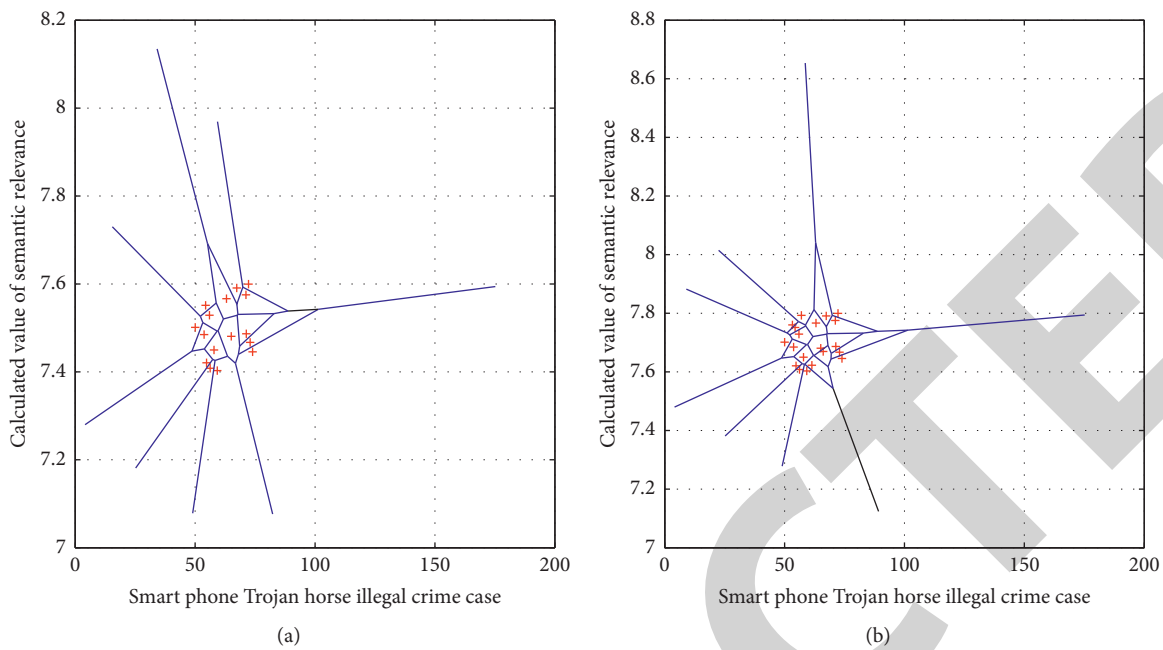


Figure 7: Data on illegal and criminal cases, benchmarks, and algorithmic calculation results. (a) Calculated value of semantic relevance. (b) Expert evaluation value.

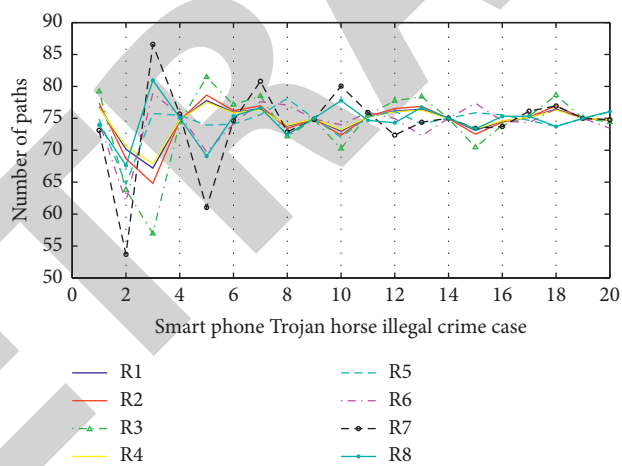


Figure 8: Distribution diagram of the sequence matching of each group of paths.

4.3. Semantic Cue Mining. For the test of semantic clue mining results, this study adopts the method of comparing with the results of manual judgment: this study randomly selects 5 pairs of vocabulary from the data of illegal and criminal cases and divides the experiment into 8 groups according to the vocabulary and denoted by Ri. We map the selected 20 words to the corresponding concept explanation page in the open domain knowledge network and manually click the hyperlink of the page body content from the i1. Special pages encountered during the process are not included in the association path. In the process of exploring the association path of each vocabulary pair, each experiment participant has 5 opportunities to obtain as many different association paths as possible. In this study, Jaccard coefficient is used as a measure of the accuracy of semantic clues.

Figure 8 shows the corresponding distribution of each group of comparison data in TOP-K in the extended example experiment for feedback. Judging from the data of 8 groups of experimental feedback, the experimental data did not appear "inverted," and the experimental results are consistent with the semantic cue criticality ranking returned by the extended algorithm, indicating that the extended feedback data are generally consistent with manual judgment.

In order to verify the reliability and stability of the feedback information provided by the extended algorithm, this work studies and calculates the mean and variance of the

Table 1: Jaccard coefficient mean and coefficient variance.

Group number/serial number	Jaccard coefficient mean					Jaccard coefficient variance				
	1	2	3	4	5	1	2	3	4	5
1	0.62	0.65	0.62	0.73	0.89	5.2	6.4	7.1	6.3	21.3
2	0.67	0.61	0.65	0.77	0.84	5.5	5.2	7.2	6.1	21.4
3	0.63	0.64	0.72	0.76	0.88	5.7	4.7	7.6	6.5	21.2
4	0.65	0.67	0.85	0.82	0.86	5.8	4.5	7.5	6.9	21.7
5	0.66	0.68	0.81	0.74	0.87	5.9	4.9	7.8	6.1	21.8

Jaccard coefficients of the comparison path and the corresponding feedback path. Table 1 provides the calculated mean and calculated variance of top 5. The data show that the minimum average Jaccard coefficient between the algorithm output and the comparison data is 0.61, and the variance is basically stable at a small level.

5. Conclusion

This work studies and analyzes the existing implantation technology of Android smart phone Trojans and analyzes the basic principles and implementation methods of obtaining root permissions on mobile phones. At the same time, it also studied the hiding technology of mobile phone Trojans, analyzed the basic principles and implementation methods of file hiding, process hiding, network connection hiding, and kernel module hiding, and analyzed the communication hiding technology and the basic method of cooperative hiding. Through the research of the broadcast receiver model of the Android platform, the background monitoring principle and implementation technology of the mobile phone Trojan horse are analyzed. Aiming at the problem of insufficient training corpus in the event relationship classification task, a tritraining-based event relationship classification method is proposed. This method trains three different classifiers through the labeled corpus, expands the training set by selecting samples with high classification confidence from the unlabeled set by majority voting, and then uses the new training set to retrain. In the process of conceptual network construction and semantic relevance calculation, "semantic strength" is introduced as a quantitative evaluation index for the contribution and reliability of each semantic element in the semantic information comparison, and corresponding measurement rules are designed. According to the semantic strength of the semantic element, the attention degree of the element obtained in the semantic calculation process is set, so that the extraction and application of the semantic information are more detailed, thereby improving the calculation accuracy of the algorithm. At present, there are not many research studies on semantic clue mining, and the experimental scheme for testing its performance is not mature. Although the mining of semantic clues is similar to the path exploration on the Internet, the integration of semantic elements, the dependence on human intelligence, and the lack of azimuth and distance factors make the general path inspection scheme difficult to apply to the subject of this article. Therefore, this study proposes a quantitative evaluation mechanism for the reliability of semantic clues. Although the author has carried out exploration and research work on this topic, due to the relatively new topic and lack of relevant research and data, the semantic clue mining algorithm of this topic still has major shortcomings and still needs to be improved.

Data Availability

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

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

The authors acknowledge Project Research and Practice on Teaching Reform of Public Security Based on Mobile Network Security Information Platform supported by special project of 2019, Guangxi Higher Education Undergraduate Teaching Reform Foundation (2019JGB402).

References

[1] J. Espin and S. Perz, "Environmental crimes in extractive activities: explanations for low enforcement effectiveness in the case of illegal gold mining in Madre de Dios, Peru," *The Extractive Industries and Society*, vol. 8, no. 1, pp. 331–339, 2021.

[2] E. Abdelzaher and Á. Tóth, "Defining crime: a multifaceted approach based on lexicographic relevance and distributional semantics," *Argumentum*, vol. 16, pp. 44–63, 2020.

[3] J. Xiong, "A method of mining key accounts from internet pyramid selling data," *Tehnički Vjesnik*, vol. 26, no. 3, pp. 728–735, 2019.

[4] P. Hensinger, "Trojan horse "digital education"–on the road to a conditioning institution set up in a school without teachers," *Current Concerns*, vol. 19, pp. 17–30, 2017.

[5] Y. Wu and J. Zhang, "Building the electronic evidence analysis model based on association rule mining and FP-growth algorithm," *Soft Computing*, vol. 24, no. 11, pp. 7925–7936, 2020.

[6] Y. Li, S. Yang, J. Zheng, Z. Zou, R. Yang, and W. Tan, "Trojan horse DNA nanostructure for personalized theranostics: can it knock on the door of preclinical Practice?" *Langmuir*, vol. 34, no. 49, pp. 15028–15044, 2018.

[7] J. Demers, "Is a trojan horse an empty signifier? the televisual politics of orange is the new black," *Canadian Review of American Studies*, vol. 47, no. 3, pp. 403–422, 2017.

Hindawi
Journal of Robotics
Volume 2022, Article ID 4167170, 9 pages
https://doi.org/10.1155/2022/4167170

Research Article

A Hybrid Mobile Robot Path Planning Scheme Based on Modified Gray Wolf Optimization and Situation Assessment

Yilun Liu and Xiaoming Li

School of Electrical and Information, Northeast Agricultural University, Harbin 150030, China

Correspondence should be addressed to Xiaoming Li; lixiaoming@neau.edu.cn

Received 7 December 2021; Revised 10 January 2022; Accepted 17 January 2022; Published 16 February 2022

Academic Editor: Shan Zhong

To better solve the problems associated with optimal pathfinding and dynamic obstacle avoidance in the path planning of mobile robots, a hybrid path planning scheme combining modified gray wolf optimization (MGWO) and situation assessment mechanism is proposed. Firstly, a MGWO algorithm is proposed to plan a global path. Secondly, different situational factors for robots in different regions are extracted from the fusion results of 2D laser measurements and image data, and a Bayesian network model of robot action selection is established. Then, the situational factors of the robot are used as evidence for reasoning. Based on the posterior probability value in the inference result, the grid to be moved is selected and the traveling direction of the robot is adjusted in order to take advantage of both global path planning and local dynamic obstacle avoidance. The simulation results show that the proposed MGWO has better optimization performance. When combined with a situation assessment mechanism, it realizes dynamic obstacle avoidance while keeping the path length as short as possible.

1. Introduction

With technological advancement and social development, the level of intelligence and automation of mobile robots has gradually improved, and it has gradually penetrated into people's daily life [1]. In the field of mobile robots, path planning is one of the most important requirements, which is the key technology to realize autonomous robot navigation [2].

The mobile robot needs to plan a short, energy-efficient, and safe path from the initial position to the target position, and it must be able to avoid static and dynamic obstacles along the way. At the same time, mobile robots should have certain computing capabilities of calculating the shortest and safest path in real time to reduce time and energy consumption [3].

This article proposes a mobile robot path planning scheme based on the MGWO and situation assessment mechanism. Firstly, a global path is planned with the proposed MGWO. Secondly, a situation assessment mechanism is used to fuse the data from multiple sensors equipped on the robots, obtaining situational factors from multisensor information and inferring the impacts of the detected obstacles on the robot movement so as to determine the next move. Moreover, the corresponding action is carried out to find the shortest path or the second shortest path from the starting position to the destination without any collision with the obstacles.

The contributions of this article are listed as follows:

(1) To solve the problem of global path planning of the mobile robot, a modified gray wolf optimization algorithm (MGWO) is proposed in which the population diversity is enhanced by logistic chaotic mapping. An adaptive adjustment strategy of control parameters is utilized to achieve a balance between search and development capabilities. With the static weighting average strategy, the population position is updated to speed up the convergence speed, thereby boosting the performance in finding an optimal global path.

(2) From the perspective of cognition, the mechanism of a situation assessment is introduced to the

application of mobile robots, which provides a new solution for the local path planning of mobile robots.

(3) A hybrid planning scheme based on the divide-and conquer concept is proposed. The initial path based on global environment information is obtained with the MGWO algorithm. When an obstacle is detected, the situation assessment mechanism is used to avoid the obstacle through local planning and then back to the global path. Repeat the process until the destination is reached.

The rest of this article is organized as follows. Section 2 introduces the related path planning methods. The details of the proposed MGWO are explained in Section 3. Section 4 describes the proposed situation assessment mechanism. In Section 5, a comprehensive comparative analysis is conducted based on the results obtained from a number of experiments. Finally, Section 6 concludes the full article.

2. Related Works

According to the amount of situational information acquired, robot path planning methods can be divided into two categories: global path planning and local path planning, where local path planning is also called dynamic path planning [4].

At present, global path planning algorithms can be generally classified as conventional algorithms and intelligent algorithms. The former includes A^* algorithm [5] and RRT algorithm [6], and the latter includes ant colony algorithm [7] and particle swarm algorithm [8]. Traditional algorithms mostly use the length of the path as the only indicator, while intelligent search algorithms use randomly generated initial solutions or sampling points to approximate the optimal solution through multiple iterations. The biggest characteristic of an intelligent algorithm is its randomness, so its solution is not unique. A lot of heuristic intelligent search methods have been proposed with respect to optimization algorithms, considering that there are a large number of NP-hard problems in path planning, and each problem has its own optimal solution, which may not be solved. Therefore, the heuristic intelligent search algorithm approximates the optimal path through iterative improvements of randomly generated initial feasible solutions [9].

Since its proposal, the gray wolf optimization (GWO) [10] algorithm has been widely studied because of its advantages of good searching performance, simple structure, and easy implementation. For example, Toufan and Niknafs [11] exploited the cosine function-based searching factors and introduced a dynamic weighting strategy to balance global and local exploring abilities, which improves the solution accuracy of GWO. Olivera et al. [12] used an exponential function to attenuate the searching factor α so that the convergence factor is changed nonlinearly and dynamically with the increase of iteration times and tried to find a better trade-off between the search and development stages to ensure that the optimal solution is approached. Albina and Lee [13] proposed balancing the searching ability with stochastic convergence factors and performing individual update through differential mutation to improve the convergence and accuracy of the original GWO.

The proposed MGWO further improves the performance of the original GWO and speeds up the convergence. However, as a global path planning method, the MGWO does not have the ability to avoid obstacles in real time. Aiming at this problem, we propose a dynamic path planning of mobile robots based on the fusion of MGWO and a situation assessment mechanism.

Local path planning is a kind of dynamic planning. The robot perceives its surroundings with sensor information and plans a collision-free path from the current node to a target subnode online. In a rapidly changing environment, it is necessary to continuously collect sensor information and replan the path to guide the robot's movement in real time. The planning range is usually limited to the detection range of the sensors. Commonly used local planning algorithms include artificial potential field [14], dynamic window [15], D^* algorithm [16], and fuzzy logic approaches [17].

Situation assessment is the high-level information fusion in a multisensor information system. It is the extraction, combination, and comprehensive processing of the high-level relationships of the information flows of different types and multiple levels obtained from the system, thereby acquiring the important cognitive information from sensor information to infer the target intentions and future trends. Situation assessment is a cognitive reasoning method based on human thinking.

Huang et al. [18] proposed a cyber situational awareness (CSA) decision-making system based on fuzzy sets, in which a team CSA with the ability to process unknown information is established in a distributed way. Arora et al. [19] proposed to code certain knowledge with a Bayesian network and perform network reasoning by Monte Carlo tree search technique to plan actions based on information perception. The experiment results prove the method has practical value. Zhou et al. [20] proposed an uncertain information fusion method for aerial situation awareness in the DST framework. Firstly, the perceived uncertain information is preprocessed with belief entropy, and the evaluated information is combined according to Dempster combination rules. This kind of technique can enhance and expand the limited cognitive ability.

3. Global Path Planning with MGWO

3.1. Problem Formulation. Given the regional environment map. $\mathbf{T} = \{\mathbf{T}_1, \mathbf{T}_2, \ldots, \mathbf{T}_n\}$ is the collection of target points. O is the starting point of the robot. The robot is required to traverse n checkpoints and return to the starting point. The traversal path is $\mathbf{R} = \{\mathbf{O}, \mathbf{T}_1, \mathbf{T}_2, \ldots, \mathbf{T}_n, \mathbf{O}\}$. Then, the path length L can be calculated as follows:

$$L = \|\mathbf{OT}_1\| + \sum_{j=1}^{n-1} \|\mathbf{T}_j\mathbf{T}_{j+1}\| + \|\mathbf{T}_n\mathbf{O}\|, \tag{1}$$

where $\|\mathbf{OT}_1\|$ denotes the distance traveled by the robot from the starting point to the first checkpoint. $\|\mathbf{T}_j\mathbf{T}_{j+1}\|$ is the distance from checkpoint j to checkpoint $j+1$. $\|\mathbf{T}_n\mathbf{O}\|$ is the

distance from checkpoint n to the starting point. It is required to find the order of the traversal checkpoints $\{\mathbf{T}_1, \mathbf{T}_2, \ldots, \mathbf{T}_n\}$ so that the shortest path is found with length min$L$.

3.2. Classic Gray Wolf Optimization. The gray wolf optimization is an algorithm that simulates the predation behaviors of predator wolves at the top of the food chain. Most gray wolves like to live in groups and have a very strict social hierarchy, as shown in the pyramid structure of Figure 1.

In the mathematical model of GWO, each wolf represents a candidate solution in the group, where α is the optimal solution and β and δ are the second and third best solution, respectively; the rest of candidate solutions are denoted by ω. In GWO, the search (optimization) is guided by α, β, and δ and followed by ω.

Let the population size of the gray wolves be N, and the search space is a Dim-dimensional space, $\mathbf{X}_i^{\mathrm{Dim}}$ is the position vector of the ith wolf in the Dim-th dimension; then, the mathematical model of the gray wolves hunting prey is expressed as follows:

$$\begin{aligned} \mathbf{D} &= \left|\mathbf{C}\mathbf{X}_p(t) - \mathbf{X}(t)\right|, \\ \mathbf{X}(t+1) &= \mathbf{X}_p(t) - \mathbf{A}\mathbf{D}, \end{aligned} \tag{2}$$

where t is the current iteration number. $\mathbf{A}$ and $\mathbf{C}$ are coefficient vectors, $\mathbf{A} = 2\mathbf{a}\mathbf{r}_1 - \mathbf{a}$, $\mathbf{C} = 2\mathbf{r}_2$, where $\mathbf{r}_1$ and $\mathbf{r}_2$ are random numbers in the range [0, 1]. a is the convergence factor, and its value linearly decreases from 2 to 0 as the number of iterations increases. $\mathbf{X}_p$ is the position vector of the prey. $\mathbf{D}$ is the positional relationship vector between the gray wolf individual and the prey. $\mathbf{X}$ is the position vector of the gray wolf.

The mathematical model of the gray wolves tracking prey locations can be expressed as follows:

$$\begin{cases} \mathbf{D}_\alpha = \left|\mathbf{C}_2\mathbf{X}_\alpha - \mathbf{X}\right|, \\ \mathbf{D}_\beta = \left|\mathbf{C}_2\mathbf{X}_\beta - \mathbf{X}\right|, \\ \mathbf{D}_\delta = \left|\mathbf{C}_2\mathbf{X}_\delta - \mathbf{X}\right|, \end{cases} \begin{cases} \mathbf{X}_1 = \left|\mathbf{X}_a - \mathbf{A}_1\mathbf{D}_\alpha\right|, \\ \mathbf{X}_2 = \left|\mathbf{X}_\beta - \mathbf{A}_2\mathbf{D}_\beta\right|, \\ \mathbf{X}_3 = \left|\mathbf{X}_\delta - \mathbf{A}_3\mathbf{D}_\delta\right|, \end{cases} \tag{3}$$

where $\mathbf{C}_1$, $\mathbf{C}_2$, and $\mathbf{C}_3$ are random vectors. $\mathbf{D}_\alpha$, $\mathbf{D}_\beta$, and $\mathbf{D}_\delta$ denote the distance of wolves α, β, and δ with other members in the wolf pack, respectively. $\mathbf{X}_\alpha$, $\mathbf{X}_\beta$, and $\mathbf{X}_\delta$ denote the positions of wolves α, β, and δ, respectively. $\mathbf{X}$ is the current position of the wolf pack. The position of wolf ω in the wolf pack is jointly determined by wolves α, β, and δ.

3.3. Modified Gray Wolf Optimization Algorithm

3.3.1. Population Initialization Based on Logistic Mapping. Chaos has the characteristics of randomness, ergodicity, and regularity. By introducing chaotic sequences, the individuals in the initial population can utilize the information in the solution space as much as possible, thereby enhancing global searching capability. In the population initialization stage, chaotic sequences are generated using logistic mapping [21], which is simple and chaotic and has good traverse uniformity to generate chaotic sequences to initialize the population position of the wolf pack. The logistic mapping can be expressed as follows:

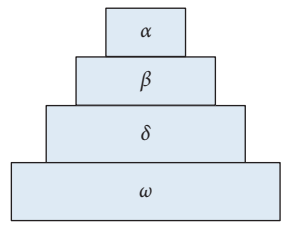


Figure 1: Social hierarchy of gray wolves.

$$x_{n+1} = a x_n (1 - x_n), \tag{4}$$

where $x_n \in [0, 1]$ is the chaotic variable, $a = 4$. Let $x_0 = 0.4288$. Figure 2 shows the bubble map after the population initialization of 30 populations using logistic mapping.

3.3.2. Adaptive Adjustment Strategy of the Control Parameters. The global searching ability and local development ability of the GWO algorithm can be controlled by a. It decays from 2 to 0 linearly as the number of iterations increases, and the iterative convergence process is not linear. Based on this fact, we propose an adaptive adjustment strategy for nonlinear control parameters, which balances the search and development capabilities of the GWO algorithm by adjusting the value of a adaptively. In early iterations, when the global search is performed, a larger value is assigned to a to realize a fast nonlinear changing rate and powerful global search ability and avoid local optimum. In later iterations, a smaller value is assigned to a to slow down the nonlinear changing rate in order to find an optimal solution within a certain region, thereby achieving strong development ability and boosting convergence speed.

The strategy can be expressed as follows:

$$a(t) = a_{\mathrm{initial}} - a_{\mathrm{initial}} \left[\frac{1}{e-1}\left(e^{\left(t/t_{\max}\right)} - 1\right)\right]^n, \tag{5}$$

where t is the current iteration times. $T_{\max}$ is the maximum number of iterations. $n \in [1, 2]$ is the linear modulation index. a_{initial} is the initial value of the control parameter a, $a_{\mathrm{initial}} = 2$. Figure 3 shows the nonlinear decaying trend of a.

3.3.3. Static Weighting Average Strategy. The main idea of the static weighting average strategy is to weigh three leader wolves in accordance with the hierarchical structure of the pyramid. Using static weighting average strategy, wolves α, β, and δ are given the weights of 0.5, 0.3, and 0.2, respectively:

$$\mathbf{X}(t+1) = \frac{5\mathbf{X}_1 + 3\mathbf{X}_2 + 2\mathbf{X}_3}{10}. \tag{6}$$

4. Situation Assessment

In the proposed situation assessment mechanism, the situational factors of the environment where the robot is

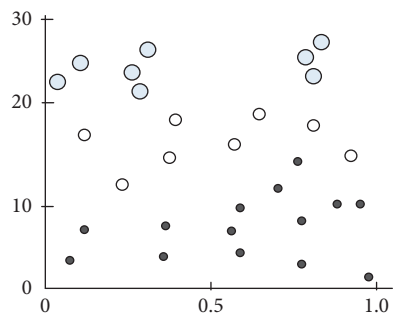


Figure 2: Population initialization with logistic mapping.

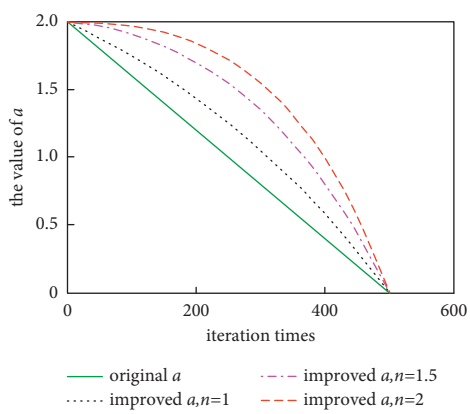


Figure 3: Comparison of the control parameter a.

located are extracted based on the 2D laser ranging data, sonar sensor measurements, and 2D image data of the mobile robot in order to infer the next move of the robot, thereby realizing local path planning within the unknown environment.

4.1. Situational Factor Extraction. Through the fusion of 2D laser ranging data and monocular camera image data, the target information located at the forefront of the mobile robot, that is, within the range of [1°, 17°], is obtained.

Figure 4 shows the coordinate system of the mobile robot, where the positive and negative directions of axis Xt represent the front and back directions of the robot movement, respectively. The positive and negative directions of axis Yt represent the right and left directions of the robot movement, respectively.

Based on the distance D between the target and the robot and the angle A (°) between the target and the positive direction of axis Yt of the robot coordinate system, the environmental factors are extracted using 2D laser ranging data. Place the angle A of the location information into a certain region in the map; if the distance D satisfies $0 \leq D \leq 220$, then the target is regarded as a situational factor of the region.

4.2. Bayesian Network Model of Robot Action Selection. In this study, the local path planning of a mobile robot is classified into three types of relatively independent action

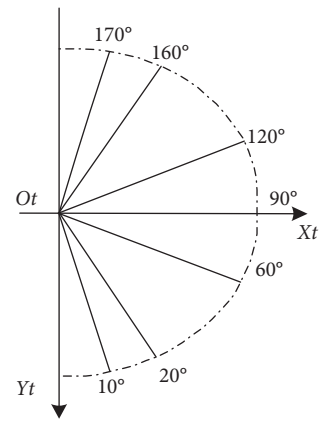


Figure 4: Situational factor extractions based on angle partition.

units (subactions): moving forward, obstacle avoidance, and escaping from a U-shaped trap.

Taking the robot coordinate system in Figure 4 as a reference, the discrimination rules of the three types of subactions are determined. For example, if there are obstacles in the right-ahead, right-front, and left-front directions of a mobile robot, then the robot will choose the subaction of obstacle avoidance.

The forebode node and the result node are extracted from the antecedent part and subsequent part of the discriminant rules, respectively, as listed in Table 1.

In this study, the obstacles in the right and left directions will be regarded as parallel to the moving direction of the robot. Therefore, the forebode node T_2 in Table 1 is determined based on the obstacles with respect to the robot position, that is the right or left obstacles. By using the diagnostic Bayesian network modeling, that is, take the forebode node as the father node of the result node and take the result node as the child node of the forebode node, the Bayesian network model of action selections is established as shown in Figure 5.

In Figure 5, the action selection node C comprehensively considers the values of the forward node, avoidance node, and escaping node. Except for node C, all nodes in Figure 5 are of the discrete node type, which are binary discrete nodes, and their value states are yes (Y) and no (N). Node C is a four-value discrete node, and its value states are S_0, S_1, S_2, and S_3. Among them, S_0 means choosing to move forward, S_1 means choosing to avoid obstacles, S_2 means choosing to escape from U-shaped traps, and S_3 means an error state. After determining the Bayesian network structure, the conditional probabilities of each node need to be given. The parameter is configured based on the discriminate rules of subactions. Take the avoidance node (S_2) as an example; the configuration is listed in Table 2.

All forebode nodes in Figure 5 are the root nodes of the Bayesian network. In this study, the value of all forebode nodes is set to (0.5, 0.5).

4.3. Action Selection Inference. For the action selection inference of the mobile robot, based on the Bayesian network model shown in Figure 5, inputting the evidence and

Table 1: Robot action selection based on Bayesian network model.

Node description	Node type	Explanation
Obstacle judging node (T_1)	Forebode node	Whether or not there is an obstacle in front of the robot within the range [1°, 17°]
Obstacle node in parallel with the robot (T_2)	Forebode node	Whether or not the obstacle detected is parallel to the moving direction of the robot
Right-front obstacle (T_3)	Forebode node	Whether or not there is an obstacle on the right front of the robot
Left-front obstacle (T_4)	Forebode node	Whether or not there is an obstacle on the left front of the robot
Right-ahead obstacle (T_5)	Forebode node	Whether or not there is an obstacle right ahead of the robot
Right obstacle (T_6)	Forebode node	Whether or not there is an obstacle on the right of the robot
Left obstacle (T_7)	Forebode node	Whether or not there is an obstacle on the left of the robot
Forward node (S_1)	Result node	If the robot chooses to move forward
Avoidance node (S_2)	Result node	If the robot chooses to avoid obstacle
Escaping node (S_3)	Result node	If the robot chooses to escape from a U-shaped trap
Action selection node (S_4)	Result node	The final action decision made by the robot

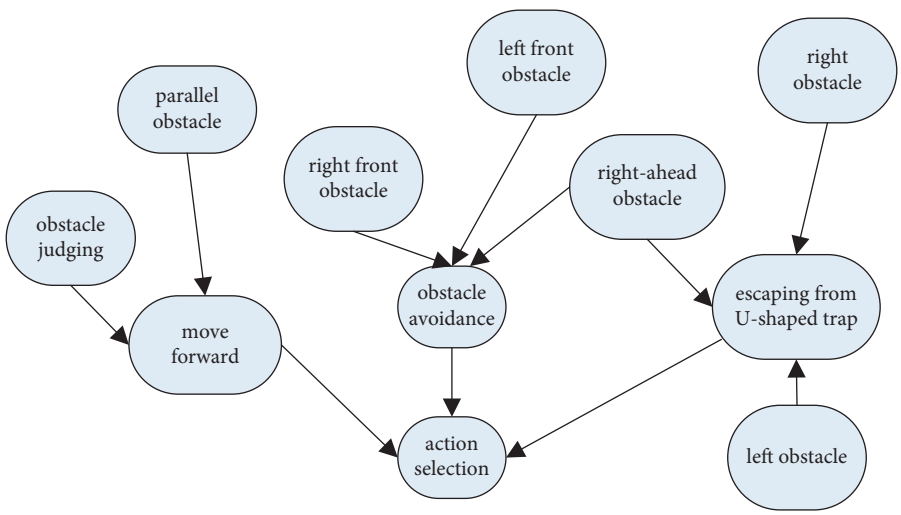


Figure 5: Bayesian network model.

Table 2: Parameter configuration for S_2 node.

T_3	T_4	T_5	S_2 Y	S_2 N
Y	Y	Y	1.0	0.0
Y	Y	N	1.0	0.0
Y	N	Y	1.0	0.0
Y	N	N	1.0	0.0
N	Y	Y	1.0	0.0
N	Y	N	1.0	0.0
N	N	Y	1.0	0.0
N	N	N	0.0	1.0

calculating the posterior probabilities of the selection nodes by the reasoning algorithm of the Bayesian network, the result node with the biggest posterior probability will be chosen as the action decision of the robot.

In the Bayesian network of action selection, the forebode nodes are evidence nodes, and evidence values need to be assigned to the corresponding forebode nodes before performing the inference.

The evidence values of the forebode nodes are obtained from the extracted situational factors of the robot. For example, if the set of situational factors in the right-ahead direction is not empty, then the evidence value of the obstacle node T_5 in the right-ahead direction is set as Y; otherwise, it is set to N. Based on the extracted situational factors, input evidences for seven forebode nodes (T_1, T_2, T_3, T_4, T_5, T_6, T_7), where the evidence values are denoted by $e1$, $e2$, ..., $e7$, form an evidence set $E = \{e_1, e_2, \ldots, e_7\}$. Taking the evidence set E as posterior conditions, the posterior probabilities of the action selection node C can be calculated with Bayesian network reasoning algorithm; then, the posterior probabilities of node C in 4 value states can be

Table 3: Action selection details.

	Subaction	Deviation angle(/°C)
1	Move forward	0
2	Escaping from U-shaped obstacle	180
3	Obstacle avoidance	30 to the right
		60 to the right
		30 to the left
		60 to the left
		30 to the left or right
		60 to the left or right
		0

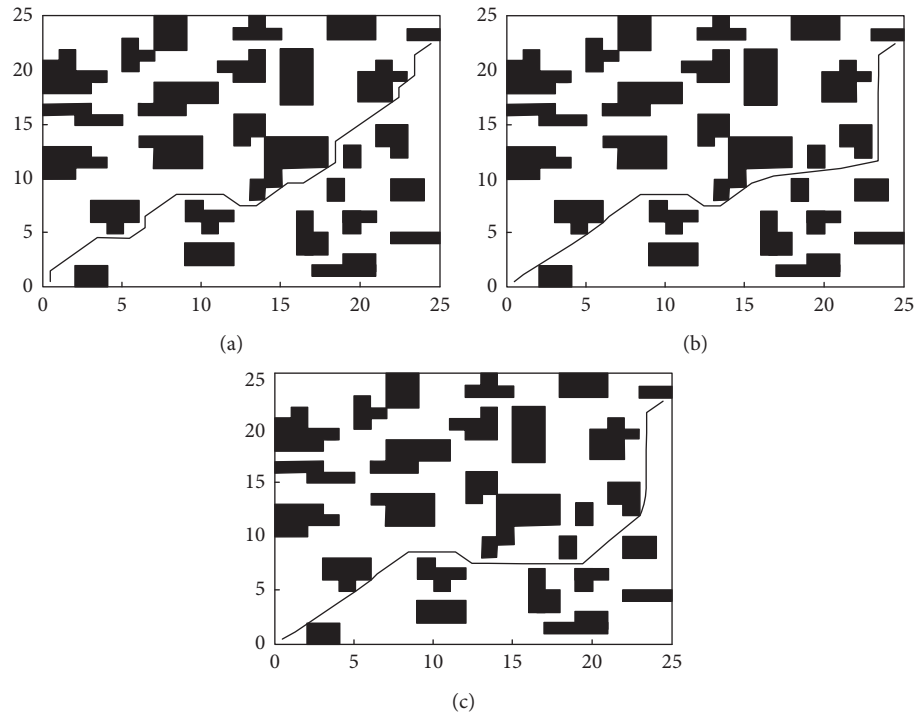


Figure 6: Path planning performance of different methods. (a) Path planned by the PSO algorithm. (b) Path planned by the GWO algorithm. (c) Path planned by the MGWO algorithm.

obtained as follows: $P(C = S_0/E)$, $P(C = S_1/E)$, $P(C = S_2/E)$, $P(C = S_3/E)$.

The Bayesian network reasoning algorithms include precise reasoning algorithm and approximate reasoning algorithm. Due to the low complexity of the Bayesian network model of action selection, this article adopts the precise reasoning algorithm of the Bayesian network, which is the variable elimination method.

Certain movement decisions will be taken in corresponding to the subaction chosen by the mobile robot for the next step. The decisions for the three types of subactions are listed in Table 3.

4.4. Hybrid Path Planning. The fundamental idea behind the hybrid path planning mechanism is based on the divide-and-conquer concept. The mechanism is divided into two scenarios. Under the condition that the global situation information is known, a global path is planned with the global path planning algorithm. On this basis, the situation assessment method is used to plan a local path to avoid obstacles.

Firstly, an initial path based on global situation information is planned with the proposed MGWO algorithm. Local path planning is carried out under the guidance of the global path. When there are no obstacles that randomly appear in the surroundings, the situation assessment method plans a trajectory along the global path. If random obstacles are detected nearby that block the robot from continuing to move along the global path, the situation assessment method focuses on obstacle avoidance, planning a local path to avoid the obstacles. After that, the robot returns back to the global path until it reaches the destination.

5. Experiment

5.1. Global Path Planning Test. Firstly, the grid map is used to model the environment information, and the global path

Table 4: Quantification results.

Methods	Longest path	Shortest path	Average length	Variance
PSO	46.77	39.55	42.67	3.11
GWO	45.98	38.57	39.88	4.02
MGWO	39.74	36.68	38.33	2.12

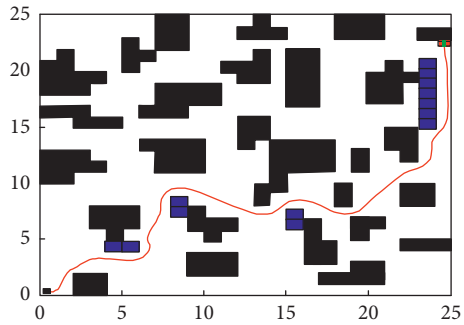


Figure 7: Path planned by the hybrid method after static obstacles added to the map.

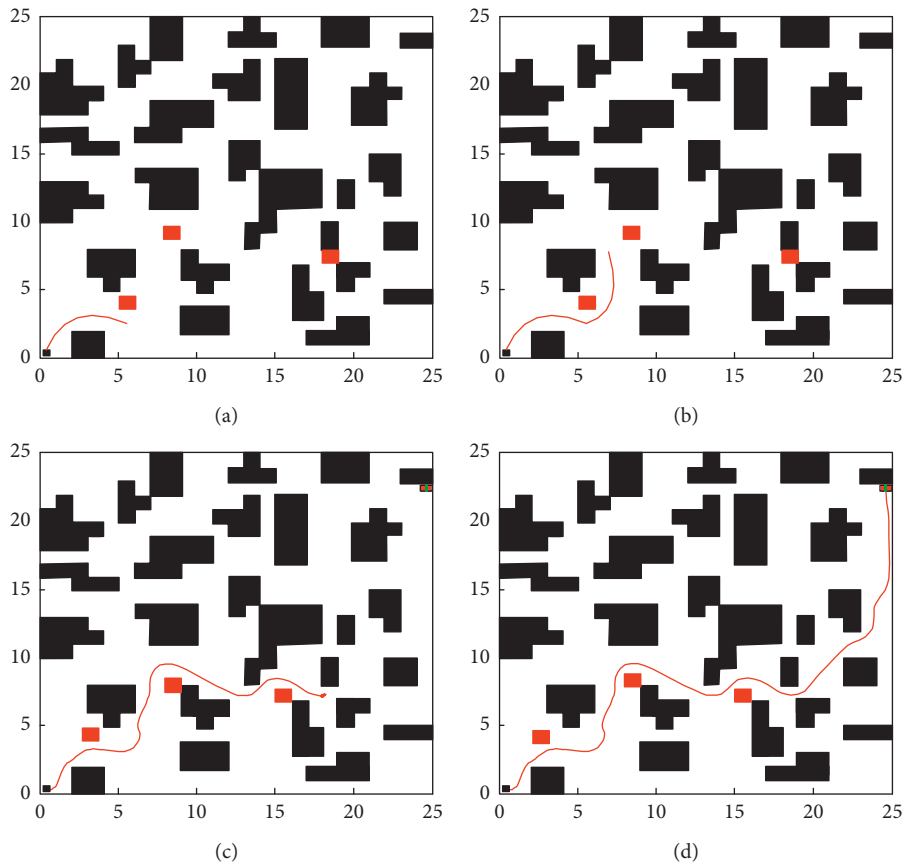


Figure 8: Path planned by the hybrid method after dynamic obstacles added to the map. (a) Hybrid path planning *a*. (b) Hybrid path planning *b*. (c) Hybrid path planning *c*. (d) Hybrid path planning *d*.

planning is carried out using the classic GWO, the proposed MGWO, and the particle swarm algorithm (PSO), respectively. In the simulation, for all the algorithms, the population size is set to 50, and the maximum number of iterations is set to 100.

5.1.1. The Shortest Collision-Free Path. Figure 6 shows the path planning performance of the three algorithms in the grid environment. The quantified results of each of the three algorithms after 20 iterations are shown in Table 4. It can be seen from the table that the proposed MGWO outperforms

the other two algorithms in terms of the longest path, the shortest path, and the average path length.

Compared with PSO and classic GWO, the longest path planned by the proposed MGWO is 4.25 m and 4 m shorter, respectively, which is a reduction of 9.35% and 8.85%, respectively. In terms of the shortest path length, the proposed MGWO has achieved a 3.08% and 1.58% reduction compared with PSO and classic GWO, respectively. Moreover, the average path length is also reduced by 7.13% and 2.88% by the proposed MGWO compared to PSO and classic GWO, respectively. In addition, the variance of the path length found by the proposed MGWO is more stable than the other algorithms.

5.2. Hybrid Path Planning Simulation. Static obstacles are added to the grid map to test the static obstacle avoidance capability of the proposed hybrid method; the 12 static obstacles added are shown in Figure 7. In the figure, blue grids represent newly added obstacles; red lines indicate the moving path of the robot. Use the situation assessment method to plan the local path for the robot. After successfully avoiding the newly added obstacles, the robot continues to move along the global path until it reaches the endpoint.

Three dynamic obstacles are added to test the dynamic obstacle avoidance capability of the proposed hybrid method. The obstacles move horizontally, and the mobile robot successfully avoids the obstacles with the planned local path, indicating the path planning task is successfully accomplished. The planned path provided by the hybrid method is shown in Figure 8.

From the experimental results, it can be seen that compared with using only local path planning, the path planned by the proposed hybrid method is able to address the local minima problem and provide a path very close to the optimal path length. In addition, the simulation results also prove that the proposed hybrid method has dynamic obstacle avoidance capability; that is, the hybrid method can evade newly added static and dynamic obstacles and successfully reach the endpoint while guaranteeing an optimal path.

6. Conclusion

This article proposes a mobile robot path planning scheme based on a modified gray wolf optimization algorithm and situation assessment mechanism. Firstly, we propose the MGWO algorithm based on the GWO algorithm, which has the characteristics of good searching ability and robust performance. In MGWO, the population diversity is enhanced by logistic chaotic mapping, the balance between search and development capabilities of the GWO algorithm is achieved by the adaptive adjusting strategy of the control parameters, and a static weighting average strategy is used to update the population position and speed up convergence. Finally, combine the paths planned globally and locally to ensure that the situation assessment method can perform local path planning along the globally planned path.

Simulation results prove that the proposed scheme has higher optimization accuracy and stability than state-of-the-art methods. The proposed MGWO algorithm can effectively reduce the path length and make the trajectory smoother. The hybrid method can ensure the optimal path to the greatest extent while taking into account the dynamic obstacle avoidance ability of the robot, which has more advantages than a single algorithm.

In the future, based on the scale of the tasks and battery capacity constraint of the robot, pathfinding can be performed based on task clustering, and large-scale tasks can be automatically and autonomously converted into time-by-step tasks so as to improve the applicability of path planning schemes in large-scale scenes.

Data Availability

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

Conflicts of Interest

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

References

[1] F. Rubio, F. Valero, and C. Llopis-Albert, "A review of mobile robots: concepts, methods, theoretical framework, and applications," *International Journal of Advanced Robotic Systems*, vol. 16, no. 2, Article ID 1729881419839596, 2019.

[2] Z. Mohd Nayab and J. C. Mohanta, "Methodology for path planning and optimization of mobile robots: a review," *Procedia Computer Science*, vol. 133, pp. 141–152, 2018.

[3] J. Cheng, Hu Cheng, M. Q.-H. Meng, and H. Zhang, "Autonomous navigation by mobile robots in human environments: a survey," in *Proceedings of the 2018 IEEE International Conference on Robotics and Biomimetics (ROBIO)*, December 2018.

[4] S. Campbell, N. O'Mahony, A. Carvalho, L. Krpalkova, D. Riordan, and J. Walsh, "Path planning techniques for mobile robots a review," in *Proceedings of the 2020 6th International Conference on Mechatronics and Robotics Engineering (ICMRE)*, February 2020.

[5] B. Fu, L. Chen, Y. Zhou et al., "An improved A^* algorithm for the industrial robot path planning with high success rate and short length," *Robotics and Autonomous Systems*, vol. 106, pp. 26–37, 2018.

[6] H. Ryu and Y. Park, "Improved informed RRT^* using gridmap skeletonization for mobile robot path planning," *International Journal of Precision Engineering and Manufacturing*, vol. 20, no. 11, pp. 2033–2039, 2019.

[7] K. Akka and F. Khaber, "Mobile robot path planning using an improved ant colony optimization," *International Journal of Advanced Robotic Systems*, vol. 15, no. 3, Article ID 1729881418774673, 2018.

[8] E. Krell, A. Sheta, A. Balasubramanian, and S. King, "Collision-free autonomous robot navigation in unknown environments utilizing PSO for path planning," *Journal of Artificial Intelligence and Soft Computing Research*, vol. 9, 2019.

[9] J. Li, S. X. Yang, and Z. Xu, "A survey on robot path planning using bio-inspired algorithms," in *Proceedings of the 2019*

IEEE International Conference on Robotics and Biomimetics (ROBIO), December 2019.

[10] S. Mirjalili, S. M. Mirjalili, and A. Lewis, "Grey wolf optimizer," *Advances in Engineering Software*, vol. 69, pp. 46–61, 2014.

[11] N. Toufan and A. Niknafs, "Robot path planning based on laser range finder and novel objective functions in grey wolf optimizer," *SN Applied Sciences*, vol. 2, no. 8, pp. 1–19, 2020.

[12] J. Oliveira, P. M. Oliveira, J. Boaventura-Cunha, and T. Pinho, "Chaos-based grey wolf optimizer for higher order sliding mode position control of a robotic manipulator," *Nonlinear Dynamics*, vol. 90, no. 2, pp. 1353–1362, 2017.

[13] K. Albina and S. G. Lee, "Hybrid stochastic exploration using grey wolf optimizer and coordinated multi-robot exploration algorithms," *IEEE Access*, vol. 7, pp. 14246–14255, 2019.

[14] U. Orozco-Rosas, O. Montiel, and R. Sepúlveda, "Mobile robot path planning using membrane evolutionary artificial potential field," *Applied Soft Computing*, vol. 77, pp. 236–251, 2019.

[15] Y. Zhang, Z. Liu, and L. Chang, "A new adaptive artificial potential field and rolling window method for mobile robot path planning," in *Proceedings of the 2017 29th Chinese Control And Decision Conference (CCDC)*, May 2017.

[16] A. Stentz, *The D* Algorithm for Real-Time Planning of Optimal Traverses*, Carnegie-Mellon Univ Pittsburgh Pa Robotics Inst, Pittsburgh, Pennsylvania, 1994.

[17] M. M. Almasri, K. M. Elleithy, and A. M. Alajlan, "Development of efficient obstacle avoidance and line following mobile robot with the integration of fuzzy logic system in static and dynamic environments," in *Proceedings of the 2016 IEEE Long Island Systems, Applications and Technology Conference (LISAT)*, April 2016.

[18] Z. Huang, C.-C. Shen, S. Doshi, N. Thomas, and H. Duong, "Fuzzy sets based team decision-making for cyber situation awareness," in *Proceedings of the MILCOM 2016-2016 IEEE Military Communications Conference*, November 2016.

[19] A. Arora, R. Fitch, and S. Sukkarieh, "An approach to autonomous science by modeling geological knowledge in a bayesian framework," in *Proceedings of the 2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, September 2017.

[20] Y. Zhou, Y. Tang, and X. Zhao, "A novel uncertainty management approach for air combat situation assessment based on improved belief entropy," *Entropy*, vol. 21, no. 5, p. 495, 2019.

[21] D. R. da Costa, O. Rene, T. Medrano, and E. Denis Leonel, "Route to chaos and some properties in the boundary crisis of a generalized logistic mapping," *Physica A: Statistical Mechanics and its Applications*, vol. 486, pp. 674–680, 2017.

Hindawi
Journal of Robotics
Volume 2022, Article ID 8286007, 6 pages
https://doi.org/10.1155/2022/8286007

Research Article

Implementation of a Community Data Processing System Based on Data Mining

Li Li

School of Philosophy and Public Administration, Henan University, Kaifeng, Henan, China

Correspondence should be addressed to Li Li; 30040056@vip.henu.edu.cn

Received 31 October 2021; Revised 20 December 2021; Accepted 23 December 2021; Published 15 February 2022

Academic Editor: Shan Zhong

With the development of social information technology, intelligent community as a new way of life is changing people's lives step by step. The community as a unit of residence in China has developed quite maturely, but the community services are not perfect, and the management is relatively mechanized. Therefore, improving work efficiency and enriching and perfecting community service is an increasingly important issue. Data mining is to extract valuable information by analyzing the internal connections, rules, and patterns of these data so as to provide more favorable decision support for community managers and provide users with more humane and modern community intelligence services. This research focuses on the implementation of community data processing systems based on data mining. Firstly, data preprocessing analysis is carried out, the realization of data storage and cache is studied, the process and characteristics of cluster analysis are studied in detail, and the simulation results of a community data processing system based on data mining are summarized. This study uses data mining technology to dig out the daily consumption data of users in the community mall, cluster the data, and analyze the consumption situation and consumption types of different types of users. In this study, data mining technology is used to mine the fault repair data of the community, and classification prediction technology is used to classify and predict different types of faults so as to help managers troubleshoot problems existing in the community environment.

1. Introduction

Intelligent community, in theory, is the intelligent control of community data, the combination of hardware equipment and network information technology to connect the whole community network, and intelligent control through mobile terminal software [1, 2]. With the rapid development of information technology and the rapid development of intelligent devices and information devices, the intelligent community is no longer an empty talk. People's pursuit of intelligent and humanized life in the new society has also become the driving force for the market expansion and development of intelligent communities [3]. Data mining, as its name implies, is actually an analysis process of mining or extracting valuable data or knowledge with unknown laws from intricate and irregular data [4]. With the rapid growth of community databases, a large amount of potentially valuable data is scattered in various database tables. If data mining cannot be used to efficiently integrate and analyze these data, it will not only waste data information but also cannot help community managers make important decisions [5, 6]. With data mining as an effective means, smart home systems with push function as the way, and with the purpose of providing high-quality and humanized community services, to improve the lifestyle of community residents to the greatest extent, provide decision-making assistance for community managers, provide users with a comfortable, convenient, intelligent, and colorful life, and embrace the arrival of the information age. Based on the existing intelligent community system, this paper makes an in-depth analysis of data mining architecture and constructs a data mining architecture suitable for the community. The clustering analysis method and classification prediction method of data mining technology are applied to the consumption data and historical repair data of community users. According to the data integration and analysis results,

the decision-making basis for community service managers is provided. In addition, in the process of promoting the modernization of social governance, there are information barriers between government functional departments and resources that cannot be reasonably integrated, which leads to many community management services not being effectively and quickly supplied. The construction of a community management service platform can change this situation to some extent and promote the modernization of social governance. Through the constructed community management service platform, to obtain a large number of community and community residents data through data mining, to help the government make decisions on this basis and improve the intelligent level and professional level of social governance.

2. Data Preprocessing

For numerical data, data mining analysis methods have inherent advantages, and they have high computational efficiency and fast execution speed. Therefore, this study tries its best to convert data into numerical data for operations, which are mainly completed in data preprocessing [7]. Data preprocessing can be divided into two categories: the first is to shield and omit some irrelevant data in the clustering algorithm; the second is to process some nonnumerical data and convert them into numerical data.

In the first type of data preprocessing, since the basic information used to identify users, such as user number and name, is meaningless to analyze the identification information of such character types in the clustering algorithm, the records are identified only when they are stored and screened during algorithm analysis. For example, the data selected by a certain cluster include user number, age, communication consumption, entertainment consumption, and domestic consumption, which is a group of five dimensions of data space. After the omission, the data set becomes the four-dimensional data of age, communication consumption, entertainment consumption, and domestic consumption, and the data of all dimensions are numerical, which meets the requirements of cluster analysis.

In the second type of data preprocessing, it is mainly aimed at the digital standardization of nonnumerical characteristic data, such as residence location, gender, and so on, and converts them into feature vectors, which are used to identify the corresponding points of the user's feature space, so as to realize the clustering algorithm analysis of the feature space. For example, 0 indicates female and 1 indicates male. Different positive integers represent different communities. The address number is composed of the building number, unit number, and room number. The address number is the floating point number added on the basis of the corresponding integer of the cell, which can not only ensure the numerical value of the data but also carries out correlation analysis for the user's living location.

3. Data Storage and Caching

Considering that the data can be reused and to ensure the consistency of various data, ARFF format files are used to cache the collected data sets to improve system performance [8]. Data set features are mainly divided into the following two types:

(1) Static characteristic data: such as name, gender, age, residential address, and other attributes that remain unchanged for a long time

(2) Real-time statistics: such as user consumption data, reports for repair, water and electricity consumption, and so on, need to query the database in real time to obtain statistical data

Data storage solution is shown in Figure 1:

According to the characteristics of these two kinds of data, this paper adopts the storage and query scheme, as shown in Figure 1, to optimize the data. Since static feature data have not changed for a long time, statistics can be queried and updated from the community management system database at intervals (monthly or even quarterly). While real-time data can be optimized based on cache file modification time, for example, the cache file stores all the consumption data of a user before December 31, 2019. For the subsequent data mining, it only needs to query the user consumption data from January 1, 2020 to the query time in the community management system database, and then add them to the cache data.

4. Clustering Analysis

In the front page, users input the initial K value of cluster analysis, the start and end time of cluster data, and select the data source to be clustered. The data of community mall can be carried out separately, or the community mall can be combined with the basic information of community users for cluster analysis. The value of k is verified in Java Script to determine whether it is empty, whether it is a numeric parameter, whether it is greater than 0, and so on. After the verification is successful, the data are passed to the action for processing [9]. If there is no corresponding cached data, the corresponding data in the data source will be queried according to the time, and the data will be input into the clustering algorithm for analysis after data pretreatment and caching. The analysis results of the algorithm should also be processed and evaluated to obtain the final results of the cluster analysis.

4.1. Analysis of Clustering Algorithm. The principle of the clustering algorithm is as follows: firstly, scan data one by one, and each data feature is classified into the same class or generated into a new class according to the distance from the scanned data; then the distance between the various basis classes is combined until a certain requirement is reached and stopped. In this paper, the classical K-means algorithm is adopted for clustering, and the cluster analysis process is shown in Figure 2:

(i) During initialization, n total data objects and K objects are given as the initial nodes of clustering.

(ii) Scan other node objects in turn, and calculate the similarity distance between these objects and the initial central node. The calculation formula uses the

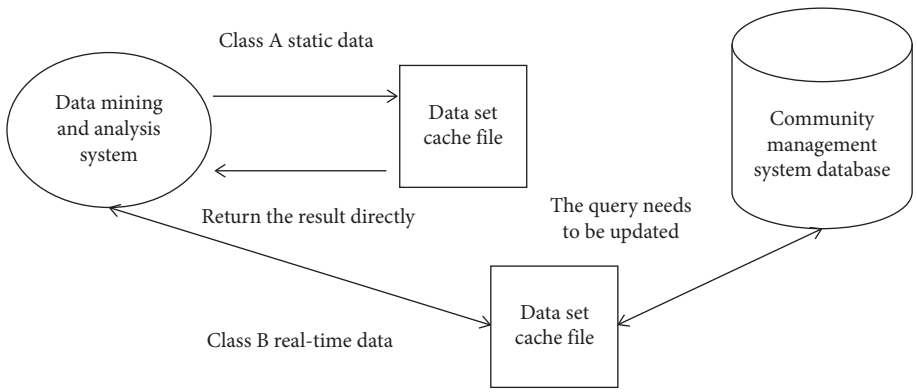


Figure 1: Data storage solution.

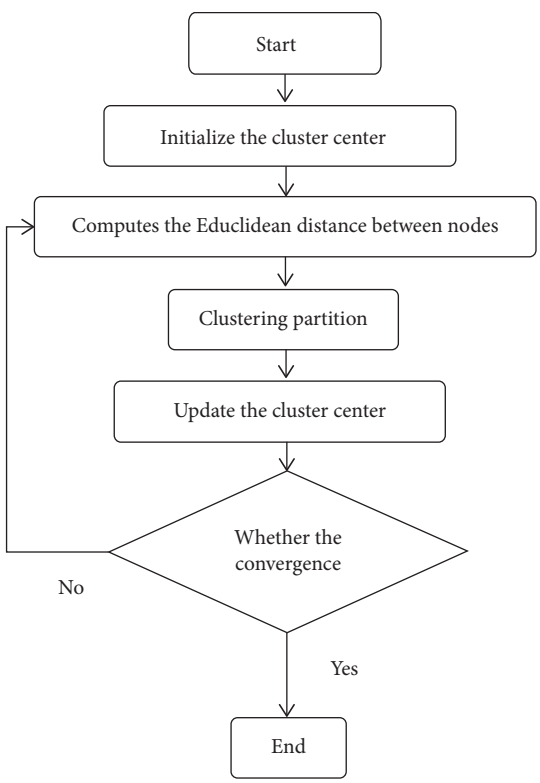


Figure 2: Cluster analysis process.

classical Euclidean distance formula to measure the index of similarity as follows:

$$d(x, y)\sqrt{\sum_{i=1}^{P}\left|x_i - y_i\right|^2}. \quad (1)$$

Among them, the similarity index meets the following mathematical function requirements: (a) $d(x, y)$ represents the distance between two objects and is a nonnegative value; (b) $d(x, x) = 0$, the distance between an object and itself is always equal to 0; (c) $d(x, y) = d(y, x)$, the distance function has symmetry; and (d) $d(x, y)$ cannot be greater than the sum of $d(x, z)$ and $d(z, y)$, because the direct distance from object x to object y cannot be greater than the distance from any third object z in the path (triangle inequality)

(iii) These objects are respectively assigned to the most similar (shortest Euclidean distance) cluster

(iv) Recalculate and update the cluster central node

(v) This process is repeated until the standard measure function (mean square deviation) converges

Consumption table of cluster analysis is shown in Table 1.

The consumption data after preprocessing and storage (as shown in Table 1) can be analyzed according to the abovementioned clustering algorithm. All the user's feature attributes constitute each corresponding point of the user's feature space, and some methods can be adopted to reduce the accuracy and improve the clustering efficiency through distance division of the feature space. Finally, find the feature space of similar user groups and form a list of users with similar features and user group classification. The closer the point is to the user, the higher the similarity to the user. In the face of the huge amount of user data in the community, especially when there are many user feature dimensions, 10,000 users will generate 100 million similarity distances, and the calculations are extremely complicated. At the same time, considering the high latitude data vector, the k-means algorithm is not highly efficient. Therefore, the system will reduce the accuracy of some very similar user attributes in the process of clustering so as to reduce the number of clustering indexes and improve efficiency. For example, users whose age ranges from 24 to 26 years old are clustered in accordance with the age of 25. In addition, users living on the same floor are classified as one, that is, the last two room numbers in the address number are omitted. For example, the user with address number 1.0231101 and the user with address number 1.0231102 are identified as 1.02311.

The k-means algorithm requires the initial value of the number of clusters to be set in advance. The more clusters there are, the more classification results there will be and the more detailed classification conditions there will be. However, the more, the better. In this system, community managers can input the number of clusters in advance on the page, and the background server will use the obtained

Table 1: Consumption table of cluster analysis.

Age	Shopping consumption	Entertainment consumption	Tourism consumption	Household consumption	Education consumption
25	800	1800	1200	400	1000
35	1000	3100	2400	800	200
40	1000	1600	2000	1000	3000
60	800	100	600	2300	0

number of clusters for clustering analysis by the K-means algorithm. By default, the number of clusters is 3.

4.2. The Key Code for Calculating the Euclidean Distance

```
Private float distance (float [] element, float [] center){
Float distance = 0.0f;
Float x = element [0] − center [0];
Float y = element [1] − center [1];
Float z = x*x + y*y;
Distance = (float) Math.sqrt (z);
Return distance;
}
```

The key code for calculating the error sum of squares criterion function:

```
Private void count Rule (){
Float jcf = 0;
For (int i = 0; i < cluster.size (); i++){
For (int j = 0; j < cluster.get (i).size (); j++){
jcf+ = error Square (cluster.get (i).get (j), center.get (i));
}}
jc.add (jcf);}
```

Update the key code of the central node of the clustering algorithm:

```
Private void set New Center (){
For (int i = 0; i < k; i++){
Int n = cluster.get (i).size ();
If (n ! = 0){
Float [] new Center = {0, 0};
For (int j = 0; j < n; j++){
New Center [0]+ = cluster.get (i).get (j) [0];
New Center [1]+ = cluster.get (i).get (j) [1];
}
New Center [0] = new Center [0]/n;
New Center [1] = new Center [1]/n;
Center.set (i, new Center);}}}
```

4.3. Cluster Analysis. The system uses multiple methods to display clustering results at multiple levels and strives to provide community managers with clear and concise data results and convenient operation interfaces. The system sets up three forms to display the cluster analysis results: a pie chart, a cluster center distribution table, and a user feature list.

(i) Pie chart: intuitive, can clearly and concisely show the proportion of each category

(ii) Cluster center distribution table: it enables community managers to understand the distribution of each cluster center and understand the characteristics of cluster user groups

(iii) User feature table: it can fully display the user group corresponding to each clustering result and its detailed information

In addition, each category should have a push service based on the information push system. The purpose of cluster analysis in this study is to cluster irregular user consumption data and user basic information data. And different types of information are pushed to different consumer groups to achieve humanization and intelligent management of the community.

Example analysis of clustering results is shown in Table 2.

According to Table 2, for young users with relatively high entertainment and shopping consumption, the community manager can push promotional activities such as mall discounts for these users based on the actual situation of surrounding commerce, advertising, and investment promotion in the community. For middle-aged and elderly users who consume less entertainment and more household services, community managers can push advertising information such as medical treatment, social security, and home promotion to them.

5. Analog Implementation

In this test, the consumption data of community users in a quarter from October to December 2019 provided by a company were used for testing. The test results are shown in Table 3.

According to Table 3, the test results show that among them, the share of users in the first category is up to 45%. This category of users has the highest entertainment consumption and belongs to the entertainment consumption type users, whose average entertainment consumption is about 2500 yuan. Therefore, for these kinds of users, they can push peripheral business, advertising, and entertainment information. The second category accounts for 23% of users, who spend the most on study and tourism, with an average of 1700 yuan on education and 1300 yuan on tourism. These users belong to educational tourism users, so they can push educational information around the community, such as English learning, primary and secondary school training, music, dance, and other educational information, as well as surrounding tourism

Table 2: Example analysis of clustering results.

Cluster type	Generation	Consumption type	Push ads or information
First	Youth	High entertainment consumption, low household consumption	Entertainment information
Second	Middle aged	Entertainment consumption medium, household consumption medium	Entertainment and household information
Third	Elderly	Low entertainment consumption, high household consumption	Household information

Table 3: The test results.

Clustering categories	Proportion (%)	Age	Entertainment	Tourism	Shopping	Education	Living	Propensity to consume
First	45		2500	1100	200	500	400	Entertainment consumption type
Second	23		400	1300	600	1700	800	Educational tourism type
Third	32		200	900	400	200	2100	Home saving

information for them. The third category accounts for 32% of the total number of users. This category of users has a high level of home consumption and belongs to the frugal type of home users, with an average home consumption of about 2100 yuan. Therefore, they can push medical, social security, home, and other service information around the community to them.

For the classification test of obstacle report data, in this test, the obstacle repair data of a community in 2020 are used as the classification training data. Use the obstacle repair data of a community in 2020 as the real test data, and use the decision tree obtained from the training data to predict and analyze the real data. The test results show that elevator failure, line failure, and sewer failure are key problems. Community managers need to focus on maintenance and maintenance to avoid failure. Lighting faults, access control faults, monitoring faults, and other faults are not critical problems. The system makes a return visit decision and judgment for specific obstacles in the community. The classification decision judgment results of common disorders are shown in Table 4.

In the data reported for repair, only 1.51% of those reported for repair need to be paid attention to and visited again. Obviously, only a small part of the repair records have the value of a return visit. The list contains all the records that need to be returned. The community administrator can view the detailed repair reports and directly push messages to the user.

According to the test results of cluster analysis and classification prediction analysis of mining and analysis systems, the application of big data analysis technology is conducive to aiming at these characteristics of user consumption data. After analysis, it was decided to use the clustering analysis method in data mining to analyze user consumption data. Big data analysis technology is used to classify the data of obstacle repair reports according to certain rules and find out their classification rules, so as to predict the distribution of different types of obstacles in the community and the proportion of obstacle repair reports according to their classification rules. In this way, community managers can not only actively eliminate obstacles but also master the distribution and regional coverage of obstacles in the whole community.

Table 4: The classification decision judgment results of common disorders.

Fault category	Times	Fraction of coverage (%)	Focus or not
Sewer faults	15	7.18	Yes
Line faults	23	2.36	Yes
Elevator faults	34	5.61	Yes
Monitoring faults	18	2.33	No
Access control faults	17	2.48	No
Lighting faults	21	3.19	No
Other faults	9	1.41	No

6. Conclusion

The community data processing system based on data mining includes three major systems: community management, data mining and analysis, and information push. Among them, community management helps community managers to complete daily management works, such as user data maintenance, fault reporting, and water and electricity payment. At the same time, it is also the main source of community user data and provides platform support for data mining and analysis systems and information push systems. Data mining and analysis, on the one hand, it mines the daily consumption data of users in the community mall. By clustering the data, it can analyze the consumption situations and consumption types of different types of users and provide support for the strategic decision-making of the community by combining the basic information of users grasped by the community, so as to provide more targeted and humanized services for users. On the other hand, it mines the fault repair data of the community and classifies and predicts different types of faults by using the classification and prediction technology, so as to help managers troubleshoot problems existing in the community environment and reduce the incidence of faults. Information push service is an important link to combine community service with a smart home system. It uses a XMPP protocol to push community information to users' mobile phone terminals by using a smart home system, so that users can get all kinds of community service information in real time, narrowing the distance between community managers and

users. Therefore, on the basis of data mining technology, we can make full use of the innate advantages of the community to obtain user information, analyze user data, and use the analysis results to provide help for the daily management and strategic decision-making of the community, so as to improve the quality of service and the sense of belonging of the community.

At the same time, due to the lack of personal research ability and the limitation of data collection, there are still many deficiencies in this study, which can be further improved in future research. As for data mining technology, this paper has not studied deeply enough. More effective algorithms can be used to analyze community data. In the management data and user consumption data of the community, there are still a lot of wasted useful information for the system to mine and analyze, such as the analysis of the user's payment data and the investigation of some abnormal users. There is still room for improvement and development. In the process of follow-up work and study, we will continue to study the technology involved in this topic and constantly improve and enrich the research results of this topic.

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] K. He and Z. Yang, "Smart community management system construction," *Bulletin of Surveying and Mapping*, vol. S2, pp. 248–250, 2016.

[2] N. Singh, V. Singh, and T. E. Carlson, "Pim-GraphSCC: PIM-based graph processing using graph's community structures," *IEEE Computer Architecture Letters*, vol. 19, no. 2, pp. 1-2, 2020.

[3] T. M. Roddenberry, M. T. Schaub, H. T. Wai, and S. Segarra, "Exact blind community detection from signals on multiple graphs," *IEEE Transactions on Signal Processing*, vol. 68, no. 99, pp. 1-2, 2020.

[4] A. Mahmood, K. Shi, S. Khatoon, and M. Xiao, "Data mining techniques for wireless sensor networks: a survey," *International Journal of Distributed Sensor Networks*, vol. 2013, pp. 185–193, 2013.

[5] H. Li, Y. Liu, R. Zhang, and G. Liu, "Design of community ECG monitoring system based on Wireless sensor Network," *Journal of Tianjin Polytechnic University*, vol. 34, no. 01, pp. 64–67, 2015.

[6] S. F. Tonellato, "Bayesian nonparametric clustering as a community detection problem," *Computational Statistics & Data Analysis*, vol. 152, 2020.

[7] J. Wang, X. Wang, and C. Zheng, "Design and implementation of intelligent community 3D display system," *Computer Technology and Development*, vol. 028, no. 009, pp. 156–161, 2018.

[8] X. Ma, J. Geng, and C. Fan, "A network community detection algorithm based on multi-view data fusion," *Computer Applications and Software*, vol. 35, no. 07, pp. 310–314, 2018.

[9] E. Luo, G. Wang, and C. Li, "Research on clustering algorithm for multi-dimensional data deduplication in big data environment," *Mini-Micro Systems*, vol. 37, no. 03, pp. 40–44, 2016.

Hindawi
Journal of Robotics
Volume 2022, Article ID 4029558, 9 pages
https://doi.org/10.1155/2022/4029558

Research Article

Computing Resource Allocation Strategy Using Biological Evolutionary Algorithm in UAV-Assisted Mobile Edge Computing

Li Wang, Xiang Yao, and Zhenqi Yuan

School of Internet of Things Engineering, Wuxi Taihu University, Wuxi, Jiangsu 214000, China

Correspondence should be addressed to Li Wang; wangl@wxu.edu.cn

Received 30 November 2021; Accepted 21 January 2022; Published 8 February 2022

Academic Editor: Shan Zhong

Aiming at the problems of high computing energy consumption and long time in traditional UAV-assisted edge computing research work, a computing resource allocation strategy using biological evolutionary algorithms in UAV-assisted mobile edge computing is proposed by introducing UAV swarms and genetic algorithms. Firstly, it analyzes the communication model for uplink transmission, the calculation model for local computing tasks, and UAV to perform computing tasks. Secondly, the objective function and overall model of system are constructed by comprehensively considering multiple constraints. Then, improved genetic algorithm is introduced into the model. On the basis of data encoding, crossover, mutation, and termination operations, the optimization performance of algorithm is greatly improved by multiple iterations of fitness function. Finally, the energy consumption of proposed algorithm and other two algorithms under the same number of iterations are compared and analyzed by simulation experiments. The experimental results show that the optimal solution, average, and variance of proposed algorithm for energy consumption are 52.354, 50.326, and 0.224, respectively, and its performance is better than other two comparison algorithms.

1. Introduction

Communication technology and Unmanned Aerial Vehicle (UAV) technology, as an important technological foundation of human information age, have developed rapidly in recent years. With the continuous advancement of related technologies, UAV technology has gradually developed from military only to military, civilian, and commercial use, and it has been widely used in various fields of society [1–3]. Following the development and widespread popularity of mobile terminal devices has provided a broad stage for a variety of intelligent applications. But at the same time, it also brings many problems that cannot be ignored, such as limited computing resources and limited battery capacity [4,5]. Therefore, studying reasonable calculation methods to reduce network congestion and transmission delay in the application of mobile devices is the focus of current industry research [6,7].

In response to related issues, the most forward-looking and research-oriented algorithm is mobile edge computing. It can marginalize cloud computing and information technology services in the process of providing auxiliary computing [8]. In addition, combined with the characteristics of high flexibility and mobility of UAVs, related problems can be better solved [9,10]. Reference [11] studied the problem of maximizing computing efficiency in mobile edge computing networks powered by wireless power in partial and binary computing offloading modes. It maximized the calculation efficiency under the maximum-minimum fairness criterion through joint optimization of energy harvesting time, local calculation frequency, offloading time, and power. However, this method did not consider frequency division multiple access and cannot realize D2D communication of different users on orthogonal frequency bands. Reference [12] used drones as computing servers to help user devices calculate its tasks or act as a repeater for further offloading computing tasks to the access point. They proposed a UAV-assisted mobile edge computing architecture, which minimizes the weighted total energy consumption under given multiple constraints.

However, this method did not consider the problem of safe interruption when the energy of user devices was limited. Reference [13] studied the weighted sum energy minimization problem under the security offload rate constraint, the calculation delay constraint, and the security interruption probability constraint. They also considered the actual passive eavesdropping scenarios and used confidentiality interruption probability to measure the confidentiality performance of computing offloading and characterized the best confidentiality offloading rate and power allocation. However, this method did not consider the balance between task delay and network energy consumption. Reference [14] proposed a UAV-assisted nonorthogonal multiple access network, which used UAVs and base stations to provide corresponding services for ground users. They realized UAV trajectory optimization through alternate user scheduling and obtain the optimal nonorthogonal multiple access precoding vector by two schemes with different constraints and maximize the total calculation rate on this basis. However, this method did not propose a corresponding solution to the highly coupled nonconvex weighting problem. Reference [15] proposed a mobile edge-assisted computing method and a relay scheme, which defined the cost function of trade-off between energy consumption and delay time through research. Combined with the minimization of cost function relative to the actual constraints, a closed-form solution was derived, which improved the throughput of uncompressed data in mobile peer-to-peer communication. However, this method did not provide an optimal solution to the nonconvex problem formed. Reference [16] gives computing power to the moving drone and provides computing offloading opportunities for mobile users with limited local processing power. It studied mobile cloud computing system based on drones and realized the quality of service requirements for offloading mobile applications on the basis of minimizing total energy consumption of mobile. However, the effective utilization of resources in the process of UAV relay providing services to mobile users in this method is low. In the multi-UAV-assisted mobile edge computing system, in [17] based on the partial computing offload mode, joint optimization of user association, CPU cycle frequency, power and spectrum resource allocation, and UAV trajectory scheduling maximizing computing efficiency problem was studied. They also proposed an iterative optimization algorithm with a double loop structure to find the optimal solution for the problem of nonconvexity and coupling between variables. However, this method did not significantly improve the communication delay problem.

Based on the above analysis, aiming at the problems of high computing energy consumption and long time in traditional UAV-assisted edge computing research work, a computing resource allocation strategy using biological evolutionary algorithm in UAV-assisted mobile edge computing is proposed. The basic idea is as follows: ① First, build a multi-UAV-assisted mobile edge computing system overall model and corresponding objective function by establishing communication model and calculation model. ② Optimize the computing resource allocation by introducing improved genetic algorithm to improve the optimization performance of algorithm. Compared with traditional computing resource allocation strategies, the innovation of proposed method lies in the following:

(1) The proposed computing resource allocation strategy studies the communication model of uplink transmission, the computing model of computing task time, and energy consumption and builds the system model and objective function on this basis

(2) The overall model and objective function of system are constructed, and the optimization performance of algorithm is improved by introducing an improved genetic algorithm

2. System Model and Optimization Goal

2.1. System Model. The overall model of system uses mobile edge computing system including u user devices and v UAVs, and user devices are randomly allocated in a specified area, where $u \in U\{1, 2, 3, \ldots, U\}$, $v \in V\{1, 2, 3, \ldots, V\}$, and the overall model of system is shown in Figure 1.

As shown in Figure 1, each user device of system model needs to use a binary offloading strategy to perform its own computationally intensive tasks. The expression for computationally intensive tasks is shown in

$$R_u = (T_u, X_u, Q_u). \tag{1}$$

In formula (1), T_u represents the total number of cycles experienced after task R_u is completed. X_u represents the amount of data offloaded after task R_u is completed. Q_u represents constraints, including time constraints and user service quality constraints.

Completing the offloading task of calculation requires three steps: data transmission, calculation execution, and result return. The focus is on the two parts of data transmission and calculation execution, and the result return part does not need to be calculated [18].

2.2. Communication Model. The following is an analysis of communication model in uplink transmission process, assuming that the offloading decision of user devices is $m_{uv} = \{0, 1\}$, where $v \in V$ and $v \in U$. The meaning of offloading decision m_{uv} is as follows: when $m_{uv} = 0$, it means that u user devices perform calculation locally; when M1, it means that the u user devices offload the calculation to v UAVs. The offloading decision is that m_{uv} satisfies the relationship shown in the following:

$$\sum_{v \in V} m_{uv} = 1, \quad \forall u \in U. \tag{2}$$

It can be seen from formula (2) that each offload task of user devices can only perform calculations on one UAV.

User devices can be divided into two forms: fixed and mobile. The position of a fixed form of devices is fixed during use. However, the position of a mobile device changes during use. Due to the extremely fast data transmission speed, the position of mobile form devices during the data

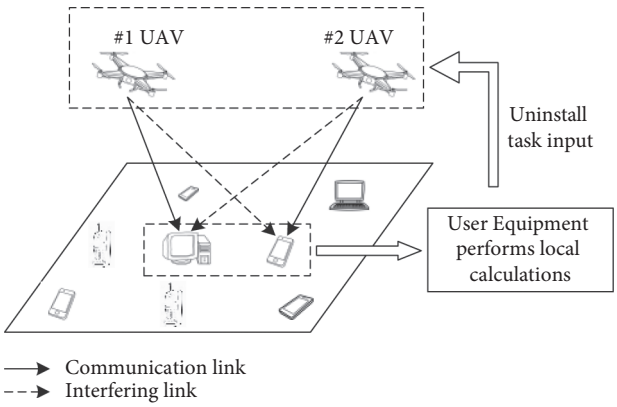


Figure 1: Multi-UAV-assisted mobile edge computing system.

transmission process can also be regarded as fixed [19]. At this time, if the u user devices satisfy $m_{uv} = 1$, then u user devices will offload the calculation to v UAVs. The transmission rate can be expressed by the following formula:

$$s = B \cdot \log_2\left(1 + \frac{P_{Q,uv} z_{uv}}{\alpha^2}\right). \tag{3}$$

In formula (3), B represents the bandwidth of signal transmission channel. $P_{Q,uv}$ represents the transmit power of signal. α represents the noise power of UAV. z_{uv} represents the gain coefficient of signal transmission channel, which can be calculated by the following formula:

$$z_{uv} = \frac{\varphi}{G_v^2 + L_{uv}^2}. \tag{4}$$

In formula (4), φ represents the small-scale fading component. G_v represents the flight height of v UAV from the ground. In the calculation process, it is considered that all UAVs have the same flight height from the ground. L_{uv} represents the horizontal distance between u user devices and v UAVs.

2.3. Calculation Model. The following is an analysis of time and energy consumption of local computing tasks and UAV computing tasks in the calculation process. First, analyze the calculation model of local computing task.

It is assumed that the local computing capability of each user devices is the same as N_{u1}, and the total number of cycles experienced after the completion of computing task R_u is T_u. Then the time t_1 required for the u user devices to complete the local calculation can be expressed by the following formula:

$$t_1 = \frac{T_u}{N_{u1}}, \quad \forall u \in U. \tag{5}$$

According to formula (5) power P_{u1} required for u user devices to complete the local calculation can be obtained, as shown in the following formula:

$$P_{u1} = c_u \cdot N_{u1}^{m_u}, \quad \forall u \in U. \tag{6}$$

In formula (6), m_u is a coefficient, and its value is usually set to 3. c_u is the effective switched capacitance, and its value is general.

Then, we analyze the calculation model of UAV to perform computing tasks.

When offloading decision $m_{uv} = 1$, user devices will offload the calculation to UAV for execution. At this time, it is necessary to clarify whether user devices are within the coverage of UAV, and the judgment basis for judging whether u user devices is within the coverage of v UAVs is shown in the following formula:

$$F_{uv} \le G_u \tan \beta_v, \quad \forall u \in U, v \in V. \tag{7}$$

In formula (7), β_v represents the communication angle of v UAVs. F_{uv} can be calculated with the following formula:

$$F_{uv} = \sqrt{(a_u - A_v)^2 + (b_u - B_v)^2}. \tag{8}$$

In formula (8), (a_u, b_u) represents the coordinate position of u user devices. (A_v, B_v) represents the plane coordinate position of v UAVs at the height of G_u from the ground.

According to equation (7), if user devices are within the coverage of UAV, then continue to analyze the calculation model of UAV to perform computing tasks. Set the data transmission time as $t_{s,uv}$ and task execution time as $t_{r,uv}$; then according to the proposed communication model, the total time t_2 spent by u user devices to execute tasks on v UAVs can be obtained, as shown in the following:

$$t_2 = m_{uv} \cdot \left(t_{s,uv} + t_{r,uv}\right), \quad \forall u \in U, v \in V. \tag{9}$$

In equation (9), the calculation methods for data transmission time $t_{s,uv}$ and task execution time $t_{r,uv}$ are as shown in

$$\begin{cases} t_{s,uv} = \dfrac{X_u}{s_{uv}}, \\ t_{r,uv} = \dfrac{T_u}{N_{uv}}. \end{cases} \tag{10}$$

In formula (10), N_{uv} represents the computing resource allocated by v UAVs to u user devices. All user devices should meet the constraint conditions shown in equation (11) in terms of time.

$$m_{uv}\frac{T_u}{N_{u1}} + \sum_{v=1}^{V} m_{uv}\left(\frac{X_u}{s_{uv}} + \frac{T_u}{N_{uv}}\right) \le t. \tag{11}$$

In formula (11), t is the total time.

At the same time, since the overall resources of UAV are limited, N_{uv} should meet the constraints shown in following:

$$\sum_{u \in U} N_{uv} \le N_{v,\max}. \tag{12}$$

It can be known from equation (12) that the resources allocated to user devices by each UAV must be less than the total amount of resources of UAV.

The energy consumed by the suspension and calculation of UAV is very small and can be ignored [18]. Based on this premise, the energy consumption of u user devices offloaded to v UAVs can be calculated, and the result is shown in the following formula:

$$P_{uv} = P_{Q,uv}\frac{X_u}{s_{uv}}, \quad \forall u \in U, v \in V. \tag{13}$$

In formula (13), $P_{Q,uv}$ represents the transmit power of signal. X_u represents the amount of data offloaded after task R_u is completed. s_{uv} represents the transmission rate of signal.

2.4. Optimization Goal. The following is to analyze and optimize the objective function of system model. For the overall system model proposed above, the corresponding objective function is given as shown in the following:

$$\min_{\{M_{UAV},M_r,M_N\}} \sum_{u\in U}\sum_{v\in V} m_{uv}\cdot P_{Q,uv}\cdot\frac{X_u}{s_{uv}} + \sum_{u\in U} P_{uv}\frac{T_u}{N_{uv}}. \tag{14}$$

In formula (14), M_{UAV} represents the position matrix of UAV. M_r represents the offloading matrix of task. M_N represents the allocation matrix of UAV resources.

The constraint conditions are shown in the following:

$$\begin{cases} (1)\, m_{uv} = \{0,1\}, \quad \forall u \in U, v \in V, \\ (2)\, \sum_{u\in U} m_{uv} = 1, \quad \forall u \in U, \\ (3)\, m_{uv}\frac{T_u}{N_{u1}} + \sum_{v=1}^{V} m_{uv}\left(\frac{X_u}{s_{uv}} + \frac{T_u}{N_{uv}}\right) \le t, \quad \forall u \in U, \\ (4)\, \sum_{u\in U} N_{uv} \le N_{v,\max}, \quad v \in V, \\ (5)\, F_{uv} \le G_u \tan\beta_v, \quad \forall u \in U, v \in V. \end{cases} \tag{15}$$

In formula (15), the constraint condition (1) indicates that all tasks can only be executed locally or offloaded to UAV for execution. Constraint (2) means that if tasks are offloaded to UAV for execution, then each task can only be offloaded to a certain UAV for execution. Constraint (3) represents the delay constraint of user devices. Constraint (4) means that the resources allocated to user devices by each UAV must be less than the total amount of resources of UAV. Constraint condition (5) represents UAV coverage range constraint.

3. Computing Allocation Strategy Based on Improved Genetic Algorithm

Due to the many constraints of objective function, the difficulty of solving the optimal computing offloading strategy is greatly increased, and it is difficult to calculate the optimal solution in polynomial time. Therefore, an improved genetic algorithm is introduced to find the optimal solution of task offloading strategy. The basic principle of genetic algorithm is biological evolution process based on natural selection and genetics mechanism. It is usually used to solve nonlinear, global, and more complex combinatorial optimization problems, and it is a high-performance computing model [20]. The main idea is to simulate the natural selection concept of survival of fittest in nature. By converting the problem-solving process into a process similar to the crossover and mutation of chromosomal genes in biological evolution, on this basis, iterative calculations are used to gradually approximate and obtain the optimal solution.

3.1. Coding and Initial Population. The traditional genetic algorithm does not directly process the parameters of solution space but first encodes parameters and then continues to process [21]. The encoding operation is to convert parameters in the solution space into individuals or chromosomes in the space. These individuals or chromosomes are arranged according to a specific structure of genes, and the process of parameter conversion is the process of encoding. After the parameters are encoded, crossover, mutation, and genetic operations can be carried out more conveniently.

In general, encoding methods include gray code encoding, binary encoding, and symbol encoding. From the analysis of communication model in Section 2.1, it can be known that the user's offloading decision m_{uv} is only 0 and 1, which can be regarded as binary variables. Thus, the choice of binary encoding can greatly reduce the complexity when choosing the encoding method. When encoding, first set the initial population set as P_0, $P_0 = \{B_1, B_2, B_3, \ldots, B_k, \ldots, B_K\}$, where B_k represents the k individual or chromosome and satisfies $|P_0| = K$. U represents the total number of user devices in network, and U binary digits are used to construct the k individual or chromosome: $B_k = \{b_1^k, b_2^k, b_3^k, \ldots, b_u^k, \ldots, b_U^k\}$, $k \in \{1, 2, 3, \ldots, K\}$. On this basis, the specific steps of population initialization algorithm used to generate initial population set P_0 can be obtained, as shown in the following:

(1) First, let $k = 1$ and randomly generate a random number that obeys a uniform distribution in the range of [0, 1] as the probability of randomly generating a gene u in the k individual or chromosome. This random number is denoted as p_u^k.

(2) Let $u = 1$.

(3) Judge whether p_u^k is less than or equal to 0.5.

(4) If p_u^k is less than or equal to 0.5, then $b_u^k = 0$; otherwise, $b_u^k = 1$.

(5) Execute $u = u + 1$ operation, and return to step (3) until $u = U$ is established, and then jump out of loop.

(6) Execute $k = k + 1$ operation, and return to step (2), until $k = K$ is established, the end.

3.2. Fitness Function. The main function of fitness function is to evaluate the performance of an individual or chromosome, and its principle basis is the degree of fitness of

individual in the natural environment [22]. The fitness function can be used to distinguish the pros and cons of different individuals or chromosomes. If the value of fitness function is larger, it means that the individual or chromosome is more adaptable in the natural environment, and it is easier to be retained in genetic evolution, and the corresponding genetic decision-making performance is better.

According to the objective function shown in equation (14), the reciprocal of total energy consumption of all user devices is used as the fitness function, as shown in the following equation:

$$F = \frac{1}{C} = \frac{1}{\sum_{i\in R}\left[\left(1 - L_j\right)C_j + L_j C_j\right] + \sum_{i,j\in R}\left[\left|L_i - L_j\right|C_{ij}\right]}. \tag{16}$$

In formula (16), C represents the total energy consumed by user devices. The L_i and L_j subtables indicate the execution locations of i and j tasks. C_i and C_j represent the total energy consumed to perform the i and j tasks, respectively. C_{ij} represents the total energy consumed in the process of receiving data.

According to the fitness function shown in formula (16), the pros and cons of different individuals or chromosomes can be evaluated and distinguished. It can be seen from formula (16) that the smaller the total energy consumption of a certain body or chromosome, the larger the fitness function value, and the easier it is to be retained in genetic evolution. Taking the fitness function value as the basic basis can determine the probability of different individuals or chromosomes appearing in the next generation. On this basis, some individuals or chromosomes are selected from the parents according to roulette selection method to form a group of offspring.

3.3. Genetic Operation

3.3.1. Cross Operation. The main principle of crossover operation is that some genes in a certain body or chromosome in the parent are exchanged according to a specific rule and produce offspring [23].

For example, in single-point crossover, the crossover process is to randomly select a point as crossover point in a certain body or chromosome in the parent. Besides, based on the crossover point, the genes on the left and right sides of crossover point in the parent or chromosome are exchanged, and two new chromosomes are generated. The principle process is shown in Figure 2.

3.3.2. Mutation Operation. The main principle of mutation operation process is based on a specific mutation probability, by replacing part of gene of an individual or chromosome in the parent with other genes. The purpose is to generate new individuals and maintain the diversity of entire population to a greater extent, thereby enhancing the overall local search capability of genetic algorithm to a certain extent [24].

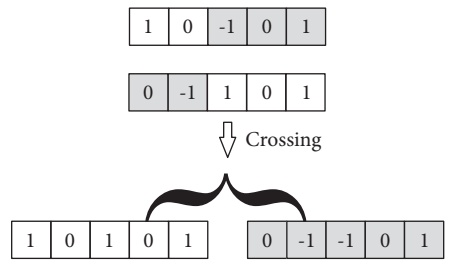


Figure 2: The basic principle of single-point crossing.

In general, variation can be divided into multiple variations such as basic position variation, uniform variation, boundary variation, and Gaussian approximation variation. In order to be consistent with the encoding method, the binary-based basic bit mutation method is selected here. The basic principle of mutation process is shown in Figure 3.

3.4. Genetic Termination Condition. After genetic crossover and mutation operations, a new offspring population can be obtained according to the parent population. On this basis, the minimum energy consumption objective function is solved according to the proposed multiple constraints, and the performance of newly generated offspring population is evaluated. Within the specified maximum number of iterations, as long as the number of iteration calculations does not exceed the set maximum number of iterations, then continue and repeat the crossover and mutation operations in genetic algorithm until the termination condition of objective function solution of genetic algorithm is met. If the number of iterative calculations reaches the set maximum number of iterations or fitness function of best fitness individual in the successive generations of offspring populations is still not significantly improved. Then it is considered that the termination condition for solving objective function is reached and the iterative calculation is stopped.

4. Experiment and Analysis

4.1. Simulation Setting. The simulation experiment platform uses MATLAB mathematical software, version 2019a. The computer hardware conditions used in the simulation are as follows: CPU is i7-7200U; running memory size is 4G. During the simulation experiment, overall horizontal area of the experiment was set as a square area with an area of 100 m × 100 m. The user devices are randomly distributed in the test area, and specific parameters of simulation experiment are shown in Table 1.

4.2. Simulation Analysis. Different parameter settings have different effects on the performance of genetic algorithm, such as the number of iterations. Too many iterations will increase the amount of calculation and prolong calculation time of algorithm. Too few iterations will cause the algorithm to fail to converge, and the optimal solution cannot be obtained. In order to make algorithm obtain the best performance, firstly, the overall performance of proposed

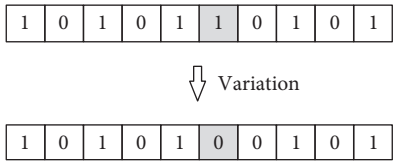


Figure 3: The basic principle of bit variation.

Table 1: Simulation experiment parameters.

Parameter	Value
Number of user devices	200/400/600
Number of UAV	4/10/15
Transmit power (W)	1
The size of transmitted data (kb)	5×10^4
The number of cycles to complete the task	10^9
Total time (s)	3
Channel bandwidth	10^6
Energy consumption for local execution (W)	1
The total resource size of drone (kb)	10^{10}
The height of drone from the ground (m)	30

algorithm is simulated and analyzed according to the number of genetic iterations. The relationship between the number of iterations and the total energy cost of algorithm is shown in Figure 4.

It can be seen from Figure 4 that the total energy cost of algorithm is different under different iteration times, and the smaller the number of iterations, the greater the total energy cost of algorithm. With the gradual increase in the number of iterations, the total energy cost of algorithm also gradually decreases. When the number of iterations reaches about 35 times, the total energy cost of algorithm no longer changes. This is because when the number of iterations is too small, it will stop the iterative calculation if the algorithm fails to converge, and the optimal solution cannot be found. Thus, the energy cost of algorithm will be greater. After 35 iterations, the algorithm has found the optimal solution, and increasing the number of iterations will not improve the performance of algorithm.

The task completion time of resource allocation strategy, the energy consumption required to calculate the specific location of UAVs, the influence of number of UAVs on the energy consumption of algorithm, and the influence of number of user devices on the energy consumption of algorithm are all important evaluation indicators to measure the strategy performance. The following is a comparative analysis of algorithm proposed in this paper and algorithms in [12] and [17] under the same conditions for the above four evaluation indicators.

Firstly, calculate the task completion time for different algorithm resource allocation strategies. The average task completion time of different algorithms under different number of tasks is shown in Figure 5.

It can be seen from Figure 5 that, in the case of different number of tasks, the algorithm proposed in this paper has the smallest average time for task completion compared with the algorithms in [12] and [17]. And as the number of tasks increases, the growth rate of average task completion time is

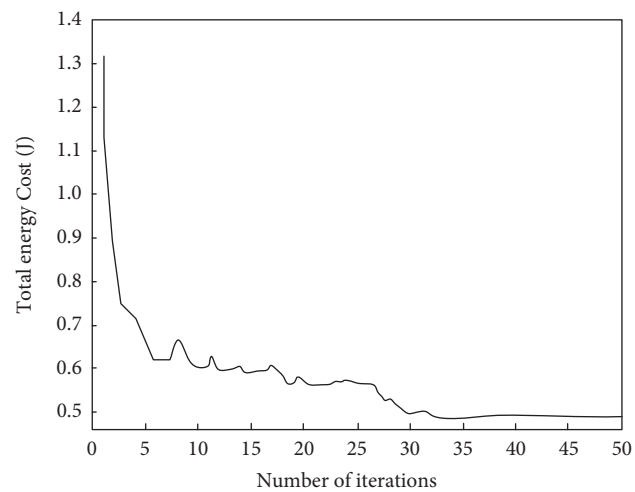


Figure 4: The relationship between the total cost of algorithm and the number of iterations.

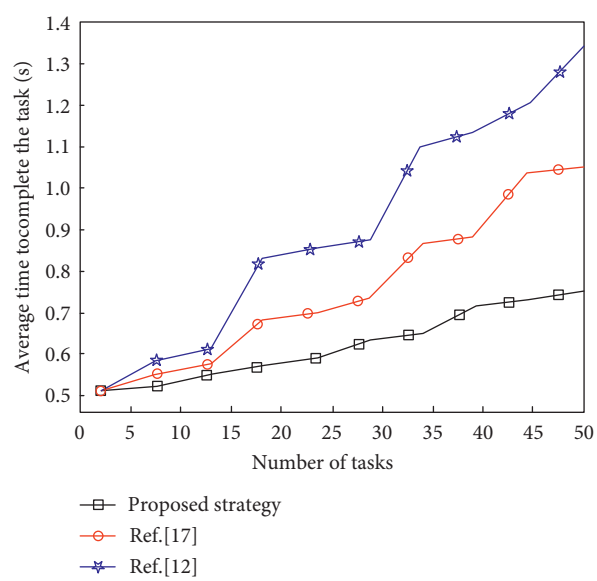


Figure 5: The average time for different algorithms to complete tasks under different number of tasks.

relatively small. This is because the computing resource allocation strategy proposed in this paper can better prioritize the migration of tasks to any UAVs with the fastest response time. Although [12] combined UAV and edge computing methods and proposed an overall architecture, it did not optimize the solution process of objective function, and the solution process took more time. Reference [17] used an iterative optimization algorithm to study the overall calculation efficiency maximization. However, the calculation models of time consumption and energy consumption of local tasks and UAV tasks were not analyzed and improved in detail, and time and energy consumption were not saved to the best.

The following is an analysis of energy consumption of tasks performed by different resource allocation strategies.

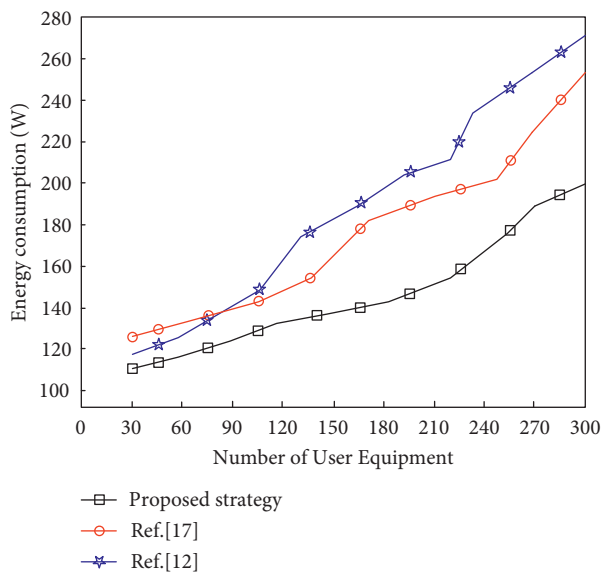


Figure 6: The relationship between energy consumption and the number of user devices of different algorithms.

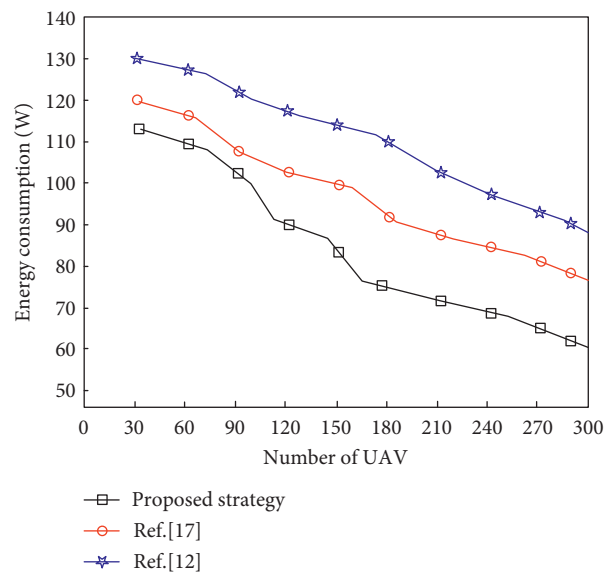


Figure 7: The relationship between energy consumption and the number of UAV of different algorithms.

When the number of UAVs is the same, the energy consumption of different strategies when different numbers of user devices are adopted is shown in Figure 6.

It can be seen from Figure 6 that, under the premise of using the same number of UAVs, the energy consumption of three algorithms will increase as the number of users increases. This is because an increase in the number of users means an increase in the number of tasks, and more energy will be consumed to perform more tasks. At the same time, it can be seen that the algorithm proposed in this paper consumes the least energy in the case of different numbers of users. And as the number of user devices increases, the growth rate of its energy consumption is also the lowest.

When the number of user devices is the same, the energy consumption of different strategies when different numbers of UAVs are adopted is shown in Figure 7.

It can be seen from Figure 7 that, under the premise of using the same number of user devices, the energy consumption of three algorithms will decrease as the number of UAVs increases. At the same time, it can be seen that the algorithm proposed in this paper consumes the least energy when using different numbers of UAVs. Besides, as the number of UAVs increases, the average rate of change in their energy consumption is also the largest.

The following is an analysis of three algorithms to calculate energy consumption of a specific UAV location under the same conditions. The energy consumed by using different algorithms to calculate UAV position is shown in Table 2.

In Table 2, Average represents the optimal solution of algorithm under the same number of iterations, Optimal represents the average value of all solutions under the same number of iterations, and Variance represents the variance of solution under the same number of iterations. It can be seen that, compared with the other two algorithms, the optimal solution and average value of algorithm under the same number of iterations are relatively small, 52.354 and 50.326, respectively. The variance value of its energy consumption is 0.224. This shows that the algorithm proposed in this paper has the best performance.

Table 2: The energy consumed by calculating the location of UAV by different algorithms.

Algorithms	Index	Energy consumed
Algorithm in this paper	Average	52.354
	Optimal	50.326
	Variance	0.224
K-means	Average	60.469
	Optimal	58.978
	Variance	0.523
Fuzzy C-Means	Average	65.642
	Optimal	64.434
	Variance	0.573

The allocation strategy proposed in this paper can always consume less energy than the other two algorithms when the number of user devices or the number of UAVs changes. This is because the improved genetic algorithm is introduced into the model, and the optimization performance of fitness function is improved by operations such as coding, crossover, and genetics and further reduces the energy consumption of computing resource allocation strategy. As the number of user devices or the number of UAVs changes, genetic operations and iterative calculations will be performed again, and the optimal resource allocation will always be maintained. The calculation model is analyzed in detail, which makes up for the defect that the objective function solving process is not optimized in [12] and the energy consumption cannot be saved to the best in [17].

5. Conclusion

Aiming at the problem of traditional UAV-assisted edge computing, this paper proposes a computing resource allocation strategy using biological evolutionary algorithms in UAV-assisted mobile edge computing. Analyze the constraints of proposed objective function in many aspects, which improves the overall computing performance of system model to a certain extent. The simulation analysis of proposed algorithm is carried out by simulation experiments, and experimental results show that, compared with traditional algorithms, the average time and energy consumption of proposed algorithm are the smallest. In addition, its average task completion time increases with the increase in the number of tasks, and the increase rate of energy consumption with the increase in number of user devices is the smallest. The optimal solution of energy consumption under the same number of iterations, the average value, and variance are the smallest, and the overall performance of algorithm is the highest. Although the proposed method has achieved some results, the flight altitude of UAV is often not fixed in the actual scene. In the future, the performance of the proposed algorithm in task calculation time and energy consumption in the case of random flight altitude of UAV will be deeply studied.

Data Availability

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

Conflicts of Interest

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

Acknowledgments

This work was supported by the Jiangsu Key Laboratory of loT Application Technology (Wuxi Taihu University) and the Natural Science Research Project of Colleges and Universities in Jiangsu Province (no. 18KJB413009).

References

[1] Y. Chen, N. Zhao, Z. Ding, and M.-S. Alouini, "Multiple UAVs as relays: multi-hop single link versus multiple dual-hop links," *IEEE Transactions on Wireless Communications*, vol. 17, no. 9, pp. 6348–6359, 2018.

[2] Y. Mao, J. Zhang, S. H. Song, and K. B. Letaief, "Stochastic joint radio and computational resource management for multi-user mobile-edge computing systems," *IEEE Transactions on Wireless Communications*, vol. 16, no. 9, pp. 5994–6009, 2017.

[3] H. Zhao, H. Wang, W. Wu, and J. Wei, "Deployment algorithms for UAV airborne networks toward on-demand coverage," *IEEE Journal on Selected Areas in Communications*, vol. 36, no. 9, pp. 2015–2031, 2018.

[4] Y. Zeng, R. Zhang, and T. J. Lim, "Throughput maximization for UAV-enabled mobile relaying systems," *IEEE Transactions on Communications*, vol. 64, no. 12, pp. 4983–4996, 2016.

[5] Y. Zeng, R. Zhang, and T. J. Lim, "Wireless communications with unmanned aerial vehicles: opportunities and challenges," *IEEE Communications Magazine*, vol. 54, no. 5, pp. 36–42, 2016.

[6] T. Oktay, H. Celik, and I. Turkmen, "Maximizing autonomous performance of fixed-wing unmanned aerial vehicle to reduce motion blur in taken images," *Proceedings of the Institution of Mechanical Engineers-Part I: Journal of Systems & Control Engineering*, vol. 232, no. 7, pp. 857–868, 2018.

[7] L. Gupta, R. Jain, and G. Vaszkun, "Survey of important issues in UAV communication networks," *IEEE Communications Surveys & Tutorials*, vol. 18, no. 2, pp. 1123–1152, 2016.

[8] J. Du, Y. Zhang, Y. Chen, X. Li, Y. Cheng, and M. V. Rajesh, "Hybrid beamforming NOMA for mmWave half-duplex UAV relay-assisted B5G/6G IoT networks," *Computer Communications*, vol. 180, no. 32, pp. 232–242, 2021.

[9] X. H. Chen, Q. J. Shi, and Y. L. Cai, "Joint cooperative computation and interactive communication for relay-assisted mobile edge computing," in *Proceedings of the 2018 IEEE 88th Vehicular Technology Conference*, pp. 1–5, IEEE Press, Piscataway, NJ, USA, 2018.

[10] W. Wu, F. H. Zhou, and P. Li, "Energy-efficient secure NOMA-enabled mobile edge computing networks," in *Proceedings of the 2019 IEEE International Conference on Communications*, pp. 1–6, IEEE Press, Piscataway, NJ, USA, 2019.

[11] F. Zhou and R. Q. Hu, "Computation efficiency maximization in wireless-powered mobile edge computing networks," *IEEE Transactions on Wireless Communications*, vol. 19, no. 5, pp. 3170–3184, 2020.

[12] X. Hu, K.-K. Wong, K. Yang, and Z. Zheng, "UAV-assisted relaying and edge computing: scheduling and trajectory optimization," *IEEE Transactions on Wireless Communications*, vol. 18, no. 10, pp. 4738–4752, 2019.

[13] W. Wu, F. Zhou, R. Q. Hu, and B. Wang, "Energy-efficient resource allocation for secure NOMA-enabled mobile edge computing networks," *IEEE Transactions on Communications*, vol. 68, no. 1, pp. 493–505, 2020.

[14] N. Zhao, X. Pang, Z. Li et al., "Joint trajectory and precoding optimization for UAV-assisted NOMA networks," *IEEE Transactions on Communications*, vol. 67, no. 5, pp. 3723–3735, 2019.

[15] M. Qin, L. Chen, N. Zhao, Y. Chen, F. R. Yu, and G. Wei, "Computing and relaying: utilizing mobile edge computing for P2P communications," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 2, pp. 1582–1594, 2020.

[16] S. Jeong, O. Simeone, and J. Kang, "Mobile edge computing via a UAV-mounted cloudlet: optimization of bit allocation and path planning," *IEEE Transactions on Vehicular Technology*, vol. 67, no. 3, pp. 2049–2063, 2018.

[17] J. Zhang, L. Zhou, F. Zhou et al., "Computation-efficient offloading and trajectory scheduling for multi-UAV assisted mobile edge computing," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 2, pp. 2114–2125, 2020.

[18] L. Liu, Y. Zhou, J. Yuan, W. Zhuang, and Y. Wang, "Economically optimal MS association for multimedia content delivery in cache-enabled heterogeneous cloud radio access networks," *IEEE Journal on Selected Areas in Communications*, vol. 37, no. 7, pp. 1584–1593, 2019.

[19] Y. Zhou, L. Tian, and L. Liu, "Fog computing enabled future mobile communication networks: a convergence of communication and computing," *IEEE Communications Magazine*, vol. 57, no. 5, pp. 20–27, 2020.

[20] Q. Chen, Z. B. Liu, and L. N. Ruan, "Energy-Efficient task caching and offloading strategy in mobile edge computing

systems," in *Proceedings of the 2nd International Conference on Security with Intelligent Computing and Big-data Services (SICBS)*, pp. 824–837, Guilin, China, 2018.

[21] A. Bonadio, F. Chiti, and R. Fantacci, "Performance analysis of an edge computing SaaS system for mobile users," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 2, pp. 2049–2057, 2020.

[22] Y. Guo, S. Wang, A. Zhou, J. Xu, J. Yuan, and C. H. Hsu, "User allocation-aware edge cloud placement in mobile edge computing," *Software: Practice and Experience*, vol. 50, no. 5, pp. 489–502, 2020.

[23] Q. Y. Fan, J. Y. Lin, and G. S. Feng, "Joint service caching and computation offloading to maximize system profits in mobile edge-cloud computing," in *Proceedings of the 2020 16th International Conference on Mobility, Sensing and Networking (MSN 2020)*, pp. 244–251, Tokyo, Japan, December 2020.

[24] H. J. Hong, W. J. Fan, and C. E. Chow, "Optimizing social welfare for task offloading in mobile edge computing," in *Proceedings of the 19th IFIP Networking Conference (Networking)*, pp. 524–528, Paris, France, June 2020.

Hindawi
Journal of Robotics
Volume 2022, Article ID 4969883, 9 pages
https://doi.org/10.1155/2022/4969883

Research Article

Expression Recognition Using Improved AlexNet Network in Robot Intelligent Interactive System

Yifeng Zhao and Deyun Chen

School of Computer Science and Technology, Harbin University of Science and Technology, Harbin, Heilongjiang 150080, China

Correspondence should be addressed to Deyun Chen; chendeyun@hrbust.edu.cn

Received 2 December 2021; Accepted 15 January 2022; Published 8 February 2022

Academic Editor: Shan Zhong

Aiming at the insufficient feature extraction in the expression feature extraction stage of traditional convolutional neural network and the misclassification of mislabeled samples, an expression recognition and robot intelligent interaction method using deep learning is proposed. First, in image preprocessing, the dimension of the color image is reduced by image gray adjustment to reduce the amount of calculation, the shadow interference is eliminated by the average method, and the image is enhanced by histogram equalization. Second, multichannel convolution is used to replace the single convolution size in the second convolution layer in AlexNet, the Global Average Pooling layer is introduced to replace the fully connected layer, and Batch Normalization is introduced to improve the feature extraction ability of the model and avoid gradient explosion. Finally, the Focal Loss is improved by setting the probability threshold to avoid the impact of mislabeling samples on the classification performance of the model. The experimental results show that the recognition accuracy of the model on FER2013 data set is 98.36%. The effectiveness of the algorithm is verified on the intelligent interactive system of service robot based on expression recognition. Compared with other expression recognition methods, the proposed method can extract more expression features and recognize facial expression more accurately.

1. Introduction

With the progress of society and the continuous development of science and technology, people pay more and more attention to intelligent robots and related fields. The wide application of robots and people's demand for robots promote the development of robot technology in a more intelligent direction [1]. In recent years, according to the requirements of China's national high-tech research and development plan in "13th five-year plan," in order to implement the "made in China 2025," robots will be taken as the key development field, and the overall deployment has been made Moreover, the research on service robot has been favored by many researchers and relevant research institutions, and it has become a hotspot in robot research [2–4]. As a key technology in the field of robotics, human-robot interaction (HRI) is a technology involving multidisciplinary fields, including human-computer interaction technology, artificial intelligence, robotics, and natural language understanding. HRI is a technology to study how to conduct friendly information interaction between human and robot [5–7]. At present, most of the interaction technology between human and robot mainly depends on human giving the specified commands to the robot to complete specific tasks. The robot cannot perceive and understand the user's purpose and intention, nor infer the user's psychological state according to the user's current external state to provide better interactive services. Robots still lack active awareness. In order for service robots to provide users with more convenient and humanized service experience, the traditional human-computer interaction technology must be upgraded so that robots can reasonably infer emotional state according to people's external state performance, so as to perceive and understand people's intentions and emotions and provide relevant services [8–11].

In recent years, the computing power of computer hardware has been greatly improved, including CPU (central processing unit), GPU (graphics processing unit), and artificial intelligence chips specially developed to accelerate the computing power of deep neural network, such as TPU (tensor processing unit) and NPU [12–15]. Because the field of computer vision needs strong computing power, especially the data-driven deep learning algorithm, which benefits from the improvement of computing power, it has achieved breakthrough results. Facial expression recognition technology is an interdisciplinary and comprehensive subject, which mainly includes behavior, psychology, sociology, and computer science [16–19]. Through the research of facial expression and emotional analysis, robots can perceive human psychological state and interpret human behavior intention. Especially in the field of human-computer interaction, intelligent interaction can be realized, so facial expression recognition technology has high scientific research and application value.

With the development of deep learning, more and more researchers pay attention to this hot field [20]. Facial expression recognition integrates many technologies such as machine learning, pattern recognition, and computer vision. Its application scenarios are mainly as follows. (1) In the field of intelligent transportation, it can detect the driver's expression in real time, judge the driver's mental state, and observe whether the driver is driving tired, or whether the driver is in negative emotional states such as anger and anxiety. If the driver is in poor condition, give an alarm and warning immediately. (2) In the field of service robots, with the continuous development of robot technology, service robots will eventually enter ordinary families. Human computer interaction technology based on emotion analysis divides the robot emotion interaction technology into three parts: robot emotion recognition, robot emotion model, and robot emotion feedback. In the whole closed-loop interaction technology, robot emotion perception technology is the basic research field.

As a humanized interaction mode of service robot, facial expression recognition technology collects, detects, and recognizes human facial expression information through the vision system and then analyzes human psychological behavior. In order to improve the accuracy of expression recognition, this paper improves the model on the basis of traditional AlexNet. The main innovations are as follows.

(1) Four different convolution kernels are used for multichannel convolution to extract features to a greater extent. The Global Average Pooling layer is used to replace the fully connected layer to avoid the problem of too many parameters.

(2) By setting the probability threshold, Focal Loss is improved and its confidence is changed, so as to reduce the attention of FL to this kind of samples and improve the classification performance of the model.

2. Related Researches

The research of expression recognition can be mainly divided into traditional feature extraction methods and deep learning methods. The traditional feature extraction mainly depends on the manually designed extractor, which requires a lot of professional knowledge. At the same time, the generalization and robustness are slightly insufficient compared with the deep learning method. The feature extraction of deep learning method updates and iterates the weights through back-propagation and error optimization algorithm to extract deeper and more abstract features in the process of learning a large number of samples. In recent years, many scholars have applied deep learning method to facial expression recognition and achieved good results. Reference [21] proposed a microexpression recognition method based on Wasserstein Generative Adversarial Network, established facial expression recognition network and facial identity recognition network, and improved the accuracy and robustness of facial expression recognition by suppressing intraclass changes. Reference [22] converted the expression image into Local Binary Pattern (LBP) feature map and then used the LBP feature map as the input of CNN for training, which had achieved good results. However, it will lead to low accuracy and insufficient robustness in unknown environment. Reference [23] combined ResNet-50 model and VGG16 model to form a new combined model for facial expression recognition and achieved good results on KDEF data set. In addition to improving the basic model and network structure, many researchers have also studied and improved the loss function. Reference [24] proposed the attention mechanism network SENet (Squeeze-and-Excitation Networks), which automatically obtained the importance of each feature channel through learning and then enhanced the features important to the current task and suppresses the features that are not useful to the current task according to the importance. Reference [25] used the method of spatial attention combined with multichannel connection to improve the convolutional neural network. First, finetune the pretrained model to obtain the feature map, add the spatial attention mechanism to highlight the expression area, and then classify the feature vectors with obvious discrimination. Reference [26] proposed a method to improve the fusion of convolutional neural network and attention mechanism, which integrated the global image features with multiple unobstructed facial regions of interest features, so as to improve the expression ability of regional features. Reference [27] proposed a facial expression recognition algorithm based on local features and deep belief network, extracted the nonuniform illumination invariant features of LSH in facial expression images, and extracted the edge detail features of facial expressions by using Gauss-Laplace operator. Reference [28] proposed a high-frequency edge digital signature feature extraction framework combined with histogram features, which obtains the dynamic digital signature descriptor of human face by projecting edge pixels vertically and horizontally. Digital signature uniquely and completely describes facial expression.

In this paper, the classical AlexNet network model is improved. Aiming at the insufficient feature extraction in the expression feature extraction stage of traditional convolutional neural network and the misclassification of mislabeled samples, an expression recognition and robot intelligent interaction method using deep learning is proposed. The

multichannel convolution is used to replace the single convolution size in the second convolution layer in the classical AlexNet. Global Average Pooling layer is used to replace the fully connected layer, and Batch Normalization operation is introduced to avoid the gradient explosion phenomenon. Focal Loss is improved by setting the probability threshold to avoid the impact of mislabeling samples on the classification performance of the model.

3. Facial Expression Recognition Based on Convolutional Neural Network

3.1. Overall Network Structure. Facial expression recognition is a typical image classification problem in computer vision, which mainly includes four parts. The flow chart of the designed facial expression recognition system is shown in Figure 1. In the facial expression recognition system, the first part is image acquisition, face detection, and image preprocessing. Fast and accurate face detection is carried out on the collected images. After the face region is detected, in order to reduce the influence of inconsistent scale size, image preprocessing is carried out. Then, Mobile Net is used to extract expression features, and expression features are extracted from the face region after image preprocessing. The effectiveness of features directly affects the accuracy of expression classification. After a lot of training and learning, the feature extraction algorithm can effectively extract expression features. The last is expression recognition. Through the final judgment and classification of the extracted expression features, the facial expression recognition is completed and the recognition results are output.

3.2. Image Preprocessing. When recognizing face information, it is necessary to collect face image through camera, and the image may be affected by factors such as illumination change, expression change, and partial shadow. Direct use of the collected image for subsequent processing will increase the amount of calculation and reduce the recognition accuracy. Therefore, it is necessary to preprocess the image before detecting face.

Image gray adjustment is to reduce the dimension of color image. After gray processing, the amount of calculation can be reduced, and the shadow interference can be eliminated by using average method. The average value of the three channels of each pixel point is taken as the processing result, that is:

$$\text{Gray}(i, j) = \frac{R(i, j) + G(i, j) + B(i, j)}{3}. \quad (1)$$

In order to improve the observability of the original image and improve the contrast of the image, histogram equalization is used to enhance the image. Its basic principle is to make the gray value of the image as evenly distributed as possible, so that the number of each gray level is basically the same. The histogram equalization function should meet two requirements.

(1) The value range of $f(x)$ is $[0, L-1]$, where L indicates gray level.

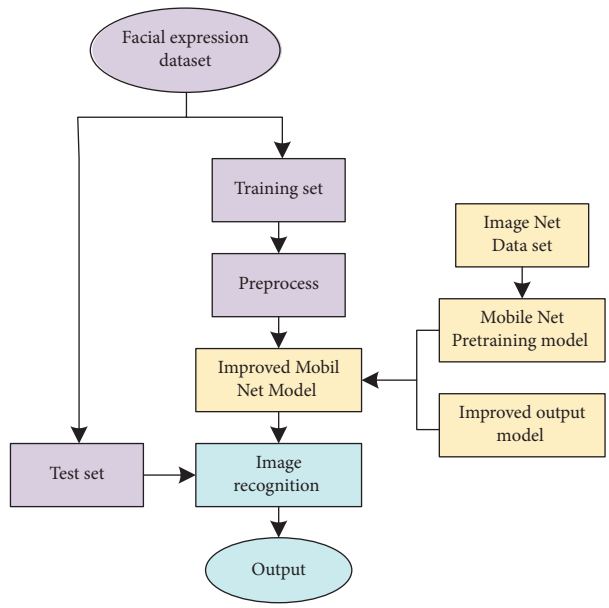


Figure 1: Overall framework of the proposed method.

(2) $f(x)$ is monotonically increasing in $[0, L-1]$.

Equation (2) is often used for histogram equalization in image processing, and this function can meet the above two requirements.

$$y = f(x) = \sum_{i=0}^{x_i} (L-1) \frac{h(x_i)}{n}, \quad (2)$$

where $h(x_i)$ is the vertical height of the gray value x_i in the histogram and n represents the total number of pixels.

3.3. Improved AlexNet Network. The convolution kernel size of convolution layer in AlexNet network is a single value, and the generated feature map features are not diverse, which will lead to insufficient feature extraction. This section makes improvements to the three defects: single size convolution kernel, too many parameters in the fully connected layer, and change of data distribution.

The traditional convolution layer uses a single-sized convolution kernel to convolute the input data for obtaining several feature maps. The multichannel convolution technology can be regarded as an extension of the traditional convolution, adding several convolution kernels of different sizes to a single convolution layer. It will make the generated feature map features more diverse, as shown in Figure 2.

Change the convolution kernel depth of 96 in the first layer to 64, and replace the convolution kernel with a single size of $5 * 5$ in the third layer with four convolution kernels. The sizes of the four convolution kernels are $1 * 1$, $3 * 3$, $5 * 5$, and $7 * 7$, respectively, and set the depth of the four convolution kernels to 64, respectively. Since the depth of the original convolution kernel with a single size of $5 * 5$ is 256, four convolution kernels with different sizes are replaced to combine the four convolution kernels with different sizes. Four different convolution kernels are used

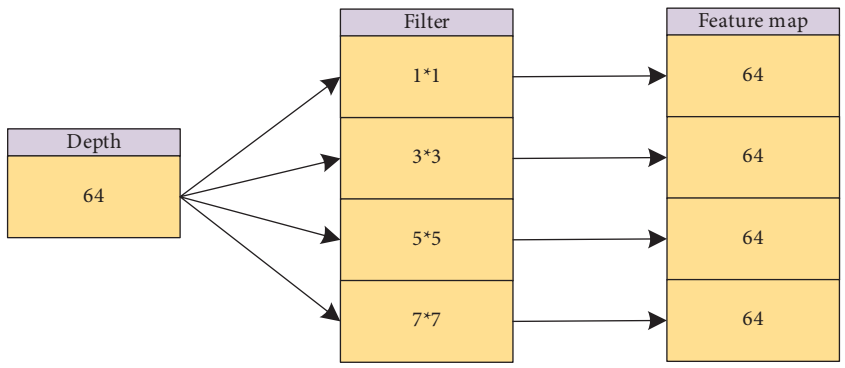


Figure 2: Multiscale convolution kernel operation.

for multichannel convolution to replace the third convolution layer of classical AlexNet network. Compared with the convolution kernel with single size, it strengthens the diversity of convolution kernels and extracts more features.

In order to avoid the problem of too many parameters, the Global Average Pooling (GAP) layer is used to replace the fully connected layer. After the fully connected layer expands the convolution layer into a vector, it is necessary to classify each feature map, and the idea of GAP is to combine the above two processes, as shown in Figure 3. The GAP is different from the average pooling. The average pooling has a filter size. Generally, it averages a subregion of the feature map and then uses the step size to slide the region. However, the GAP has no filter size, which is for the whole feature map.

In order to apply the Global Average Pooling to the classical AlexNet network, first, a convolution layer is added. The kernel size is $3 * 3$, the stripe is 1, the padding is 1, and the depth is 64. At this time, the output is $6 * 6 * 64$. Then, add a maximum pooling layer with filter of $3 * 3$ and stripe of 2, and then add an LRN layer. At this time, the output is $3 * 3 * 64$. Add a convolution layer with kernel size of $3 * 3$, stripe of 1, padding of 1, and depth of 5. At this time, the output is $3 * 3 * 5$. Finally, add GAP, that is, Global Average Pooling for the whole five feature maps, so as to obtain five features.

In the process of network training, the update of the parameters of the previous layer will affect the change of the data distribution of the latter layer. In order to solve the change of data distribution and ensure that the output data of each layer are on the same distribution, the normalized preprocessing is carried out first, and then the normalized data enter the next layer of the network. Therefore, Batch Normalization (BN) is introduced. The mean value of the data processed by BN algorithm is 0 and the variance is 1. The data normalization formula is shown in the following equation:

$$x^{(k)} = \frac{x^{(k)} - E\left[x^{(k)}\right]}{\sqrt{\mathrm{Var}\left[x^{(k)}\right]}}, \tag{3}$$

where $x^{(k)}$ is the input neuron and $E[x^{(k)}]$ is the average value of neurons in training data.

In order to improve its expression ability, learnable parameters γ and β are introduced.

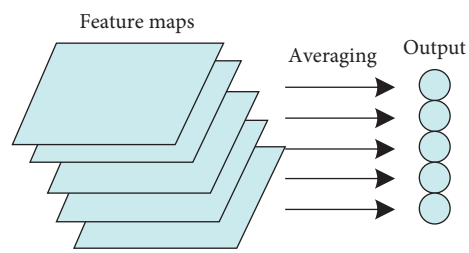


Figure 3: Schematic diagram of Global Average Pooling operation.

$$y^{(k)} = \gamma^{(k)} x^{(k)} + \beta^{(k)} \tag{4}$$

If the data set contains massive data, the calculation will be too complex. A simplified way is used to replace the mean and variance of the whole data set with the mean and variance of a batch, which greatly reduces the amount of calculation.

The calculation formulas of the mean and variance of a batch are shown in the following equations, respectively:

$$\mu = \frac{1}{n} \sum_{i=1}^{n} x_i, \tag{5}$$

$$\sigma^2 = \frac{1}{n} \sum_{i=1}^{n} (x_i - \mu)^2, \tag{6}$$

with the mean and variance, the normalization operation can be carried out. The process is shown in formula (8):

Normalization operation:

$$x_i = \frac{x_i - \mu}{\sqrt{\sigma^2 + \varepsilon}}, \tag{7}$$

where μ is the mean and σ is the variance.

The output of Batch Normalization can be obtained according to equations (4) and (8):

$$y_i = \gamma x_i + \beta. \tag{8}$$

3.4. Sample Imbalance and Focal Loss Improvement. Sample imbalance is an important problem in the field of machine learning. This problem will cause minor samples to

drown in major samples and reduce the importance of minor samples. In practical classification problems, major samples are simple and easy to classify, while minor samples account for only a few, which are difficult to classify. Because of simple samples with dominant position, the loss of simple samples should be small, but the number is large, which makes a major contribution to the loss. The samples that are difficult to classify are easy to be ignored by the model. These uncertain samples often lead to overfitting of the network, nonconvergence of the network in the initial stage, and damage to the model learning useful information.

In the expression recognition task, Cross Entropy (CE) is the commonly used loss measurement function, as follows:

$$\mathrm{CE} = -\ln\ p_i, \tag{9}$$

$$p_i = \frac{\exp(a_i)}{\sum_{j=1}^{7}(a_i)}, \tag{10}$$

where p_i is the probability of the label corresponding to the model prediction result and a_i represents the output of the neuron corresponding to the Softmax layer.

Focus Loss (FL) function reduces the weight of the samples easy to classify, making the model focus more on the samples difficult to classify during training. FL formula is as follows:

$$\mathrm{FL} = -\alpha(1 - p_i)\gamma\ \ln\ p_i. \tag{11}$$

Among them, the role of balance parameter α is to control the weight of unbalanced samples to the total loss and balance the number of samples in different categories. Focusing parameter γ is used to control the weights of the samples easy to classify and the samples difficult to classify.

In the calculation of FL, if there are some errors in the sample label of the data set, or the noise itself is very large, the model will learn the wrong information and reduce the accuracy of the model due to the increase of weight. Aiming at the problem that FL cannot deal with mislabeled samples, the threshold judgment is set through the confidence of samples and real labels, and the mislabeled samples are selected to change their confidence, so as to reduce FL's attention to such samples and improve the classification performance of the model.

$$\begin{cases} p_{\mathrm{top}} = \varepsilon, p_{\mathrm{top}} > c, \text{ and } y_p \neq y_t, \\ \mathrm{FL}, \text{ others}, \end{cases} \tag{12}$$

where p_{top} is the maximum probability that the prediction is true in several types of samples, the superparameter c is the probability threshold, y_t is the real label of the sample, and y_p is the prediction label of the sample.

If the maximum probability p_{top} mapped by the sample is greater than this threshold c, it is considered that the confidence of the sample is very high, and then the sample prediction label is compared with the real label. If the comparison shows that the sample prediction label is equal to its real label, it indicates that the sample is a simple sample with high confidence, and execute Focal Loss. If the sample prediction label is not equal to its real label, it indicates that the sample is a mislabeled sample with high confidence. Set its prediction probability to the minimum; that is, discard the sample. Aiming at the problem of mislabeled samples, this algorithm improves Focal Loss by setting threshold parameters to judge the discrete probability of Softmax output. It selects and discards mislabeled samples and improves the classification performance of the model. The flow chart of the improved Focal Loss algorithm is shown in Figure 4.

4. Design of Intelligent Interactive System for Service Robot Based on Expression Recognition

In the design of service robot human interaction system, if robot emotional interaction with people is achieved, it is a key link for robot to perceive and understand human emotional expression. Human emotional expression mainly includes expression, language, and body action. This paper mainly uses the way of recognizing facial expression to make the robot understand human emotional expression and give corresponding interactive feedback. The intelligent interactive system of service robot designed in this paper mainly consists of expression recognition, voice interaction, action interaction, and so on. The overall design block diagram of the interactive system is shown in Figure 5.

This paper adopts a facial expression recognition method based on deep learning. In the process of expression recognition, the most important step is to extract expression features, so after detecting the face, locating and tracking the key feature points of the face should be done, so as to extract the effective features representing the changes of facial expression. When recognizing face information, it is necessary to preprocess the image before detecting face. In feature extraction, four different convolution kernels are used for multichannel convolution to replace the third convolution layer of classical AlexNet network, which can extract picture features in a higher dimension. Sample imbalance is an important problem in the field of machine learning. Using the Focus Loss function to reduce the weight of the samples easy to classify makes the model focus more on the samples difficult to classify in training. The voice interaction module NAOqiAPIs provides rich application module interfaces, and voice-related application modules are provided by the Audio module. The main modules and functions provided by Audio are shown in Table 1, which implements recognition of the speaker's emotion through speech. ALDialog and ALTextToSpeech modules are used to create a knowledge base to make speech interaction strategies according to different expression categories and convert the text knowledge base into speech. The Motion application module provided by NAOqiAPIs can be used to set various poses and actions of the robot. The main modules and functions provided by Motion module are shown in Table 2.

5. Experiment and Analysis

5.1. Experimental Environment and Data Set. FER2013 facial expression data set consists of 35886 facial expression images, including 28708 training images, 3589 public

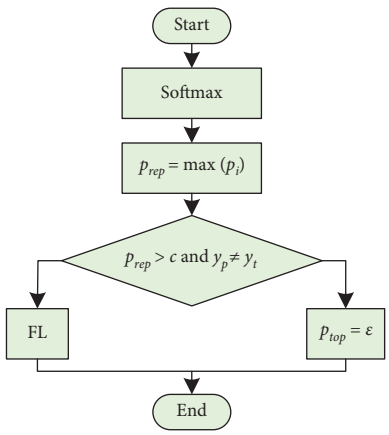


Figure 4: Flow chart of improved FL algorithm.

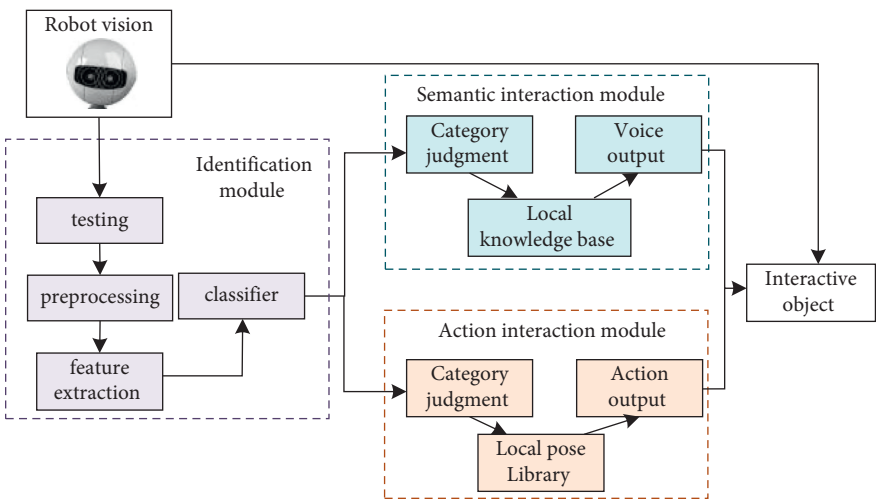


Figure 5: Overall block diagram of interactive system.

Table 1: Audio module function introduction.

AL audio player and AL audio recorder	Audio recording and playback
AL sound detection and AL sound localization	Sound detection and student source location
AL speech recognition	Speech recognition
AL text to speech	Text to speech
AL dialog	Create a basic knowledge base for communication dialogue
AL voice emotion analysis	Speaker's emotion through speech recognition

Table 2: Motion module function introduction.

AL robot posture	Provides a predefined pose for the robot
AL navigation	Make the robot move safely and stop when an obstacle is detected
AL motion	Write your own way of movement according to different methods

validation images, and 3589 private validation images. Each image is a grayscale image with fixed size of $48 * 48$. There are 7 kinds of expressions in the data set, which, respectively, correspond to digital labels 0–6. The labels corresponding to specific expressions are as follows: 0: Angry, 1: Disgust, 2: Fear, 3: Happy, 4: Sad, 5: Surprised, and 6: Neutral. Figure 6 shows an example of FER2013 expression database.

5.2. Curve of Accuracy and Loss Value. The batch size of the training network is set to 64, the epoch of training is set to 5000 steps, the image size of facial expression is $48 * 48$, and the starting value of learning rate is 0.001. The network is optimized by Adam optimizer.

Figure 7 is the accuracy change diagram of the facial expression experiment based on improved AlexNet network, and Figure 8 is the loss change diagram of the facial

Figure 6: Example of FER2013 expression database.

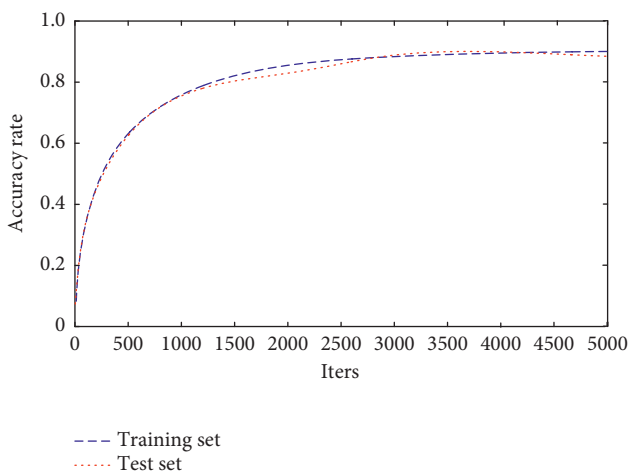


Figure 7: Experimental accuracy change diagram of facial expression recognition based on improved AlexNet network.

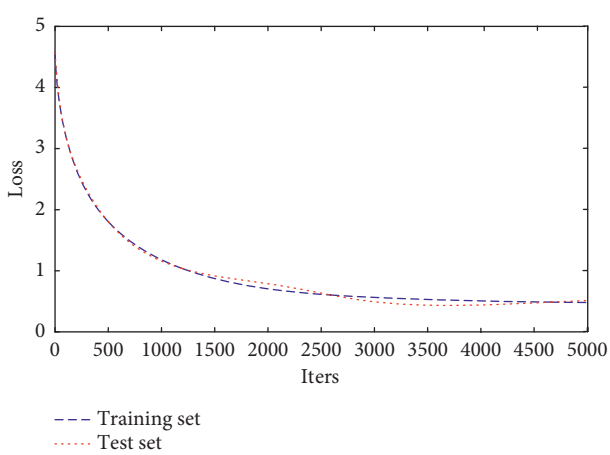


Figure 8: Experimental loss change diagram of facial expression recognition based on improved AlexNet network.

expression experiment based on improved AlexNet network. The experiment took about 13 hours. As can be seen from Figure 7, the curve changes obviously before step 2000, and the accuracy curve of the experiment rises rapidly. After step 2000, the change is slow and finally tends to converge. In step 3000, the accuracy rate of the training set is 90.26% and that of the test set is 89.30%. As shown in Figure 8, the value of the loss function decreases gradually with the increase of the number of training times. After reaching a certain number of training times, the loss value tends to be constant.

5.3. Identification Accuracy Comparison and Confusion Matrix. In order to prove the accuracy of the proposed method, [27] and [28] are selected for comparison. The comparison result is shown in Table 3. The accuracy of the proposed algorithm reaches 98.36%, while the accuracy of the comparative algorithm in [27, 28] reaches 96.69% and 97.17%, respectively. It is proved that the accuracy rate of the proposed method in this paper is better than the comparison algorithm on FER2013 data set. References [27, 28] only extract the edge detail features of facial expression, but lack feature processing, so the accuracy is low. The proposed method introduces the multichannel convolution technology and replaces the classical single-sized convolution kernel with four different-sized convolution kernels of $1 * 1$, $3 * 3$, $5 * 5$, and $7 * 7$. Compared with the single-sized convolution kernel, multichannel convolution strengthens the diversity of convolution kernel and can extract more features. In addition, Global Average Pooling and Batch Normalization are added. Therefore, the accuracy of facial expression recognition has been greatly improved by the improved AlexNet network.

In order to further verify the algorithm, the confusion matrix is drawn according to the experimental results on FER2013 data set. The column represents the predicted class, the row represents the real class, the diagonal value is the prediction accuracy of this class, and the rest is the probability of prediction error. It can be seen from the confusion matrix in Figure 9 that the classification results of the algorithm in this paper are evenly distributed, and all kinds of expression samples are more likely to be classified into their classes. Comparing the accuracy rates of seven kinds of expressions, it is found that the accuracy rates of Sad, Happy,

Table 3: Comparison of different algorithms.

Model	Accuracy rate (%)
The proposed method	98.36
Reference [27]	96.69
Reference [28]	97.17

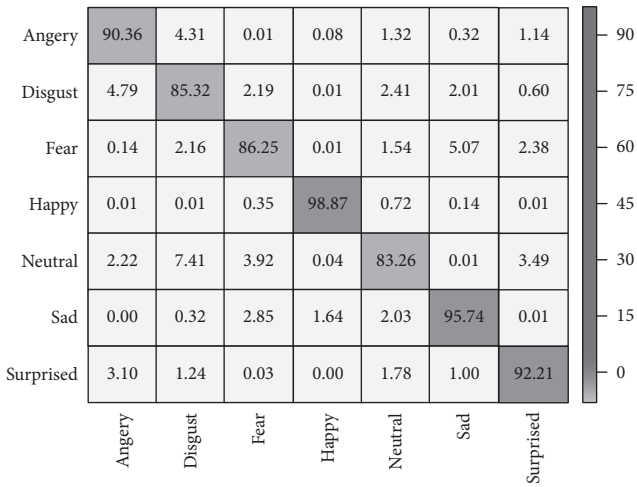


Figure 9: Confusion matrix of FER2013.

and Surprised are higher, the accuracy rate of Neutral is the lowest, and they are easy to be disturbed by Disgust expressions. It also verifies the similarity of the feature map of the two expressions. The feature separation is not obvious in pixel space, and it is also in line with human cognition. This is because this paper improves the Focal Loss by setting the probability threshold to change its confidence, so as to reduce FL's attention to this kind of samples and improve the classification performance of the model.

5.4. Interaction Effect with Robot. After the Facial Expression Rec instruction box is executed, the Switch Case instruction box distributes the expression recognition results to the eight Animated Say instruction box corresponding to the eight expression classes. The Animated Say instruction box stores predefined interaction languages and actions. For example, the expression class of the recognition interaction object is Happy, and the voice text stored in the instruction box is "You look happy, can you share it with me." The robot will make an action while talking. The expression class is Surprised, and the voice text stored in the instruction box is "Why are you so surprised? What's wrong with me." Similarly, the robot makes an action while talking. In the process of robot local test, the processing time of 200 facial expressions by the system is shown in Figure 10. Because the processor performance of the robot is worse than that of the PC, the average time for expression recognition is about 1.24 s, which is worse than that of the PC. Although the robot can detect and track the face of interactive object, the illumination of the face photographed by the camera changes greatly with the rotation of the face, which is a main reason for incorrect expression recognition.

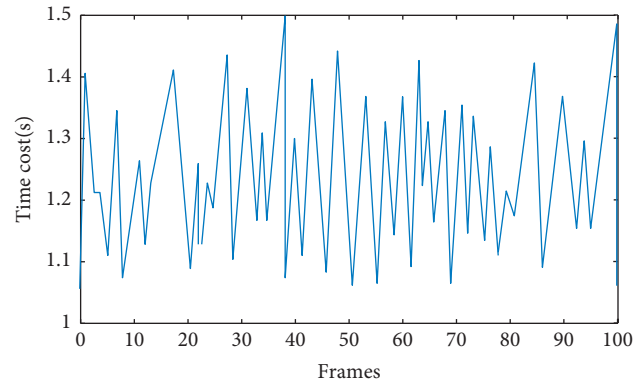


Figure 10: System running time on Robot.

6. Conclusion

Based on the above statement, this paper proposes an expression recognition and robot intelligent interaction method using deep learning. The proposed method uses four different convolution kernels for multichannel convolution to implement feature extraction to a greater extent, replaces fully connected layer with Global Average Pooling layer to avoid the problem of too many parameters, and improves Focal Loss by changing its confidence and improves the classification performance of the model. The experimental results show that the model improves the accuracy and can recognize facial expression more accurately.

In the design of robot intelligent interaction system, in order to make the robot interact with people more accurately, context analysis can be carried out in the future, so that the robot can provide more suitable interactive feedback information according to the context.

Data Availability

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

Conflicts of Interest

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

Acknowledgments

This work was supported by the Natural Science Foundation of Heilongjiang Province (no. LH2021F029) and Higher Education Reform Project of Heilongjiang Province (no. SJGY20200311).

References

[1] S. Kwak, K. San, and J. Choi, "Can robots be sold? The effects of robot designs on the consumers' acceptance of robots," in *Proceedings of the 2014 9th ACM/IEEE International Conference on Human-Robot Interaction*, pp. 220-221, IEEE, Sapporo, Japan, March 2014.

[2] E. Kwon and G. J. Kim, "Humanoid robot vs. projector robot: exploring an indirect approach to human robot interaction,"

in *Proceedings of the 5th ACM/IEEE International Conference on Human-Robot Interaction*, pp. 157-158, Osaka, Japan, March 2010.

[3] K. Haring, D. Silvera, and T. Takahashi, "How people perceive different robot types: a direct comparison of an android, humanoid, and non-biomimetic robot," in *Proceedings of the 2016 8th International Conference on Knowledge and Smart Technology*, pp. 265–270, IEEE, Chiang Mai, Thailand, February 2016.

[4] A. Ajoudani, A. M. Zanchettin, S. Ivaldi, A. Albu-Schäffer, K. Kosuge, and O. Khatib, "Progress and prospects of the human-robot collaboration," *Autonomous Robots*, vol. 42, no. 5, pp. 957–975, 2018.

[5] N. Wang, Y. Zeng, and J. Geng, "A brief review on safety strategies of physical human-robot interaction," in *Proceedings of the ITM Web of Conferences*, pp. 155–163, EDP Sciences, Moscow, Russia, November 2019.

[6] R. Liu and X. Zhang, "A review of methodologies for natural-language-facilitated human–robot cooperation," *International Journal of Advanced Robotic Systems*, vol. 16, no. 3, pp. 1729–1736, 2019.

[7] A. R. Wagner, P. Robinette, and A. Howard, "Modeling the human-robot trust phenomenon," *ACM Transactions on Interactive Intelligent Systems*, vol. 8, no. 4, pp. 1–24, 2018.

[8] M. Awais, M. Y. Saeed, and M. S. A. Malik, "Intention based comparative analysis of human-robot interaction," *IEEE Access*, vol. 12, no. 5, pp. 1996–2008, 2020.

[9] N. Mavridis, "A review of verbal and non-verbal human-robot interactive communication," *Robotics and Autonomous Systems*, vol. 63, no. 1, pp. 165–175, 2014.

[10] T. Xue, W. Wang, and J. Ma, "Progress and prospects of multimodal fusion methods in physical human–robot interaction: a review," *IEEE Sensors Journal*, vol. 2, no. 9, pp. 10–19, 2020.

[11] X. Liu, S. Ge, and F. Zhao, "A dynamic behavior control framework for physical human-robot interaction," *Journal of Intelligent and Robotic Systems*, vol. 101, no. 1, pp. 1–18, 2021.

[12] D.-X. Zhou, "Deep distributed convolutional neural networks: Universality," *Analysis and Applications*, vol. 16, no. 6, pp. 895–919, 2018.

[13] H. Patel, A. Thakkar, M. Pandya, and K. Makwana, "Neural network with deep learning architectures," *Journal of Information and Optimization Sciences*, vol. 39, no. 1, pp. 31–38, 2018.

[14] W. Luo, J. Lu, X. Li, L. Chen, and K. Liu, "Rethinking motivation of deep neural architectures," *IEEE Circuits and Systems Magazine*, vol. 20, no. 4, pp. 65–76, 2020.

[15] A. Shrestha and A. Mahmood, "Review of deep learning algorithms and architectures," *IEEE Access*, vol. 7, no. 6, pp. 53040–53065, 2019.

[16] X. Zhao and S. Zhang, "A review on facial expression recognition: feature extraction and classification," *IETE Technical Review*, vol. 33, no. 5, pp. 505–517, 2016.

[17] N. V. K. Medathati, H. Neumann, G. S. Masson, and P. Kornprobst, "Bio-inspired computer vision: towards a synergistic approach of artificial and biological vision," *Computer Vision and Image Understanding*, vol. 150, no. 2, pp. 1–30, 2016.

[18] S. Alyamkin, M. Ardi, A. C. Berg et al., "Low-power computer vision: status, challenges, and opportunities," *IEEE Journal on Emerging and Selected Topics in Circuits and Systems*, vol. 9, no. 2, pp. 411–421, 2019.

[19] A. Voulodimos, N. Doulamis, and A. Doulamis, "Deep learning for computer vision: a brief review," *Computational Intelligence and Neuroscience*, vol. 8, no. 5, pp. 126–175, 2018.

[20] A. Gupta, "Current research opportunities for image processing and computer vision," *Computer Science*, vol. 20, no. 3, pp. 15–26, 2019.

[21] C. Xu, Y. Cui, Y. Zhang, P. Gao, and J. Xu, "Person-independent facial expression recognition method based on improved Wasserstein generative adversarial networks in combination with identity aware," *Multimedia Systems*, vol. 26, no. 1, pp. 53–61, 2020.

[22] H. Zhang, Z. Qu, and L. Yuan, "A face recognition method based on LBP feature for CNN," in *Proceedings of the 2017 IEEE 2nd Advanced Information Technology, Electronic and Automation Control Conference*, pp. 544–547, IEEE, Chongqing, China, March 2017.

[23] P. Dhankhar, "ResNet-50 and VGG-16 for recognizing facial emotions," *International Journal of Innovations in Engineering and Technology*, vol. 13, no. 4, pp. 126–130, 2019.

[24] J. Hu, L. Shen, and G. Sun, "Squeeze-and-excitation networks," *IEEE*, in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 7132–7141, Salt Lake City, UT, USA, June 2018.

[25] S. Xie, H. Hu, and Y. Wu, "Deep multi-path convolutional neural network joint with salient region attention for facial expression recognition," *Pattern Recognition*, vol. 92, no. 2, pp. 177–191, 2019.

[26] Y. Li, J. Zeng, and S. Shan, "Occlusion aware facial expression recognition using CNN with attention mechanism," *IEEE Transactions on Image Processing*, vol. 28, no. 5, pp. 2439–2450, 2018.

[27] Y. Yaermaimaiti, "Facial expression recognition based on local feature and deep belief network," *Journal of Decision Systems*, vol. 3, no. 15, pp. 1–13, 2021.

[28] K. Talele and K. Tuckley, "Facial expression recognition using digital signature feature descriptor," *Signal, Image and Video Processing*, vol. 14, no. 4, pp. 701–709, 2020.

Hindawi
Journal of Robotics
Volume 2022, Article ID 7778592, 8 pages
https://doi.org/10.1155/2022/7778592

Research Article

Personalized Travel Route Recommendation Model of Intelligent Service Robot Using Deep Learning in Big Data Environment

Xiang Huang

Hunan Mass Media Vocational and Technical College, Changsha, Hunan 410100, China

Correspondence should be addressed to Xiang Huang; 332374935@qq.com

Received 8 December 2021; Accepted 15 January 2022; Published 29 January 2022

Academic Editor: Shan Zhong

Aiming at the problems that the traditional model is difficult to extract information features, difficult to learn deep knowledge, and cannot automatically and effectively obtain features, which leads to the problem of low recommendation accuracy, this paper proposes a personalized tourism route recommendation model of intelligent service robot using deep learning in a big data environment. Firstly, by crawling the relevant website data, obtain the basic information data and comment the text data of tourism service items, as well as the basic information data, and comment the text data of users and preprocess them, such as data cleaning. Then, a neural network model based on the self-attention mechanism is proposed, in which the data features are obtained by the Gaussian kernel function and node2vec model, and the self-attention mechanism is used to capture the long-term and short-term preferences of users. Finally, the processed data is input into the trained recommendation model to generate a personalized tourism route recommendation scheme. The experimental analysis of the proposed model based on Pytorch deep learning framework shows that its Pre@10, Rec@10 values are 88% and 83%, respectively, and the mean square error is 1.537, which are better than other comparison models and closer to the real tourist route of the tourists.

1. Introduction

With the continuous improvement of social living standards, people's demand for tourism and leisure is also increasing year by year. The development and prosperity of the tourism industry have made going out to travel increasingly popular. In the preparation of outbound travel, the route formulation in the travel strategy is an extremely important and critical step [1]. People can check relevant travel guides or collect relevant information on the internet, however, it will waste time and be inefficient. Moreover, the information they find will not match their current needs [2]. Therefore, it is difficult to find information that fits the purpose, and systematically, there will be big differences between the information provided by different people. In addition, the rapid increase in the number of users of social networks has caused a rapid increase in network information. When users face massive network data, they cannot realize a quick selection of information [3, 4]. Therefore, the author hopes to be able to automatically obtain personalized travel recommendations that meet their specific needs to help users quickly filter useless information in a large amount of travel information and improve the efficiency and comfort of users in integrating information [5]. The personalized route recommendation platform is diversified. Among them, intelligent service robots in scenic spots occupy an important position among many platforms because of their high efficiency and convenience. Therefore, its recommended algorithm model is also very important.

The purpose of personalized travel route recommendation is to recommend a travel route composed of multiple points of interest (POI) based on the user's personalized interests and the user's own travel restrictions. In the personalized recommendation of POI, the key is to comprehensively model the user interests and similarities between the users. A comprehensive analysis of the user's personalized preference for different POIs and the degree of similarity correlation between each POI is used to determine the user's personalized interest [6]. In addition, similarity matching is used to find the degree of association between

the users. By the analysis of similar users and similar mobile patterns, personalized POI recommendations are made to users. At the same time, personalized travel route recommendation also needs to take into account the user's personalized factors and generate a travel route that meets their own travel restrictions for each user [7].

There have been many pieces of research on personalized travel route recommendation at home and abroad. Reference [8] introduced the evolution process of the travel recommendation system in detail and conducted research on its characteristics and current limitations. At the same time, the key algorithms used in the classification and recommendation process and the indicators that can be used to evaluate the performance of the algorithms are also discussed. In terms of recommendation methods, the most common method is the recommendation technology based on collaborative filtering. Realize the personalized recommendation by mining the similarity between the users. Reference [9] uses the sequential pattern mining algorithm to generate various fine-grained candidate POI routes from POI access sequences to realize the recommendation of the best tourism route. However, the overall recommendation efficiency is low, and it is slightly insufficient for route planning with complex and multiple points of interest. Reference [10] proposed a personalized and content-adaptive cultural heritage route recommendation to achieve the best cultural heritage experience of context-aware routes. Use the first-order Markov model to convert motion as the time of the problem to realize route recommendation. The overall recommendation effect is good, however, it takes a long time and is not suitable for immediate recommendation. Reference [11] proposed a new travel route mining method on the basis of considering the theme level and characteristics of scenic spots, in which the scenic spots are subject layered according to the location information of the popular scenic spots. The travel path data set is constructed, and the travel routes are mined in combination with the subject level, however, there are still deficiencies in the consideration of user interest point matching.

With the rapid increase in the demand for tourist route recommendation in recent years, thanks to the rapid advancement of computer technology and communication technology, machine learning technology has been widely used in the field of automatic recommendation of demand [2]. Among them, deep learning algorithms have achieved excellent results in many fields. It can effectively process unstructured multimedia data. Some scholars have begun to try to use convolutional neural networks to solve the feature engineering problems faced by the recommender systems [12]. Reference [13] proposed a matrix factorization algorithm based on two-stage clustering. Using the social network subgraph integrated with preference similarity scores, combined with geographic spatial influence, the cluster refinement of preference embeddings is extended to the cluster refinement of geographic preference embeddings. In this way, the best route recommendation under complex conditions is realized. Reference [14] proposed a personalized travel recommendation scheme based on a weighted multi-information constraint matrix factorization scheme. However, the accuracy rate of the optimal travel route needs to be improved. Reference [15] proposed a route recommendation method based on interest topic and distance matching, which obtains the best travel path by analyzing the user's real historical travel footprint and scenic spot residence time and combined with the given travel time limit. However, this method has poor timeliness and adaptability and cannot be applied to the independent intelligent robot platform.

Based on the above analysis, the traditional tourism route recommendation model is difficult to pay attention to the long-term preferences of users and the poor recommendation effect caused by sparse data. This paper proposes a personalized travel route recommendation method for intelligent service robots using deep learning in a big data environment. It can be applied to intelligent service robots placed in the halls of scenic spots to realize a personalized travel route recommendation. To fuse multisource heterogeneous data, the proposed model uses the Gaussian kernel function, node2vec model, and other technologies to construct the embedded representations of users, time, space, POI score, access frequency, and social relationships. Send it to the deep learning network for analysis. It solves the problem of low recommendation accuracy caused by sparse data. The experimental results based on the pytoch deep learning framework show that the proposed model integrates user preference characteristics, geographical factor characteristics, and theme factor characteristics and can better complete tourism route recommendation, with a Pre@5 of 95%.

2. POI Recommended Problem Description

The user "sign-in" record data in social networks contains a large amount of high-value information data about POI and user preferences, which provides an opportunity for in-depth research on personalized POI recommendations. However, in practical applications, there are some personalized differences in users' preferences for POI categories [16]. The existing POI recommendation methods are mostly implemented by content-based or model-based collaborative filtering technology. The subject of POI and the relationship between the subjects are not fully considered. Therefore, in the user's personalized recommendation, combined with the theme factors of POI, more effective features are obtained from the limited user access information, and appropriate models are selected to achieve distinguishable user preference modeling. These are the keys to improve the effectiveness of personalized POI recommendations [17].

Finding effective features from the check-in data is the key to improve the quality of POI recommendations. Traditional methods only learn the linear or low-order interaction between the features, and they cannot effectively integrate the features in a location-based social network (LBSN) [18]. In recent years, with the rapid development of deep learning, it can intelligently learn high-order characteristics and interact from the input of specific tasks. Therefore, a deep neural network recommendation

framework that combines the DNN network with the LDA topic model and matrix factorization algorithm is proposed, named DLM. The user preference feature, geographic factor feature, and probability topic feature in LBSN are integrated into the POI recommendation task using word-embedding technology. High-level interactions between the features are learned through neural networks, and personalized recommendations are made to users.

3. Proposed Model

3.1. Overall Framework. The proposed implementation framework of the recommended model is mainly divided into four modules. They are data preprocessing, deep prediction model construction, network training, and final recommendation list generation. The main implementation framework is shown in Figure 1. The model preprocesses the acquired data and uses the Gaussian kernel function and node2vec model to model it to obtain the corresponding POI location and social relationship embedded representation. Both of them are input into the self-attention module to capture user preferences to obtain an ideal personalized tourism recommendation scheme.

The process of data acquisition and preprocessing is very complex. Mainly by crawling the relevant website data, the basic information data, the comment text data of travel service items, and the basic information data and comment text data of users are obtained. Then, these data are preprocessed. Perform data cleaning on the crawled data to filter out the incomplete data and junk data.

3.2. Construction of Deep Prediction Models. The construction of in-depth prediction models. Neural network technology is mainly used to construct a network model and process the preprocessed data. Use the feature extraction ability of deep learning to obtain the corresponding features, and use the model to predict the user's rating of the tourism service item [19].

Train the network. For the constructed model network, use the training sample data for supervised network training. Mine the potential factors between the users and tourism service items and learn the expression of the interaction relationship between the users and tourism service items to train the model.

Generate a personalized recommendation list. Test the experimental data and input the experimental data into the trained model. The model predicts the user's rating of travel service items and sorts them according to the size of the rating. Generate a personalized recommendation list for each user to complete the user's recommendation.

3.3. Data Acquisition and Data Preprocessing

3.3.1. Data Collection. The collected data mainly includes two parts.

The first part is the user's basic information and the comment text data. The user's basic information data mainly includes the user's gender, age, occupation, city, and

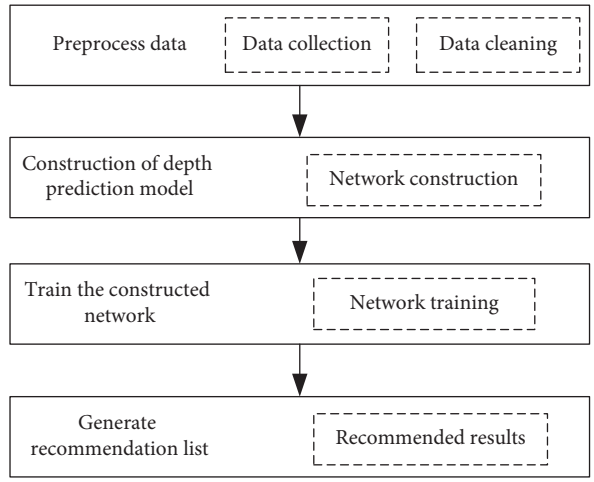


Figure 1: Main implementation framework based on depth prediction model recommendation.

historical comment items. The comment data mainly includes the user's comment text information on past travel service items and corresponding scores. This part of the content is mainly used to extract user behavior characteristics, analyze user preferences, and build user characteristic models.

The second part is the basic information of tourism service items and the comment data of tourism service items. The basic information of the travel service item includes the name, location, and label of the travel service item. The comment data mainly refers to the comment text information and the corresponding score obtained by the tourism service item. This part of the data is used to extract the attribute characteristics of the tourism service items and construct the characteristic model of the tourism service items.

To obtain these data, web crawlers are used to crawl the related travel websites. The crawler uses the Scrapy web crawler framework to crawl the website. Scrapy is a distributed crawler framework based on Python. Scrapy is highly flexible and controllable and can easily implement distributed crawlers. At the same time, Scrapy encapsulates the implementation details of a lot of crawlers, which can focus more on data extraction.

3.3.2. Data Cleaning. To ensure that the recommended results are valid, the data should be complete and reliable. Therefore, the crawled data must first be cleaned up to filter out the incomplete data and junk data. During data cleaning, the steps followed will be as follows:

(1) Firstly, filter out the users with incomplete basic information. Incomplete basic information means that the basic information characteristics cannot be found. For tourism service items and users, it is impossible to dig out the characteristic influence of its basic information. Therefore, ensuring the integrity of basic information plays an important role in model building.

(2) Spam comments need to be filtered. By observing the scraped comment data, it can be found that the general comment has only one word, and words that do not indicate good or bad mood can be filtered out as spam comments as these data have no positive effect on the establishment of the model.

(3) To filter the content of the comments, filter out the special symbols in the comments. These symbols are not helpful in digging out the characteristics of the comment content. After the data cleaning is completed, complete and reliable data is obtained. These data will be further processed.

3.4. Network Building. The proposed model contains two components: information embedding and information interaction. Its structure is shown in Figure 2. For the information embedding module, firstly, the input data is multi-hot coded, and the user POI check-in sequence model is constructed to generate the potential representation matrix. Then, in the auxiliary information extraction part, POI geographic location information is extracted by the Gaussian kernel function, and the data information is normalized by softmax. In the information interaction module, the deeper interaction of data is obtained based on self-attention to learn the long-term and short-term preferences of the users, and the information is fused using three bottleneck layers to obtain the final prediction.

3.4.1. User Representation Embed. The purpose of a given user's check-in record is to learn the potential representation of the POI sequence to improve the accuracy of POI recommendations. In the proposed model, a transition matrix is designed. In this way, the user's spatiotemporal intentions and the potential characteristics of POI are mapped in the feature space. The input is a user sign-in vector $\mathbf{P}_i^u = \left\{p_{t_1}^u, p_{t_2}^u, \ldots, p_{t_{i-1}}^u\right\}$ characterized by multi-hot. When $p_{t_k}^u$ is 1, it means that user u has visited POI_p at time t_k. The mapping process is as follows:

$$\boldsymbol{\varphi}^u = f_1\left(\omega_u p_i^u + b_u\right), \tag{1}$$

where φ^u represents the latent representation vector of user u. ω_u and b_u are the weight and bias of user u, respectively.

3.4.2. Spatio-Temporal Information Embedding. In the user's check-in record, the user's behavior is usually limited to several specific areas, which is a well-known geographical cluster phenomenon in the user's check-in activities [20]. From this phenomenon, it can be inferred that the users prefer to visit the unreachable POI near the POI they have visited before, and users' liking depends on the attributes and distance of the two POIs [21, 22]. If the attributes are the same, then the closer the distance between the two POIs, the greater the possibility that the user will visit the unvisited POI [23]. To combine the geographical distance attribute of the POI, the Gaussian kernel function is used to extract the neighbor perception influence of the sign-in POI, which is expressed as follows:

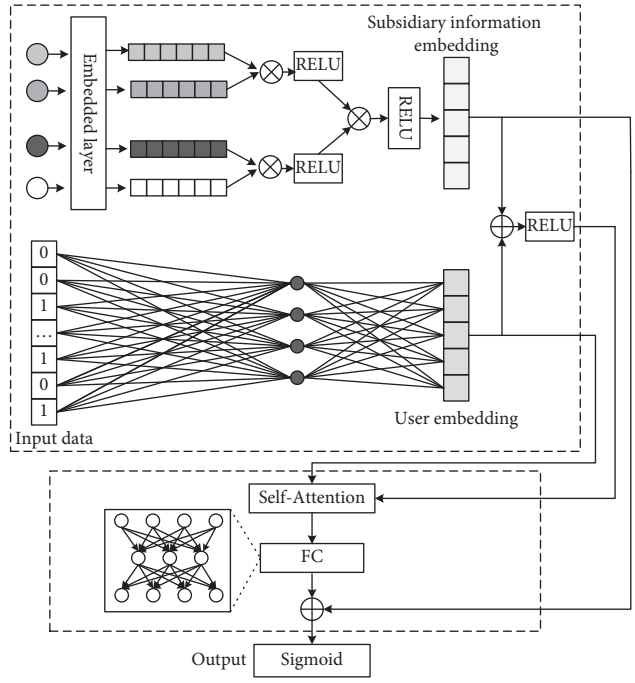


Figure 2: Overall framework of the proposed method.

$$\boldsymbol{\kappa}\left(L_{t_i}, L_{t_j}\right) = \exp\left(-\frac{\left\|L_{t_i} - L_{t_j}\right\|^2}{2\sigma^2}\right). \tag{2}$$

Among them, L_{t_i} and L_{t_j} are the geographic coordinates of the two POIs visited by the user. The value range of the Gaussian kernel is $\kappa(L_{t_i}, L_{t_j}) \in [0, 1]$. σ is the bandwidth, which controls the radial range of action. Finally, by calculating the paired Gaussian kernel value of each POI pair, the Gaussian kernel value vector $\kappa \in R^N$ can be obtained.

3.4.3. POI Score and Access Frequency Embedding. By preprocessing the LBSN data set, users' ratings and access frequency can be obtained. Use normalization to scale its value to (0, 1). The sum of all check-in score probabilities and access frequency probabilities of each user is 1, respectively. Therefore, it is easier to characterize the user's preference for the POI they have visited, which is expressed as follows:

$$\begin{aligned} \phi_r\left(L_j|L_i\right) &= \frac{\exp\left(L_i\right)}{\sum_{j=1}^{N}\exp\left(L_j\right)}, \\ \phi_q\left(L_j|L_i\right) &= \frac{\exp\left(L_i\right)}{\sum_{j=1}^{N}\exp\left(L_j\right)}. \end{aligned} \tag{3}$$

By calculating the score probability value $\phi_r(L_j|L_i)$ and the access frequency probability value $\phi_q(L_j|L_i)$ of each user separately, the probability value vectors $\varphi_r \in R^N$ and $\varphi_q \in R^N$ can be obtained.

3.4.4. User Social Relationship Embedding. Select the node2vec model to extract the user relationship features. For each user, use the node2vec model to generate m random walk sequences of length n. Then, train by Skip-Gram with hierarchical Softmax. Finally, find users related to the current user, so that a latent representation matrix $\mathbf{s} \in R^N$ of the user's social relationship can be obtained.

3.4.5. Attached Information Display. To handle the complex interaction between the user and POI and further POI recommendations, the various auxiliary information is integrated. It is expressed as follows:

$$\boldsymbol{\varphi} = f_2\left(\mathrm{ReLU}\left(\boldsymbol{\kappa} \odot \boldsymbol{\varphi}_r\right) \odot \mathrm{ReLU}\left(\boldsymbol{\varphi}_q \odot \mathbf{s}\right)\right), \tag{4}$$

where φ is the latent representation vector of ancillary information. $\odot$ is the element dot product.

3.5. Network Training. Use pairwise optimized Bayesian Personalized Ranking (BPR) loss function to learn model parameters. The objective function is as follows:

$$\mathrm{Loss} = \sum_{\left(u, L_i, L_j \in O\right)} -\ln\, \delta\left(\widehat{y}_{uL_i} - \widehat{y}_{uL_j}\right) + \lambda \|\Phi\|_2^2, \tag{5}$$

where $O = \left\{(u, L_i, L_j) | (u, L_i) \in E^+, (u, L_j) \in E^-\right\}$ represents the paired training data. E^+ indicates that there is a record of access records. E^- Q represents an unobserved visit record. $\widehat{y}_{uL_i}$ and $\widehat{y}_{uL_j}$, respectively, represent the user's preference for the target POI L_i and L_j. $\delta(\cdot)$ is the sigmoid function. $\Phi = \left\{\varphi, \{\omega^c\}_{c=1}^{C}\right\}$ represents all model parameters that can be trained. λ represents the L_2 regularization parameter that controls overfitting. The small batch Adam optimization algorithm is used to optimize the prediction model and update the model parameters.

Although deep learning models have strong representation capabilities, they often have the problem of overexpression [24]. Dropout is an effective solution to prevent neural networks from overfitting. To make the model generalize well to unobserved data, Dropout is used in training. Dropout randomly deletes specific nodes and the neural network nodes of the information self-encoding layer with a certain probability.

3.6. Generate Recommendation List. After the network model is trained using the above optimization algorithm, the vector information of tourism service items and users can be input into the network model. Through the trained network model, the user's prediction score of the tourism service item is obtained. This value is used as a basis and arranged according to size. Recommend the top N travel service items with higher scores to the user, generate a personalized recommendation list, and complete the recommendation.

4. Experiment and Analysis

The data used in the experiment is from a tourism website, and the basic information data example of its users is shown in Table 1. There will be many comments in a scenic spot. Similarly, a user will also have multiple comments, and each comment has a corresponding user and scenic spot identity (ID) and corresponding score. Because of the limited size of the table, there is no detailed display here (the comments in the table only show the content of the comment text).

The proposed model is run through the pytorch 1.2.0 framework, where the dropout rate is set to 0.5.

4.1. Evaluating Indicator. Two broad indicators are used to evaluate the performance of different recommendation models, namely accuracy and recall (these two indicators are represented by Pre@N and Rec@N, respectively), and the calculation is as follows:

$$\begin{aligned} \mathrm{Pre@}N &= \frac{1}{\vartheta} \sum_{u \in U} \frac{\mathrm{Top} - N \cap K}{N}, \\ \mathrm{Rec@}N &= \frac{1}{\vartheta} \sum_{u \in U} \frac{\mathrm{Top} - N \cap K}{K}, \end{aligned} \tag{6}$$

where ϑ represents the number of users. N represents the number of recommended POIs. Top-N represents the list of the first N points of interest recommended by the recommendation model to the target user. K represents the real check-in list in the user test set, i.e., the POI set that the user has actually accessed in the actual historical access record.

In addition, for user u and tourism service item l in the test set, $\widehat{r}_{ul}$ represents the predicted score generated by the recommendation algorithm, r_{ul} represents the actual score of user u on item i, and hence, the mean square error (MSE) can be defined as follows:

$$\mathrm{MSE} = \frac{1}{G} \sum_{g=1}^{G} \left(r_{ul} - \widehat{r}_{ul}\right)^2, \tag{7}$$

where G is the number of observed values in the test set.

4.2. Performance Comparison with Comparison Algorithm

4.2.1. Influence Analysis of Model Characteristics. Since each feature in the recommendation model will have a certain impact on the recommendation results, user preferences (UP), geographic factor (GF), and thematic factor (TF) are successively added to the proposed model, and experiments are carried out. The comparison results under Pre@5, Pre@10, and Pre@20 are shown in Figure 3.

As can be seen from Figure 3, the proposed model integrates user preference features, geographical factor features, and topic factor features, and its recommendation accuracy is better than the model with user preference and geographical factors. Taking Pre@5 as an example, its accuracy is as high as 95%. At the same time, it can also be seen that the recommendation effect obtained by the fusion of the three-factor features is significantly better than that of the single factor features or the fusion of the two-factor features.

4.2.2. Comparative Analysis of Accuracy and Recall. In addition, by comparing the accuracy and recall of the proposed model with the models in reference [9, 11, 14] on

Table 1: Data examples of tourism service items.

ID	Name	Label	Commentary
i135728	Huashan	Huashan is the western mountain of the five mountains. It has a unique landscape and is known as the most dangerous mountain in the world.	. . .“Huashan is really worth going. It is super dangerous and the first mountain in the world” . . .
i126239	Lugu lake	Lugu lake, with its natural and primitive folk customs and beautiful natural scenery, is known as the “magical oriental daughter country”	. . .“A beautiful and mysterious place, one of the holy lakes worth visiting in your life” . . .
i135460	Forbidden city	The forbidden city is one of the largest and best preserved wooden ancient buildings in the world	. . .“The forbidden city after snow is really beautiful. I Highly recommend it!” . . .

the data set, the comparison results in the case of Pre@5, Pre@10, and Pre@20 are shown in Figure 4, and the comparison results in the case of Rec@5, Rec@10, and Rec@20 are shown in Figure 5.

It can be seen from Figures 4 and 5 that the accuracy and recall of the proposed model are significantly better than those of the other recommended models. Taking Pre@10 and Rec@10 as examples, their values are 88% and 83%, respectively, while those of the other models are less than 80%. The proposed model uses the Gaussian kernel function to obtain the pairwise distance between the corresponding POIs in the user check-in record, selects the node2vec model to extract the network structure characteristics of the user's social relationship, and captures the user's preferences by the self-attention mechanism. Hence, the overall recommendation effect is good. However, reference [9] uses sequential pattern mining algorithm to build the POI knowledge base and massive structured POI access sequence to realize the recommendation of the best tourism route, but there is no influence of geography and other factors. Therefore, the accuracy and recall rate of the recommendation scheme are low. Taking Pre@20 and Rec@20 as examples, both are less than 50%. Reference [11] formed a standardized travel data set by preprocessing the data, such as word segmentation and denoising, stratified the scenic spots according to the location information of popular scenic spots, and recommended travel routes in combination with the theme level and scenic spot characteristics, however, it did not deeply mine the users. Hence, the performance of the recommendation model was poor. Reference [14] proposed a personalized travel recommendation model based on weighted multi-information constraint matrix decomposition scheme, which comprehensively describes users and travel locations using photos, user access sequences, and text tags, and it allocates different weights in combination with the common access probability based on geographical distance, which can achieve better travel route recommendation. However, because of the traditional method, the recommendation performance is lower than that of the proposed model using deep learning. Taking Rec@5 as an example, it is reduced by 9%.

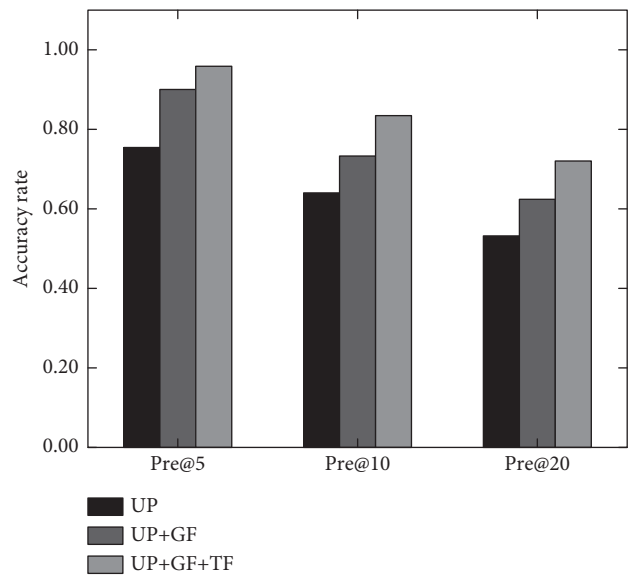


Figure 3: Comparison of @N results before characteristic factors of the proposed model.

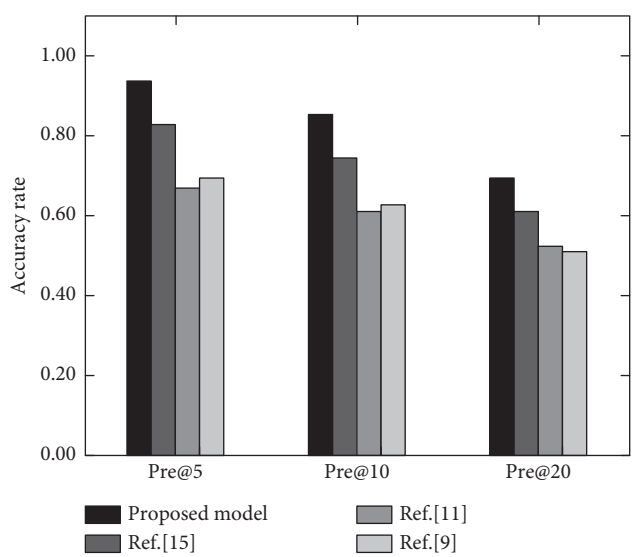


Figure 4: Comparison of accuracy of pre @N data sets.

4.2.3. MSE Analysis of Different Models. To demonstrate the recommended performance of the proposed model, it is compared with reference [9, 11, 14]. The results are shown in Table 2.

It can be seen from Table 2 that the MSE values of reference [9, 11] are almost the same, only 0.031. Because of the lack of in-depth analysis of geography, user preferences, and other factors, the recommended results deviate greatly

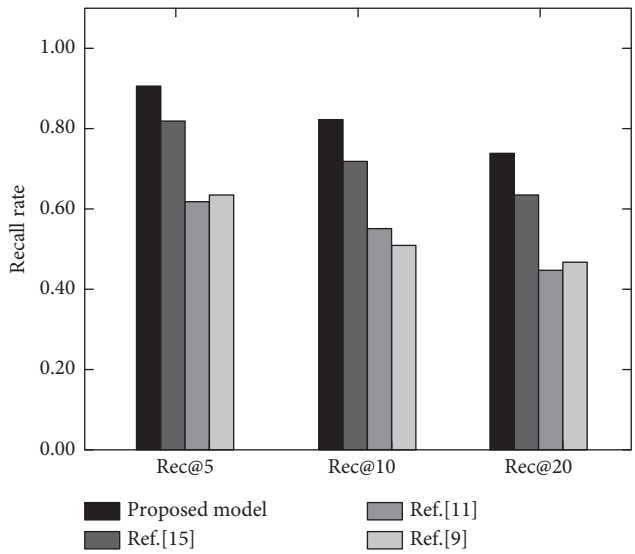


Figure 5: Comparison of accuracy of rec @N data sets.

Table 2: MSE comparison results of different models.

Model	MSE
Reference [9]	1.924
Reference [11]	1.893
Reference [14]	1.705
Proposed model	1.537

from the actual route. Reference [14] uses the weighted multi-information constraint matrix decomposition method to realize personalized travel recommendation, which considers many factors, however, it lacks a powerful learning algorithm. Therefore, the accuracy of the recommendation result is not high, and the MSE is 1.705. On the basis of preprocessing like data cleaning, the proposed model uses the deep learning algorithm to extract user features and carry out corresponding learning classification. It not only considers the comment text information of users and tourism service items but also adopts the basic information of users and tourism service items. Therefore, an ideal recommendation scheme is obtained, and its MSE is only 1.537. In conclusion, the above results demonstrate the effectiveness and superiority of the proposed recommendation model.

5. Conclusion

In recent years, with the popularization of the internet and the continuous development of information technology, people's demand for tourism is richer and more diverse. Effective and timely tourism service recommendation is of great significance to provide efficient and high-quality personalized tourism service recommendation. Therefore, based on the deep learning algorithm in the big data environment, a personalized tourism route recommendation model that can be applied to the intelligent service robot in the scenic hall is proposed. Aiming at the problems of long travel timespan of users and dynamic changes of preferences, the proposed model uses the self-attention mechanism module to filter the POI features related to the user's long-term preferences in each sequence, thereby improving the accuracy of travel route recommendation. The experimental results based on the pytoch deep learning framework show that the proposed model completes data feature extraction and prediction using a deep learning network based on a self-attention mechanism, and it comprehensively considers all kinds of data information. Therefore, its Pre@10 and Rec@10 values are 88% and 83%, respectively, and the mean square error is 1.537, which has certain advantages in the tourism route recommendation.

At present, the extensive use of the knowledge map makes it possible to extract potential interactive representations that human beings cannot notice through this technology to make an effective recommendation has become a hot issue in research. Data often have diverse and heterogeneous representations, such as the type of POI, the access time of POI, the traffic time of POI, the cost of POI, the location of POI, etc. Mining the attribute information of these entities using the knowledge map technology can make the recommendation system develop further.

Data Availability

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

Conflicts of Interest

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

References

[1] L. Cai, W. Wen, B. Wu, and X. Yang, "A coarse-to-fine user preferences prediction method for point-of-interest recommendation," *Neurocomputing*, vol. 422, no. 3, pp. 1–11, 2021.

[2] Z. Zhang, C. Zou, R. Ding, and Z. Chen, "VCG: exploiting visual contents and geographical influence for Point-of-Interest recommendation," *Neurocomputing*, vol. 357, no. 9, pp. 53–65, 2019.

[3] D. Yha, D. Bca, T. C. Jing, and Y. Zeng, "Privacy-preserving point-of-interest recommendation based on geographical and social influence," *Information Sciences*, vol. 543, no. 8, pp. 202–218, 2021.

[4] D. Yu, W. Wanyan, and D. Wang, "Leveraging contextual influence and user preferences for point-of-interest recommendation," *Multimedia Tools and Applications*, vol. 80, no. 8, pp. 1–15, 2021.

[5] L. Chang, W. Chen, J. Huang, and C. Bin, "Exploiting multi-attention network with contextual influence for point-of-interest recommendation," *Applied Intelligence*, vol. 51, no. 5, pp. 1–14, 2021.

[6] X. Xiong, S. Qiao, N. Han et al., "Where to go: an effective point-of-interest recommendation framework for heterogeneous social networks," *Neurocomputing*, vol. 373, no. 2, pp. 56–69, 2020.

[7] Y. Si, F. Zhang, and W. Liu, "An adaptive point-of-interest recommendation method for location-based social networks based on user activity and spatial features," *Knowledge-Based Systems*, vol. 163, no. 1, pp. 267–282, 2019.

[8] S. Renjith, A. Sreekumar, and M. Jathavedan, "An extensive study on the evolution of context-aware personalized travel

recommender systems," *Information Processing & Management*, vol. 57, no. 1, pp. 102078.1–102078.19, 2020.

[9] C. Bin, T. Gu, Y. Sun, and L. Chang, "A personalized POI route recommendation system based on heterogeneous tourism data and sequential pattern mining," *Multimedia Tools and Applications*, vol. 78, no. 24, pp. 35135–35156, 2019.

[10] G. Alexandridis, A. Chrysanthi, G. E. Tsekouras, and G. Caridakis, "Personalized and content adaptive cultural heritage path recommendation: an application to the Gournia and Çatalhöyük archaeological sites," *User Modeling and User-Adapted Interaction*, vol. 29, no. 1, pp. 201–238, 2019.

[11] S. Du, H. Zhang, H. Xu, J. Yang, and O. Tu, "To make the travel healthier: a new tourism personalized route recommendation algorithm," *Journal of Ambient Intelligence and Humanized Computing*, vol. 10, no. 9, pp. 3551–3562, 2019.

[12] K. Song, M. Ji, S. Park, and I.-C. Moon, "Hierarchical context enabled recurrent neural network for recommendation," *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 33, no. 7, pp. 4983–4991, 2019.

[13] L. R. Divyaa and N. Pervin, "Towards generating scalable personalized recommendations: integrating social trust, social bias, and geo-spatial clustering," *Decision Support Systems*, vol. 122, no. 7, pp. 113066.1–113066.17, 2019.

[14] D. Lyu, L. Chen, Z. Xu, and S. Yu, "Weighted multi-information constrained matrix factorization for personalized travel location recommendation based on geo-tagged photos," *Applied Intelligence*, vol. 50, no. 1, pp. 1–15, 2020.

[15] X. Cheng, "A travel route recommendation algorithm based on interest theme and distance matching," *EURASIP Journal on Applied Signal Processing*, vol. 2021, no. 1, pp. 1–10, 2021.

[16] L. A. Yan, A. Cyc, W. B. Ran, and C. S. Victor, "IMCRec: a multi-criteria framework for personalized point-of-interest recommendations," *Information Sciences*, vol. 483, no. 7, pp. 294–312, 2019.

[17] A. Expósito, S. Mancini, J. Brito, and J. A. Moreno, "A fuzzy GRASP for the tourist trip design with clustered POIs," *Expert Systems with Applications*, vol. 127, no. 8, pp. 210–227, 2019.

[18] C. Villavicencio, S. Schiaffino, J. Andres Diaz-Pace, and A. Monteserin, "Group recommender systems: a multi-agent solution," *Knowledge-Based Systems*, vol. 164, no. 1, pp. 436–458, 2019.

[19] R. M. D'Addio, R. S. Marinho, and M. G. Manzato, "Combining different metadata views for better recommendation accuracy," *Information Systems*, vol. 83, no. 7, pp. 1–12, 2019.

[20] S. A. Yu, B. Hs, L. A. Chao, and L. Yin, "LSVP: a visual based deep neural direction learning model for point-of-interest recommendation on sparse check-in data," *Neurocomputing*, vol. 446, pp. 204–210, 2021.

[21] D. Yu, K. Xu, D. Wang, and T. Yu, "Point-of-Interest recommendation based on user contextual behavior semantics," *International Journal of Software Engineering and Knowledge Engineering*, vol. 29, no. 11, pp. 1781–1799, 2019.

[22] Z. Huang, X. Lin, H. Liu, B. Zhang, Y. Chen, and Y. Tang, "Deep representation learning for location-based recommendation," *IEEE Transactions on Computational Social Systems*, vol. 7, no. 3, pp. 648–658, 2020.

[23] X. Wang and S. Kadolu, "Modeling uncertainty to improve personalized recommendations via Bayesian deep learning," *International Journal of Data Science and Analytics*, vol. 2, no. 3, pp. 1–11, 2021.

[24] X. Liu, C. Andris, and S. Rahimi, "Place niche and its regional variability: measuring spatial context patterns for points of interest with representation learning," *Computers, Environment and Urban Systems*, vol. 75, no. 5, pp. 146–160, 2019.

Hindawi
Journal of Robotics
Volume 2022, Article ID 7115551, 9 pages
https://doi.org/10.1155/2022/7115551

Research Article

Adaptive Image Denoising Method Based on Diffusion Equation and Deep Learning

Shaobin Ma,[1,2] **Lan Li**,[1] **and Chengwen Zhang**[1,2]

[1]*School of Digital Media, Lanzhou University of Arts and Science, Lanzhou 730010, China*
[2]*VR Technology R&D and Promotion Center, Lanzhou University of Arts and Science, Lanzhou 730010, China*

Correspondence should be addressed to Shaobin Ma; 1000265@luas.edu.cn

Received 4 November 2021; Revised 23 November 2021; Accepted 24 November 2021; Published 6 January 2022

Academic Editor: Shan Zhong

Effective noise removal has become a hot topic in image denoising research while preserving important details of an image. An adaptive threshold image denoising algorithm based on fitting diffusion is proposed. Firstly, the diffusion coefficient in the diffusion equation is improved, and the fitting diffusion coefficient is established to overcome the defects of texture detail loss and edge degradation caused by excessive diffusion intensity. Then, the threshold function is adaptively designed and improved so that it can automatically control the threshold of the function according to the maximum gray value of the image and the number of iterations, so as to further preserve the important details of the image such as edge and texture. A neural network is used to realize image denoising because of its good learning ability of image statistical characteristics, mainly by the diffusion equation and deep learning (CNN) algorithm as the foundation, focus on the effects of activation function of network optimization, using multiple feature extraction technology in-depth networks to study the characteristics of the input image richer, and how to better use the adaptive algorithm on the depth of diffusion equation and optimization backpropagation learning. The training speed of the model is accelerated and the convergence of the algorithm is improved. Combined with batch standardization and residual learning technology, the image denoising network model based on deep residual learning of the convolutional network is designed with better denoising performance. Finally, the algorithm is compared with other excellent denoising algorithms. From the comparison results, it can be seen that the improved denoising algorithm in this paper can also improve the detail restoration of denoised images without losing the sharpness. Moreover, it has better PSNR than other excellent denoising algorithms at different noise standard deviations. The PSNR of the new algorithm is greatly improved compared with the classical algorithm, which can effectively suppress the noise and protect the image edge and detail information.

1. Introduction

The noise image is mainly caused by the imperfect system and equipment. In the process of transmission, the image polluted by noise will affect people's visual sense to varying degrees, sometimes even leading to the loss of many image features, making the image blurred, affecting the useful information of the image, and thus hindering people's normal recognition. The image polluted by noise will have a great adverse effect on the subsequent image processing, mainly including image segmentation, extraction, detection, and recognition. Therefore, it is very necessary and important to use a good denoising algorithm for image denoising.

Image denoising can improve the accuracy of human visual recognition information and is a necessary condition for us to correctly identify images. In nature, due to various internal or external reasons, it is almost impossible to find natural pictures that are completely free of noise pollution, and the images are more or less polluted. This will make it harder for us to identify the images. In particular, some noise images, such as medical images and security images, which affect our significant judgment, need us to carry out image noise reduction to obtain clear image features. For those heavily polluted images, they lose the original content of the image, so effectively improving the accuracy of image recognition is a necessary condition for image recognition. Image denoising can improve image quality and is a

prerequisite for further image processing. In image processing, image denoising is often the first and most important step. Obtaining high-quality and high-definition images through noise reduction is a strong guarantee and good foundation for subsequent image processing.

Traditional image denoising methods have been proposed for a long time and have been used for a long time now, but most of these algorithms are not very satisfactory, especially in the process of denoising, the details of the image will be lost, and most of the denoising performance and algorithm complexity are to be improved. Although the deep learning technology for image denoising has many scholars' research, neural network research, due to the barriers of hardware, has not been developed, making the neural network technology no longer have complex network model of computation for too much worry, because high-performance GPU multicore parallel computing is well suited for the neural network model and is a necessary prerequisite for the rise of deep learning. Although a lot of people are beginning to study abroad, after all, the direction is the forward direction, and theory and technology are not very mature. Although the use of the image denoising aspect really has achieved good results, there are still many problems, so continuing research and perfect image denoising theory and improving the effect of denoising are very necessary. Based on diffusion equation and deep learning (CNN) algorithm, this paper adopts multifeature extraction technology to study the richer features of the input image in the deep network and designs an image denoising network model based on deep residual learning of convolutional network, which has better denoising performance. From the comparison results, it can be seen that the improved denoising algorithm in this paper can also improve the detail restoration of denoised images without losing sharpness. Under different noise standard deviations, the PSNR of the proposed algorithm is superior to other excellent denoising algorithms.

The first part is the introduction, the second part is related work, the third part is the Adaptive Diffusion equation and deep learning algorithm for image dryness, the fourth part is example verification, and the fifth part is the conclusion.

2. Related Work

Gaussian filtering first introduces the diffusion equation into image processing [1]. On the basis of some theoretical and numerical operations, differential operators can be obtained by transforming local filter operators. Differential operators can be transformed by local filter operators. In the case of two-dimensional images, the general processing method is to treat the pair diffusion as a uniform linear diffusion process; that is to say, the diffusion coefficients at all points of the image are the same. However, there are still loopholes in this diffusion equation [2], whose uniform diffusion makes it impossible to retain important details such as edges while removing noise. In view of the defects of the uniform diffusion characteristics of Gaussian filtering, it is natural to come up with an ideal method: reduce the diffusion at the edge of the image according to the prior information of the image, so as to remove the noise from the image and protect the information at the edge [3]. Therefore, we use the gradient operator as the operator of the edge detection to construct a monotone decreasing function [4] in which the diffusion coefficient changes with the gradient of the original image; that is, the magnitude of the gradient is inversely proportional to the diffusion coefficient. This adaptive denoising method can not only remove the noise in the flat area of the image but also maintain the protection of the edge of the image, which is called the nonuniform diffusion equation. However, this method still has loopholes and shortcomings. First, the gradient at the edge is large, so the diffusion coefficient is small [5]. The smaller the diffusion coefficient is, the better the edge is maintained, but the effect of noise removal cannot be achieved. Then, the gradient will produce deviation for image edge detection affected by noise.

A nonlinear diffusion equation [6] is proposed, which takes the gradient operator obtained in the previous step as the edge detection operator in the iterative process to reduce the impact of noise on edge judgment. However, this method cannot deal with the noise of the image edge. The diffusion coefficient is improved and designed as a matrix, which makes the diffusion coefficient larger in the tangential direction of the edge and smaller in the vertical direction of the edge, thus overcoming the shortcomings of the PM model [7]. A common problem with low-order models is the "ladder effect." In view of the characteristics of the low-order model, a fourth-order diffusion equation is introduced [8]. In this model, Laplace is used to measure marginal areas, so there is no "ladder effect" in vision. However, the edge protection ability of the model is not good, and it is easy to produce speckle noise. In [3], an improved fourth-order diffusion equation model is proposed, which uses the modulus of a gradient to replace the absolute Laplace value as the operator of image edge detection. This model has a faster convergence speed and better denoising effect but blurred image edges. An anisotropic diffusion fourth-order denoising algorithm is proposed [9], which diffuses in different degrees in the normal and tangential directions and preserves the details of the image well, but this model has gradient artifacts. At present, the research on image denoising algorithms based on diffusion equations is also making continuous progress, mainly focusing on second-order diffusion equations, fourth-order diffusion equations, and higher-order diffusion equations [10]. The high-order weighted gradient variation model for image denoising, after the convolution of a ladder as a weighted function of the second derivative, newly established a high-order variational function and get a fourth-order partial differential diffusion model [11]; the model is effective in the noise, eliminates the staircase effect, and can have good protection effect on the edge. An improved semiadaptive threshold anisotropic diffusion filter denoising algorithm is proposed [12, 13]. In this model, the local difference method is adopted to distinguish damaged pixels from noiseless ones, and some damaged pixels are replaced by predenoised pixels of the Gaussian filter. Then, an anisotropic diffusion model with a

semiadaptive threshold in the diffusion coefficient function is used to obtain the restored image. In order to achieve the semiadaptive threshold value of each diffusion, the gradient value of the destroyed pixel is introduced into the threshold value [14], which makes the diffusion of the smooth region larger and the diffusion of the boundary region smaller. This method can improve PSNR by 30% and structure similarity (SSIM) by 5%. It has a good effect on edge protection and noise removal. It overcomes the defect of blurring image edges by the fourth-order isotropic diffusion equation. An adaptive image denoising algorithm based on the fourth-order diffusion equation [15] is proposed, which uses the image gradient model to construct the measure function of the image feature information, and uses the feature detection function to adjust the normal and tangential image diffusion coefficient indexes adaptively according to the different features of the image. Isotropic diffusion is used to remove noise in flat and inclined regions, and anisotropic diffusion is used to protect the features of image edge points. Experimental results show that both denoising and edge preservation are taken into account [16]. A high-order nonlinear diffusion image smoothing method based on curvature mode is proposed [17]. Methods such as the combination of gradient and curvature, total variational coupling, and quasi-normal distribution diffusion are applied to image denoising.

Diffusion equation and deep learning use Local Receptive Field structure designed for image data [18], compared with Plain Multilayer Perceptrons (MLPs); this structure enables diffusion equation and deep learning to obtain good results while greatly reducing the parameters. This method is very useful when the amount of training data is relatively small because the total number of parameters is relatively small, so it is more difficult to overfit the training data. On the other hand, multilayer perceptron networks have more potential than diffusion equations and deep learning: multilayer perceptron networks can be used to approximate arbitrary functions [19], while diffusion equations and deep learning can only learn specific types of functions due to structural constraints. Another class of algorithm denoising autoencoder also uses the neural network as a denoising tool [20]. Denoising autoencoders are special neural networks that learn by unsupervised learning. Features are represented by the learned units of the hidden layer. Since the input-output of this neural network can be easily generated, it only needs to add noise and other pollution processing to the input, and good data features can be learned, so denoising autoencoder has become a very important algorithm in deep learning. However, the purpose of designing this kind of neural network is mainly to extract data features through unsupervised learning and to train the deep neural network layer by layer, rather than for denoising. Another difference is the fact that the noise added to the network during learning is usually not Gaussian white noise but pepper and salt noise or the Dropout of some of the input nodes. For stacked denoising autoencoder [21, 22], which is commonly used in deep learning, the noise will be added to the output node of the previously learned layer, which is different from the scheme that only adds noise to the initial input and then learns all parameters at the same time in denoising applications. In fact, the most effective neural network in the field of denoising is the most basic multilayer perceptron model. As a popular solution to denoising problems, methods based on image priors can extend various methods based on image priors [23], such as BM3D [24]. These methods are used to obtain image prior knowledge directly from the input image for image denoising. Although some denoising effects have been achieved, there are still some limitations. Firstly, these methods make use of the prior noise image, which will cause some error, so it is difficult to obtain all the features of the image, resulting in the limitation of denoising performance. Second, most methods use only the internal information of the input image and do not use any external information, so there is still room for improvement. In addition, for denoising methods based on discriminant learning [25, 26], especially those based on diffusion equation and deep learning [8, 27, 28], denoising networks with paired training data sets are trained and image noise distribution is learned in the hidden layer, thus improving denoising performance. In order to remove the known Gaussian noise level, pairs of training samples are used to train the denoising network to achieve advanced denoising performance.

3. Adaptive Diffusion Equation and Deep Learning Algorithm for Image Dryness

3.1. Image Denoising Based on Diffusion Equation. The process of recovering a clear image from an image containing noise is called image denoising so that the image information after denoising is as close as possible to the original clear image information, and the information error between them is as small as possible. Therefore, image denoising is an inverse process of image processing, as well as image deblurring and superresolution reconstruction. The noise model of the image is shown in Figure 1.

The low rank of hyperspectral images can be explained from the perspective of the linear spectral mixture model. Because of the high correlation between spectral features, each spectral feature can be represented by a linear combination of a small number of pure spectral endmembers, that is, a linear spectral mixture model. If each band of hyperspectral image data is expanded into a vector, then all bands together form a matrix X.

Diffusion equations are related to many unknown multivariate functions and their partial derivatives. For fixed positive integer k, we use the symbol Dku to represent all k partial derivatives of u.

$$Dk = \frac{Dku}{Dx1, Dx2, \ldots, Dxk}, \tag{1}$$

where $(x_1, x_2, \ldots, x_n)$ is any permutations of k elements in the set $\{1,2,3, n\}$. Therefore, we can regard Dku as a vector in n-dimensional Euclidean space and write down its length as follows:

$$Dk\ u = \sum_i \frac{Dku}{Dx1, Dx2, \ldots, Dxi}. \tag{2}$$

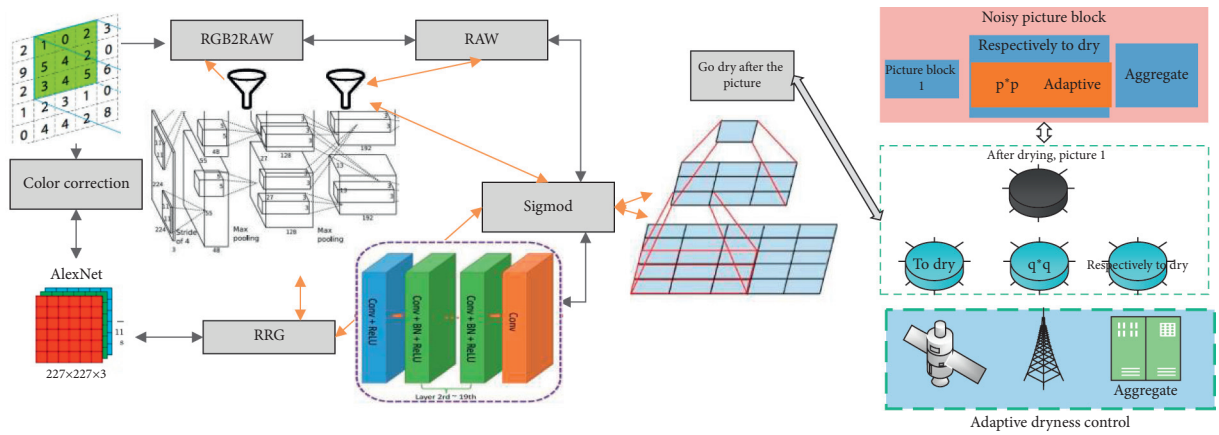


Figure 1: Research framework of adaptive image denoising based on the image block.

In particular, when $k = 1$, we call the n-dimensional vector:

$$Du = \frac{Du}{Dx1}\frac{Du}{Dx2}\cdots\frac{Du}{Dxi}. \tag{3}$$

That is the trace of u's Hessian matrix, the sum of the diagonal elements of u's Hessian matrix. Write the divergence of F as

$$\mathrm{div}F = \sum \frac{DF\ i}{Dx\ i}. \tag{4}$$

Let the general form of the diffusion model be

$$\mathrm{div}F\Delta I = \frac{DF}{DI}. \tag{5}$$

The AOS format of the deformable diffusion model is

$$F^{n+1} = (F - 2\Delta I(F^n))I^n. \tag{6}$$

The block-based denoising algorithm divides the image into small blocks for denoising, and because the noise level of each image block is different, the selection of the denoising threshold is also different. In the process of denoising, noise mainly corresponds to a small singular value. Therefore, the larger the singular value is, the less it should shrink as the singular value shrinks.

3.2. Diffusion Equation Denoising Algorithm. Based on the PM model, the fitting diffusion coefficient was established to overcome the defects of texture detail information loss and edge degradation caused by excessive diffusion intensity. Then, in order to enable the threshold function to control the threshold value according to the maximum gray value and iteration times of the image, the threshold function is designed adaptively. Based on the advantages and disadvantages of the above two diffusion coefficients, a new diffusion coefficient is established, as follows:

$$g_i = \chi g_1 + \gamma g_2. \tag{7}$$

The diffusion coefficients of each fit are shown in Figure 2.

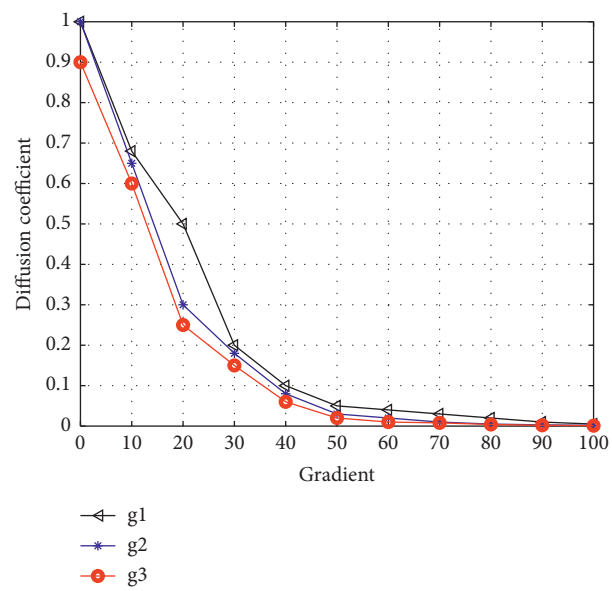


Figure 2: Each fitting diffusion coefficient.

In the process of diffusion, pixels' gray value of the area and adjacent area has a kind of relationship, usually paying a predetermined constant gradient threshold k and adopting the predetermined good constant gradient threshold, which is not conducive to certain areas on the edge of the image detail information and images, according to the pixel location of the dynamic design gradient threshold k. Through the above analysis, a one-dimensional gradient threshold function k is designed which varies with the diffusion time and the number of diffusions.

$$k(i) = \frac{1}{1 + \gamma i}. \tag{8}$$

In the tangent direction, the image needs strong diffusion in both the edge and the inner region, where noise can be eliminated and the interrupted edge can be connected on the edge. In the normal direction, the inner area of the image also needs strong diffusion to eliminate noise, but in the edge, no diffusion is needed as much as possible to maintain

the edge features. Therefore, the diffusion coefficient B is always set as 1 in the tangential direction, and a new diffusion coefficient c is designed by excluding one ill-conditioned condition in the normal direction combined with the diffusion coefficient of Tv flow:

$$c = \frac{1}{\sqrt{1 + \Delta D^2}}. \tag{9}$$

The neural network and deep learning model are very suitable for learning the function from picture block to picture block due to their strong expression ability. The frame of denoising using a neural network is shown in Figure 3. The hidden layer activation function is distinguished from the output layer activation function because the choice of the output layer activation function depends on the needs of the problem.

Specifically, we first randomly select a clean image block knife from the image data set and then artificially add Gaussian white noise to generate the corresponding noise image block X. Then, we take the vector-induced noise image block X as the input of the neural network and the corresponding vector-induced clean image block as the output of the neural network, update the parameters of the neural network through backward propagation, and gradually learn the required model through iteration.

3.3. Adaptive Image Denoising Algorithm Based on Diffusion Equation and Deep Learning. The initial value of the weight matrix of the neural network has a significant influence on the training process and the final result. For multilayer networks, we want the initial values of the network to satisfy randomness while ensuring that the input outputs of each hidden layer have the same statistical characteristics as possible. In order to achieve this, it is necessary to associate the initial value of parameters of each layer with the number of nodes of the corresponding layer and adjust it adaptively according to the number of nodes of each layer.

All training samples were checked in a cycle, and the parameters of the whole network were updated by backward conduction every time one training sample was checked. The advantages are that the updating speed of the weight matrix is greatly accelerated, and the network parameters are easier to escape from the subideal local minimum region because of the strong randomness. The disadvantage is that the error curve fluctuates more in the training process, the convergence conditions are more complex, and it is more difficult to converge. Relatively speaking, these shortcomings of stochastic gradient descent are easier to solve. For example, it is a good solution to solve the problem that the error is difficult to converge by gradually decreasing the learning rate with the increase of the number of iterations. Due to the advantages of training speed, a stochastic gradient descent algorithm has been widely used.

The setting of learning rate: a larger learning rate may make the neural network model learn faster and may also make the neural network diverge. In this paper, the initial learning rate was selected by the experimental method: 0.1 was initially tried, 0.05 was selected if the neural network diverged, and then 0.02 was tried to gradually reduce the initial learning rate. If the network converges but the training speed is slow, then we try to increase the learning rate. In order to ensure the convergence of the neural network, the current learning rate is multiplied by 0.99 in each cycle when the small-batch processing algorithm is implemented. Inertia parameter is a common acceleration method in neural network training, which can help the neural network to leave the flat region in function space faster. The specific operation method is to add part of the last weight update to each update of the weight matrix, namely,

$$W(i+1) = \lambda \frac{DE}{DW} + \beta W(i). \tag{10}$$

The relation between the activation value of the element and input current is as follows:

$$f(i) = \log \frac{E + DW - V}{E + DW - V_h}. \tag{11}$$

As shown by the similarity measure,

$$d(E, D) = \frac{|E - D|^2}{k * k}. \tag{12}$$

The EPLL algorithm first learns the prior knowledge of the image and obtains the edge texture characteristics of the image through sample learning, which is constrained by the regular term of image denoising to obtain the denoising image with obvious texture characteristics. On the other hand, the image blocks of the default denoised images in the EPLL algorithm have significant similarities with the image blocks in the samples. When the EPLL value is large, the denoised image is considered to conform to the characteristics of the sample library, and the subjective vision is also better. The sum calculation method of likelihood probability is as follows:

$$EP(x) = \sum_i \log\ p * P(x). \tag{13}$$

The logarithmic likelihood probability of all image blocks is summed to obtain the total likelihood probability of the denoised image and sample database image. The image degradation process is known, and the properties of likelihood probability are understood. The image denoising problem is transformed into a solution formula:

$$f_p(x, y) = \frac{\gamma}{2}|Ax - y|^2 - EP(x). \tag{14}$$

The solution of the target equation is approximated by introducing the auxiliary variable splitting equation into semiquadratic partition:

$$C(x, y) = \frac{\gamma}{2}|Ax - y|^2 - \sum \log\ p(y). \tag{15}$$

The adaptive image drying model has higher denoising intensity, which solves the "blocky" effect in the denoising process. In order to further protect the edge texture and other details and give consideration to denoising intensity, an adaptive orthogonal diffusion filtering model was proposed to overcome the disadvantages of high denoising

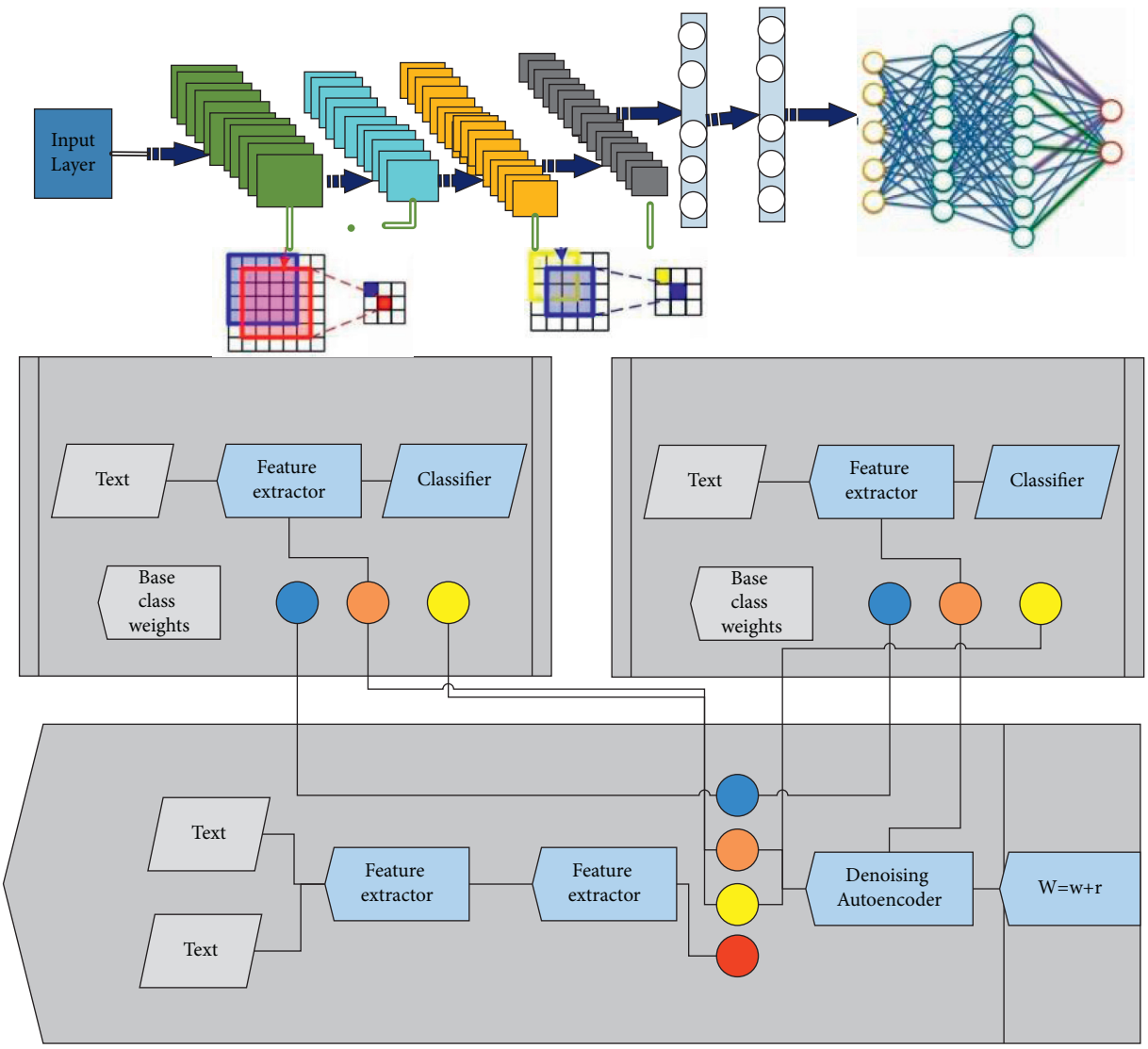


Figure 3: Frame of denoising using neural network.

intensity of the orthogonal diffusion model and the loss of edge texture and other details. The experimental results show that the PSNR model improves about 30 dB compared with the classical model (TV flow diffusion model). Compared with the TV stream diffusion model, the new model is more flexible than the TV stream diffusion model, which can adjust the diffusion system according to different parts of the image processing, control the smoothness degree, and process clear images more reasonably.

4. Example Verification

In order to verify the rationality and effectiveness of the above algorithms, Lady (600×600) and Nudist graphs with Gaussian random noise were analyzed, and experimental simulation was carried out with Matlab software and semi-implicit additive operator splitting numerical algorithm, and their MSE and PSNR were compared. In order to study the suppression effect of diffusion coefficient on noise and the retention degree of important details such as image edge texture, noises of different sizes and variances were first added to the image, and the results are shown in Figure 4.

Based on the PM model, the diffusion coefficient in the diffusion equation is improved, and the fitting diffusion coefficient is established to overcome the defects of texture detail loss and edge degradation caused by excessive strength. Then, according to the maximum gray value of the image and the number of iterations, the threshold function is designed automatically. Experimental results show that the peak signal-to-noise ratio of the new algorithm is improved by about 16 dB compared with the classical algorithm. The proposed algorithm can effectively suppress the noise and protect the edge and detail information of the image.

Notice that the side length P of the noisy image block can be inconsistent with the side length Q of the denoised image block. Such setting is similar to the convolution operation of the image. For the pixel in the middle of the image block, the information of surrounding pixels can be used more than that of the pixel from the corner of the image block. Therefore, the prediction of the original value of the pixel in

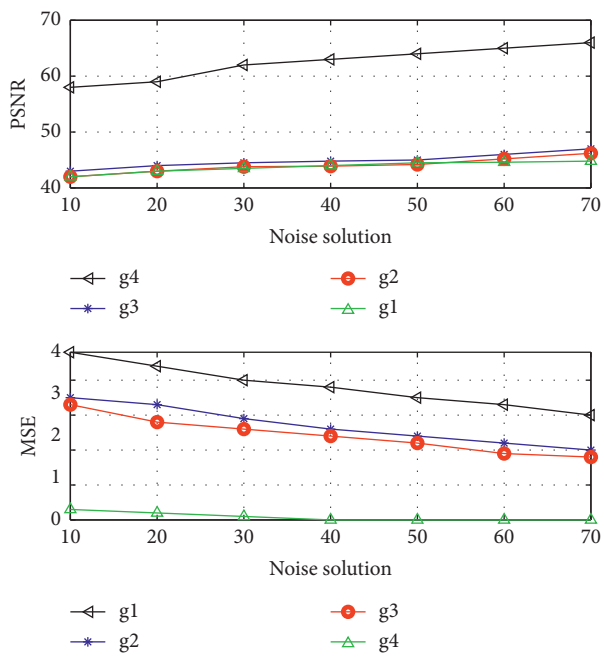


Figure 4: PSNR and MSE simulation diagrams of different variances of diffusion coefficients in lady image.

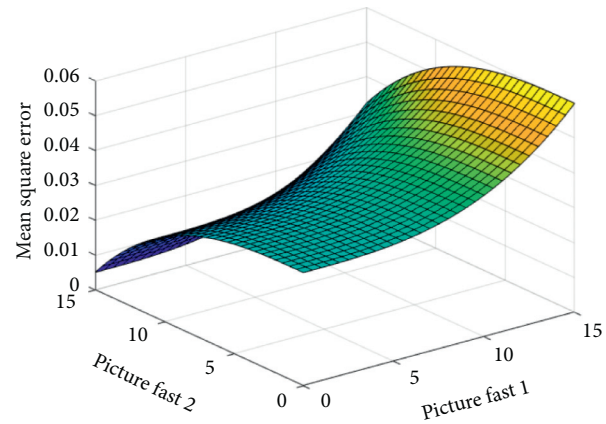


Figure 5: Mean square error at each position of the 12×12 image block after denoising.

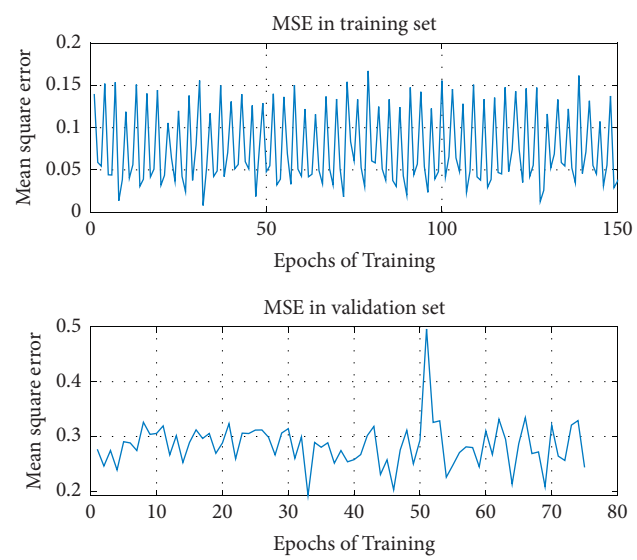


Figure 6: Changes of mean square error of diffusion equation and deep learning in the training set and test set with the number of iteration cycles during training.

the middle part of the image block will be more accurate, while the error of the edge part will be relatively larger. The actual situation in the image block is an example of mean square error (MSE) in each position as shown in Figure 5; in the case where the original noise block size is 12×12 images, we use the fourth chapter putting forward the algorithm of denoising model of training, more after denoising images after denoising and ideal cleaning image blocks of error, and perform statistics after getting the picture.

As can be seen from Figure 5, the mean square error in the middle of the image block is the lowest, and the mean square error becomes higher toward the edge, among which the mean square error at the corner is the highest. Numerically, the mean square error (MSE) in the corner of the image block is approximately 25% greater than that in the middle. On the other hand, in many algorithms, the edge length of the noisy image block is the same as that of the denoised image block. Considering that there will be some overlap in the process of splitting and aggregation, the same size of denoised image block and noise image block means that the same number of noise image blocks provides more estimates of the clean value of each pixel.

Figure 6 records the curves of the mean square error of denoising image blocks in the training of the two models during the first 150 cycles with the number of iterations. It should be noted that from the 150th cycle to the 1000th cycle, the shape of each curve is basically the same as that from the 100th cycle to the 150th cycle.

It can be seen from Figure 6 that the model using linear rectification function can fit the training data set better than the model using hyperbolic tangent function. Firstly, in the whole training process, the model using linear rectifying function has a lower mean square error in both training set and test set than the model using hyperbolic tangent function. Secondly, it only takes a few iterations for the former to stabilize the mean square error on the test set, and then the error on the training set continues to decline, while the error on the test set slowly rises; that is, slight overfitting appears. However, the latter fluctuates more and the range is larger in the training process, and there is no overfitting trend in the whole process. As the learning rate decreases (multiplied by 0.99 per cycle), it gradually converges.

Based on the diffusion equation and deep learning, the Adam algorithm is used to replace the traditional gradient descent algorithm in the backpropagation algorithm like the DnCNN algorithm, multifeature extraction technology is adopted for feature extraction of the first layer of the neural network model, and the improved linear rectifier function is used as the activation function. Deep learning technology promotes the training speed and convergence speed of the whole model and has great advantages in convergence and speed.

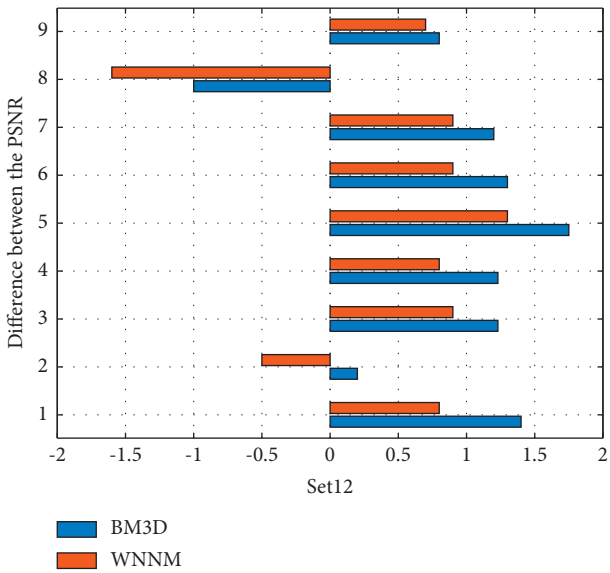


Figure 7: Loss function diagram.

As shown in Figure 7, we can find from the curve that the loss value of the blue line with low noise standard after convergence is much lower than that of the blue line with high noise standard after convergence, which conforms to the objective law, because the bigger the noise is, the greater the interference between pixels is, and the more difficult the model optimization is. In addition, when the noise is 35, we found that after 316 times of iteration the curve almost tends to be convergent, while a curve in high noise for 75, almost 1000 times or so, tends to be convergent, and the iteration is only a small drop, tens of thousands of times compared with diffusion equation and deep learning model with almost 4000 times. The loss function curve converges and becomes stable. The improved diffusion equation and deep learning model based on adaptive dryness in this paper have advantages in convergence speed and model training speed.

5. Conclusion

In this paper, the diffusion coefficient in the diffusion equation is improved and the fitting diffusion coefficient is established to overcome the defects of texture detail loss and edge degradation caused by excessive diffusion intensity. Then, the threshold function is adaptively designed and improved so that it can automatically control the threshold of the function according to the maximum gray value of the image and the number of iterations, so as to further preserve the important details of the image such as edge and texture. Finally, the simulation results show that the peak signal-to-noise ratio of the new algorithm is greatly improved compared with the classical algorithm, which can effectively suppress the noise while protecting the image edge and detail information. On the basis of image denoising based on neural network denoising, a linear correction function is proposed as the activation function of the neural network hidden layer, and the parameter setting of the model is discussed in detail. Experimental results show that the denoising performance of this algorithm reaches the best level at present, especially suitable for image denoising in a high noise environment.

Data Availability

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

Conflicts of Interest

There are no conflicts of interest in this article.

Acknowledgments

This study was supported by the Industrial Support and Guidance Project of Colleges and Universities in Gansu Province: Research and implementation of digital comprehensive display system for Dadiwan prehistoric civilization based on virtual reality (Grant no. 2019C-09).

References

[1] K. Satya and T. Jayachandra, "Deep learning approach for image denoising and image demosaicing," *International Journal of Computer Application*, vol. 168, no. 9, pp. 18–26, 2017.

[2] S. Kumar, M. Sar Fa Raz, and M. K. Ahmad, "Denoising method based on wavelet coefficients via diffusion equation," *Iranian Journal of Science and Technology. Transaction A, Science*, vol. 42, no. 6, pp. 41–56, 2017.

[3] X. Y. Meng, L. Che, and Z. H. Liu, "Towards a partial differential equation remote sensing image method based on adaptive degradation diffusion parameter," *Multimedia Tools and Applications*, vol. 76, pp. 17651–17667, 2017.

[4] K. M. Santosh, "Denoising method based on wavelet coefficients via diffusion equation," *Iranian Journal of Science & Technology. Transaction A Science*, vol. 42, no. 6, pp. 98–104, 2018.

[5] S. Zhai, Z. Weng, and X. Feng, "An adaptive local grid refinement method for 2D diffusion equation with variable coefficients based on block-centered finite differences," *Applied Mathematics and Computation*, vol. 268, pp. 284–294, 2015.

[6] J. Yu, J. Yin, J. Yin, S. Zhou, S. Huang, and X. Xie, "An image super-resolution reconstruction model based on fractional-order anisotropic diffusion equation," *Mathematical Biosciences and Engineering*, vol. 18, no. 5, pp. 6581–6607, 2021.

[7] J. Yu, L. Tan, and S. Zhou, "Image denoising based on adaptive fractional order anisotropic diffusion," *KSII Transactions on Internet and Information Systems*, vol. 11, no. 1, pp. 436–450, 2017.

[8] M. Jin, X. Feng, and K. Wang, "Gradient recovery-based adaptive stabilized mixed FEM for the convection-diffusion-reaction equation on surfaces," *Computer Methods in Applied Mechanics and Engineering*, vol. 380, no. 255, pp. 113798–113809, 2021.

[9] R. Li, T. Zeng, H. Peng, and S. Ji, "Deep learning segmentation of optical microscopy images improves 3-D neuron reconstruction," *IEEE Transactions on Medical Imaging*, vol. 36, no. 7, pp. 1533–1541, 2017.

[10] R. Lin, R. Zhang, C. Wang, X.-Q. Yang, and H. L. Xin, "TEMImageNet training library and AtomSegNet deep-

learning models for high-precision atom segmentation, localization, denoising, and deblurring of atomic-resolution images," *Scientific Reports*, vol. 11, no. 1, pp. 5386–5398, 2021.

[11] L. Wang, S. Zhou, and X. Lin, "A novel adaptive image zooming method based on nonlocal Cahn-Hilliard equation," *Knowledge-Based Systems*, vol. 166, pp. 166–189, 2018.

[12] Y. Jin, X. B. Jiang, Z. K. Wei, and Y. Li, "Chest X-ray image denoising method based on deep convolution neural network," *IET Image Processing*, vol. 13, no. 11, pp. 1970–1978, 2019.

[13] 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.

[14] N. Salamat, M. Missen, and V. Prasath, "Recent developments in computational color image denoising with PDEs to deep learning: a review," *Artificial Intelligence Review*, vol. 54, no. 8, pp. 1–32, 2021.

[15] D. Liu, W. Wang, and X. Wang, "Poststack seismic data denoising based on 3-D convolutional neural network," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 8, no. 9, pp. 1–32, 2019.

[16] Z. Li, S. Zhou, and J. Huang, "Investigation of low-dose CT image denoising using unpaired deep learning methods," *IEEE Transactions on Radiation and Plasma Medical Sciences*, vol. 5, no. 6, pp. 51–81, 2020.

[17] M. Xie, Z. Zhang, W. Zheng, Y. Li, and K. Cao, "Multi-frame star image denoising algorithm based on deep reinforcement learning and mixed Poisson-Gaussian likelihood," *Sensors*, vol. 20, no. 21, pp. 5983–5998, 2020.

[18] R. Cai, "Research progress in image denoising algorithms based on deep learning," *Journal of Physics: Conference Series*, vol. 1345, pp. 42055–42067, 2019.

[19] W. Lu, J. A. Onofrey, Y. Lu et al., "An investigation of quantitative accuracy for deep learning based denoising in oncological PET," *Physics in Medicine and Biology*, vol. 64, no. 16, pp. 165019–171335, 2019.

[20] C. Wu and T. Gao, "Image denoise methods based on deep 1eaming," *Journal of Physics: Conference Series*, vol. 1883, no. 1, pp. 12112–12123, 2021.

[21] K. Yan, L. Chang, M. Andrianakis, V. Tornari, and Y. Yu, "Deep learning-based wrapped phase denoising method for application in digital holographic speckle pattern interferometry," *Applied Sciences*, vol. 10, no. 11, pp. 4044–4057, 2020.

[22] C. Alla Takam, O. Samba, A. Tchagna Kouanou, and D. Tchiotsop, "Spark Architecture for deep learning-based dose optimization in medical imaging," *Informatics in Medicine Unlocked*, vol. 19, pp. 100335–100355, 2020.

[23] S. Zhong, W. Weng, K. Chen, and J. Lai, "Deep-learning steganalysis for removing document images on the basis of geometric median pruning," *Symmetry*, vol. 12, no. 9, pp. 1426–1438, 2020.

[24] S. Zhang, S. Xu, and L. Tan, "Stroke lesion detection and analysis in MRI images based on deep learning," *Journal of Healthcare Engineering*, vol. 2021, no. 5, 9 pages, Article ID 5524769, 2021.

[25] D.-I. Eun, R. Jang, W. S. Ha, H. Lee, S. C. Jung, and N. Kim, "Deep-learning-based image quality enhancement of compressed sensing magnetic resonance imaging of vessel wall: comparison of self-supervised and unsupervised approaches," *Scientific Reports*, vol. 10, no. 1, pp. 13950–13967, 2020.

[26] C. Chen and Z. Xu, "Aerial-image denoising based on convolutional neural network with multi-scale residual learning approach," *Information*, vol. 9, no. 7, pp. 324–345, 2018.

[27] X. Xiao, C. Yang, and X. Yang, "Adaptive learning-based projection method for smoke simulation: adaptive Projection Method based on Machine Learning," *Computer Animations and Virtual Worlds*, vol. 29, no. 3-4, pp. e1837–e1845, 2018.

[28] B. Li and W. Xie, "Image denoising and enhancement based on adaptive fractional calculus of small probability strategy," *Neurocomputing*, vol. 175, no. 29, pp. 704–714, 2016.

Hindawi
Journal of Robotics
Volume 2021, Article ID 3965689, 9 pages
https://doi.org/10.1155/2021/3965689

Research Article

Collaborative Task Offloading Strategy of UAV Cluster Using Improved Genetic Algorithm in Mobile Edge Computing

Hong Wang

Department of Electronic Information Technology, Sichuan Modern Vocational College, Chengdu, Sichuan 610207, China

Correspondence should be addressed to Hong Wang; whiam@163.com

Received 16 November 2021; Revised 7 December 2021; Accepted 13 December 2021; Published 29 December 2021

Academic Editor: Kaijian Xia

Aiming at the problem that traditional fixed base stations cannot provide good signal coverage due to geographical factors, which may reduce the efficiency of task offloading, a collaborate task offloading strategy using improved genetic algorithm in mobile edge computing (MEC) is proposed by introducing the unmanned aerial vehicle (UAV) cluster. First, for the scenario of the UAV cluster serving multiple ground terminals, a collaborative task offloading model is formulated to offload the tasks to UAVs or the base station selectively. Then, an objective function and related constraints are put forward to minimize the time delay and energy consumption by analysis of those in the communication and computing process in the system while considering many factors. Then, the improved genetic algorithm is introduced to solve the optimization problem, obtaining the optimal collaborative task offloading strategy. To verify the performance of the proposed method, simulations are conducted on MATLAB. Simulation results showed that the joint utilization of UAV and MEC improves the offloading efficiency of the proposed strategy. When the number of UAVs is 12, the total utility is up to 1.83 and the task completion time does not exceed 110 ms. In this case, the task can be reasonably offloaded to UAVs or accomplished locally.

1. Introduction

With the rapid development of big data, cloud computing, artificial intelligence, and Internet of things (IoT) technology and the maturity of 5G communication technology, the inevitable trend of interconnection of all things has greatly accelerated the evolution of IoT systems toward multiple source isomerism and intelligent online processing, such as smart city, smart home, and telemedicine [1, 2]. However, it is bound to bring great pressure to the network link and data center when transmitting all the data of massive IoT terminals to the cloud computing center for centralized processing and then returning the results. In addition, the cloud computing center may probably refuse to provide services due to the overload, which will greatly affect the user experience. In this context, mobile edge computing (MEC) is introduced [3].

The allocation of network resources in edge computing is similar to that in traditional cloud computing, as it mainly involves the reasonable allocation of computing resources and energy for user terminals and services. Additionally, in order to meet the requirements of energy consumption and packet loss rate, it needs to consider the offloading of users' tasks [4, 5]. Mobility management and computation offloading are two main technologies in mobile edge computing in recent research. Mobility management mainly concerns the resource discovery and switching, while computation offloading solves the problems of when to offload computation tasks, what to offload, and how much to offload [6]. Due to the flexibility of edge computing equipment, the service platform is no longer limited to fixed communication base stations or roadside units. Unmanned aerial vehicles (UAVs) with computing capability can also provide computing and relay services for mobile users [7]. Based on the strong mobility of UAVs, it can avoid the problems caused by many geographical factors and fill the blind area in the signal covering of communication base stations [8].

There has been some research investigating on task offloading and resource optimization allocation in MEC.

Saleem et al. proposed an online algorithm with low complexity, which realizes dynamic computation offloading based on Lyapunov optimization and minimizes the task execution time [9]. The performance needs to be further improved in the complex environment where multiple nodes and tasks processing are involved. A novel privacy-preserving and cost-efficient task offloading scheme based on the general Lyapunov optimization was given by He et al. [10]. This scheme can protect users' privacy while ensuring the best user experience. However, the processing efficiency of computing tasks is low, which is not suitable for large-scale multitask collaborative allocation. By applying the Lagrange multiplier method, Yang et al. formulated the offloading strategy and resource allocation optimization into a nonlinear equation with linear inequality constraints and then proposed the bisection search algorithm to solve them effectively, which improves the processing efficiency. Nevertheless, the communication delay has a great impact on the quality of service [11]. In order to reduce the influence of long distance between users and cloud servers on the quality of service, Ito and Koga proposed a flow splitting and aggregation scheme to lessen the offloading delay of cloud servers in edge computing, while the energy consumption still needs to be optimized [12]. Most of the aforementioned methods focus on the single-objective optimization, such as delay in computation task processing, energy consumption, and resource allocation. The overall multiobjective optimization still needs to be deeply studied.

With the continuous development of computer technology and advancement of communication technology, the advantages of heuristic algorithms in solving optimization problems have been gradually highlighted. In Reference [13], the vehicular mobile edge computing system was studied. Two independent heuristic algorithms were adopted to minimize the average delay of computation task processing while considering the limitations of vehicle mobility and energy consumption, but the energy consumption optimization in the computation process is not considered. Hossain et al. proposed a novel fuzzy-based collaborative task offloading scheme. By introducing the fuzzy logic method, the target MEC server is selected for task offloading, so as to accommodate more computation workload in the MEC system and reduce the dependence on remote cloud [14]. However, it only offloads the tasks to fixed edge servers, lacking the consideration of mobile computing servers, such as UAVs. Chen et al. formulated the task offloading as an integer nonlinear programming problem and proposed an efficient task offloading and channel resource allocation scheme based on differential evolution algorithm, which can significantly reduce the energy consumption and has a good convergence performance [15].

The computation tasks mentioned above are offloaded to intelligent terminals such as base stations, cars, and mobile phones, thus in some blind areas that are not covered by communication signals, those tasks cannot be processed. Based on the above analysis, aiming at the problem that most existing offloading strategies are not suitable for UAV clusters and taking the coverage of communication signals into consideration, a collaborate task offloading strategy of UAV cluster using improved genetic algorithm in MEC is proposed. In order to solve the signal coverage problem in the traditional cloud computing, the proposed strategy uses the high mobility of UAVs and designs a collaborative task offloading model, where tasks can be partially offloaded to the UAV cluster with computing and communication capabilities or transmitted to the remote base station, which improves the task processing efficiency.

2. System Model and Optimization Object

2.1. System Model. The system model of collaborate task offloading of UAV cluster is shown in Figure 1, which is composed of N number of mobile terminals (MTs), M number of UAVs, and one base station (BS). In order to decrease the computation delay and energy consumption of MTs, the tasks can be partially offloaded to the UAV cluster with computing and communication capabilities or transmitted to the remote base station indirectly. Assuming that mobile terminals are randomly distributed in the test area, UAVs can monitor tasks and obtain data by taking photos or videos to meet the needs of real application scenarios.

The goal of the system is to minimize the weighted delay and energy consumption of the whole system. In order to achieve this, the first thing is to obtain the delay and energy consumption of the whole system. Specifically, the whole process includes task offloading and the return of computation results, as shown by the arrows in Figure 1. Moreover, task offloading consists of two parts in the link, which are the communication and computation process, respectively. Generally, the wireless channel between UAV and MT or BS is mainly the line of sight (LoS) link, so the Doppler frequency shift can be well compensated when the UAV moves.

2.2. Communication Model. The UAV-assisted MEC system is mapped into a three-dimensional Euclidean coordinate system, where the coordinates of MT and BS are located at (0, 0, 0) and (L, 0, 0), respectively. In addition, the rotor UAV can fly or hover at a fixed height H (which can be the minimum height required to avoid ground buildings, etc.), and the UAV flies at a constant speed $v > 0$ (which satisfies the constraints of maximum speed of the UAV). In order to facilitate the analysis, it is assumed that UAVs can also execute computation tasks during the flight, but UAVs must hover at a fixed position to receive input data of tasks from MTs or transmit data to the BS, as shown in Figure 2.

In this paper, the initial hover position of the UAV when executing the first task group is expressed as $u_{k,i}^{0} = (x_{k,i}^{0}, y_{k,i}^{0}, H)$; the hover position of the UAV when collecting the input data of k-th task group of MT is expressed as $u_{k,i}^{C} = (x_{k,i}^{C}, y_{k,i}^{C}, H)$; the hover position of the UAV when transmitting the data of k-th task group to BS is expressed as $u_{k,i}^{B} = (x_{k,i}^{B}, y_{k,i}^{B}, H)$. Note that the hover position of the UAV after the execution of k-th group is the initial hover position when the UAV executes the $(k+1)$-th task group. Thus, the channel gains from UAV to MT and BS can be formulated respectively:

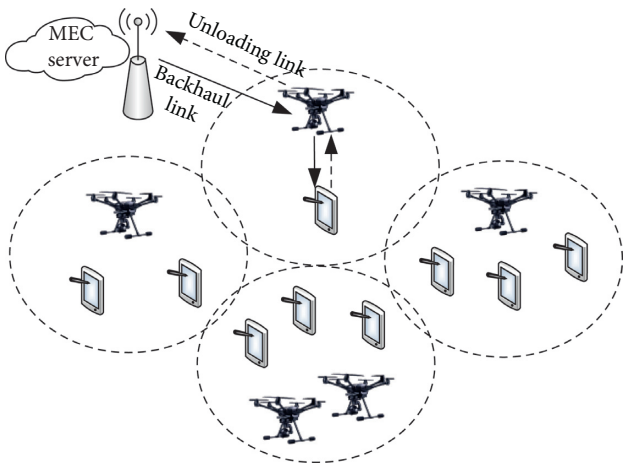


Figure 1: System model of collaborate task offloading of the UAV cluster.

$$g_k^C = \frac{p_0}{\left(x_k^C\right)^2 + \left(y_k^C\right)^2 + H^2},$$
$$g_k^B = \frac{p_0}{\left(x_k^B\right)^2 + \left(y_k^B\right)^2 + H^2}, \tag{1}$$

where p_0 is the received power when the reference distance is 1 m and the transmission power is 1 W.

p^C and p^B stand for the transmission power of the MT and the UAV, respectively, and it is assumed that each data link used for communication in the system is allocated with equal bandwidth W. Therefore, the maximum data transmission rates of the data upload link from MT to UAV and the relay link from UAV to BS are, respectively, as follows:

$$v_k^C = W\log_2\left(1 + \frac{p^C g_k^C}{\sigma^2}\right),$$
$$v_k^B = W\log_2\left(1 + \frac{p^B g_k^B}{\sigma^2}\right), \tag{2}$$

where σ^2 stands for the noise power from the UAV and the BS receiver.

2.3. Computation Model. From the perspective of a system, the task computation process mainly occurs on the devices like MT, UAV cluster, and BS. Due to the high-performance processing server and sufficient storage resources on the BS, the system does not take the delay and energy consumption into account when tasks are executed on the BS. Therefore, compared with tasks locally executed on MTs or offloaded to UAVs, the delay and energy consumption are much smaller when tasks are executed on the BS for the whole system [16]. The specific calculation process is as follows.

2.3.1. Local Execution Model. As the name implies, local execution means that the task is executed on MTs. In the process of local execution, the delay of the task can be represented as

$$T_n^{\text{local}} = \frac{G_n}{f_n}, \tag{3}$$

where G_n is the CPU cycle of the task on the mobile terminal n. f_n is the CPU frequency of the node n. By applying dynamic voltage and frequency scaling (DVFS) technology, the MT can control the energy consumption by reducing the frequency. Hence, the energy consumption of mobile terminal n can be written as

$$E_n^{\text{local}} = T_n^{\text{local}} \eta_n f_n^3, \tag{4}$$

where η_n is the effective capacity coefficient of the mobile terminal n. The volume of data executed locally, which is represented as D_n^{local}, can be calculated as

$$D_n^{\text{local}} = \frac{T_n^{\text{local}} f_n}{G_n}. \tag{5}$$

2.3.2. Execution on the UAV Model. Assuming that UAVs also apply DVFS technology to improve the energy efficiency, the delay and energy consumption when task is executed on UAV m can be calculated, respectively, as

$$T_m = \frac{G_m}{f_m},$$
$$E_m = T_m \eta_m f_m^3, \tag{6}$$

where G_m is the CPU cycle of the task on the UAV m. f_m is the CPU frequency of UAV m. η_n is the effective capacity coefficient of the UAV m. The bits of data executed on UAV m can be calculated as follows:

$$D_m = \frac{T_m f_m}{G_n}. \tag{7}$$

2.3.3. Energy Consumption Model when UAV Flies. E^{fly} represents the energy consumption when the UAV is in flight, which can be formulated as follows:

$$E^{\text{fly}} = \lambda \left\| v_u \right\|^2, \tag{8}$$

where $\lambda = 0.5$ and v_u is the flight speed of the UAV.

2.4. Model Constraints Description. In the UAV cluster collaborate task offloading system, considering the generally limited energy of UAVs, it is necessary to clarify the constraints of the energy consumption per unit time [17]. In this scenario, the UAV no longer needs to move in space, thus it only requires to calculate the energy consumed by the UAV when providing computing services for the ground terminal and hovering in the air [18].

$$T_m \eta_m f_m^3 + E_m^{\text{fly}} \le E_m^u, \quad \forall m \in M. \tag{9}$$

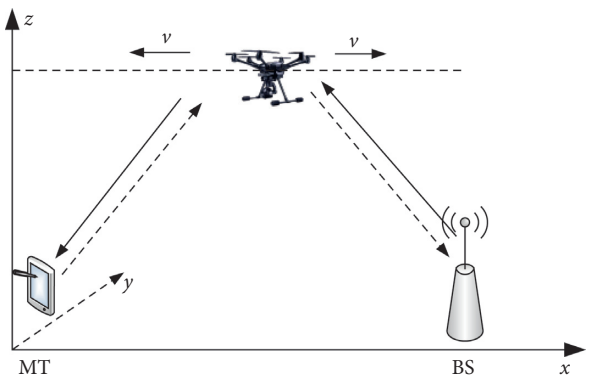


Figure 2: UAV-assisted MEC system model.

The inequality constraint represented in (9) shows the energy limitation of UAVs. The first term of the summation formula on the left of the inequality sign represents the total energy consumption of UAV $m \in M$ for the computation process, and the second term represents the energy consumption when the UAV is hovering.

Likewise, the energy of the ground terminal is also constrained. The energy consumption of the ground terminal mainly comes from local execution and transmission of data to the UAV, which is shown as

$$T_n^{\text{local}} \eta_n f_n^3 + \left(1 - \varpi_n^0\right) \lambda_n s_n e_n \le E_n^e, \quad \forall n \in N, \tag{10}$$

where ϖ^0 is the proportion of tasks offloaded to the UAV. λ_n is the number of tasks arrived on the MT in unit time, and s_n is the average size of tasks executed on the MT. e_n is the energy consumed by the MT when sending one bit data to the UAV through air interface. The inequality constraint in (10) represents the limited energy of the ground terminal. The two summation terms on the left of the inequality represent the energy consumption when the task is executed locally and task data are transmitted to the UAV, respectively. The right part of the inequality is the upper limit of available energy of the local ground terminal.

Additionally, not all the UAVs require to meet the communication conditions with each ground terminal in multi-UAVs scenario, but a communication link between them must be ensured when the ground terminal offloads the task to a UAV. Hence, we can write

$$\varpi_n^m \left[\left(x_m - a_n\right)^2 + \left(y_m - b_n\right)^2 \right] \le \varpi_n^m \left(H_m^u \tan \Theta^u\right)^2. \tag{11}$$

Equation (11) shows the constraint about the communication connectivity between the ground terminal and the UAV. If the offloading variable ϖ_n^m is ignored, the left part of the inequality represents the Euclidean distance between the projection point of the UAV's hovering position on the ground plane and the ground terminal, and the right part of the inequality represents the radius of the circular coverage area formed by the UAV's radar signal on the ground. When the inequality condition in (11) is satisfied, the UAV can communicate with the ground terminal. Noted that the offloading variables are multiplied on both sides of the inequality. If $\varpi_n^m \ne 0$, it means that the ground terminal $n \in N$ offloads ϖ_n^m proportional of the task to the UAV $m \in M$. At this time, the variables ϖ_n^m on both sides of the inequality can be reduced to meet the conditions described above. If $\varpi_n^m = 0$, it means the ground terminal does not offload any part of the task to the UAV. Thus, both sides of the inequality are 0 and the inequality holds. As there is no real data flow, whether the communication link between the terminal and the UAV exists does not affect the solution of the problem and the connectivity constraint no longer needs to be considered.

Finally, the ground terminal may transmit the task to multiple UAVs at the same time, so the conservation of data flow can be written as

$$\sum_{m=0}^{M} \varpi_n^m = 0 = 1, \quad \forall n \in N. \tag{12}$$

The equality constraint in (12) shows that the sum of the proportion of transmitted data and the proportion of local data should be equal to 1 for each ground terminal.

2.5. Optimization Object. Combined with the detailed description of the objectives and constraints of the problem mentioned above, the system model can be formulated as follows:

$$\begin{aligned} P:\ & \min\left[\omega_1\left(E_n^{\text{local}} + E_m + E^{\text{fly}}\right) + \omega_2\left(T_n^{\text{local}} + T_m\right)\right], \\ \text{s.t. } & 0 \le \varpi_n^m \le 1, \quad n = 1, 2, \ldots, N, m = 1, 2, \ldots, M, \\ & x_m^{\min} \le x_m \le x_m^{\max}, \quad m = 1, 2, \ldots, M, \\ & y_m^{\min} \le y_m \le y_m^{\max}, \quad m = 1, 2, \ldots, M, \\ & H_{m,\min}^u \le H_m^u \le H_{m,\max}^u, \quad m = 1, 2, \ldots, M, \end{aligned} \tag{13}$$

where ω_1 and ω_2 are the weights of energy consumption and delay, respectively.

In addition to the constraints from (9) to (12), those new constraints in (13) are the initial feasible region of the optimization variables so as to find the initial search window of the algorithm.

3. Improved Genetic Algorithm-Based Computation Allocation Strategy

Considering the task dependency between multiple sites of the UAV cluster, the complexity of solving the optimal computation offloading strategy is greatly improved, and it is hard to derive the optimal solution in polynomial time. Therefore, an improved genetic algorithm is introduced to find the optimal solution of task offloading strategy [19, 20].

3.1. Encoding and Population Initialization. The first step to realize a genetic algorithm is to encode the solution space of the problem. In the proposed UAV cluster collaborative task offloading model, the application is divided into K subtasks to be offloaded. Therefore, in the genetic algorithm, each chromosome should be made up from K genes. The tasks deployed at each site can be executed locally, transmitted to

edge nodes, or accomplished on the cloud. Hence, each gene has three possible values (cloud: −1, edge: 0, local: 1) [21]. Each chromosome represents a possible solution for the computation offloading. Figure 3 illustrates a chromosome composed of K genes.

3.2. Fitness Function. The quality of an individual (solution) in the genetic algorithm is evaluated by the fitness value, which is defined as the reciprocal of the total cost of computation offloading for the application to judge the merits of each computation offloading scheme. A higher fitness value represents a lower total cost and a better computation offloading scheme [22]. The fitness function can be formulated as follows:

$$\text{fitness} = \frac{1}{\omega_1\left(E_n^{\text{local}} + E_m + E^{\text{fly}}\right) + \omega_2\left(T_n^{\text{local}} + T_m\right)}. \tag{14}$$

In each process of the iterative evolution, the strategy of elite selection is adopted to retain the elite solution for the next generation of populations. In the i-th iteration, the fitness value of each individual in the i-th population is calculated first, and these individuals will be arranged in a descending order according to their fitness value. Then, the latter half of individuals with lower fitness values are discarded, and the remaining half of individuals with higher fitness values in the population are selected into the next crossover operation [23].

3.3. Genetic Operation. After completing the encoding process of chromosomes and the selection of fitness function, the following part focuses on three operations related to chromosomes, including selection, crossover, and mutation [24, 25].

3.3.1. Selection. In the process of population evolution, an individual with a higher fitness value should be selected for the next generation of populations, while an individual with lower fitness value should be eliminated by the environment. The selection operation simulates the biological evolution that only the fittest can survive in nature, which is used to find individuals with larger fitness values in the population and inherit their excellent genes to the next generation.

The roulette wheel method is often adopted for the selection process, namely the selection is made by randomly rotating the wheel disc. Each individual is likely to be selected repeatedly. Usually, the number of selected individual chromosomes is determined by the proportion size of the fitness value of each chromosome. Let $P(z_j)$ be the probability that a chromosome z_j is selected for the next generation, as shown in the following equation:

$$P\left(z_j\right) = \frac{f\left(z_j\right)}{\sum_{j=1}^{2M} f\left(z_j\right)}, \tag{15}$$

where z_j is the chromosome in the population and $\sum_{j=1}^{2M} f(z_j)$ is the sum of fitness values of all chromosomes in the population.

1	0	0	-1	…	-1

Figure 3: A chromosome composed of K genes.

3.3.2. Crossover. Crossover operation combines two individual chromosomes in different ways to generate a new individual and improves the fitness of the new population by retaining the genes with better parents for the next generation of populations. The crossover operator recombines the genes of the parent populations to generate new chromosomes, introduce the diversity, and expand the search range of potential solutions [25].

The two-point crossover method is often used for the crossover operation in the genetic algorithms. The chromosomes are randomly selected with a certain crossover probability P_0 $(0 \le P_0 \le 1)$, and then two crossover points are randomly set at the corresponding position of the selected parent chromosomes. The column gene fragments from the first crossover point to the second crossover point are exchanged with each other to generate a new pair of offspring chromosomes. Such crossover operation can inherit the characteristics of the parent generation well while maintaining the diversity of the evolution.

3.3.3. Mutation. Mutation operation simulates gene mutation in the process of evolution. Mutations in biological evolution can enrich the diversity of new individuals and make the same species have different phenotypes. Although the crossover operation expands the search range, it is limited to various combinations of existing genes. The mutation operation makes individuals generate new genes, introduce new possibilities, and maintain the diversity of the populations [26].

Individuals in the population usually mutate with a certain mutation probability P_1 $(0 \le P_1 \le 1)$. Generally, individuals with lower fitness values have a larger mutation probability, while individuals with higher fitness values should be set a lower mutation probability to maintain the superiority of genes. In this paper, chromosomes are applied with 0-1 encoding while satisfying all constraints. Hence, for each chromosome in the population, the mutation operation specifically refers to the mutation of some columns of genes in the chromosome with a certain mutation probability (that is, 0 changes to 1 and 1 changes to 0). It should be noted that the position of the mutated gene is random, but the new generated individual should meet all constraints in order to ensure that it is still the feasible solution of the optimization problem, otherwise the mutation is failed and should be eliminated directly.

4. Experiment and Analysis

In the experiment, a network composed of 8 UAVs and 26 MTs is set up and the measured area is 10×10 km. MTs are randomly distributed in the test area, and their positions remain stable for a certain time. UAVs collaborate with mobile terminals to execute and transmit tasks, and the

coordinates of the base station are (−5, −5). The main simulation parameters are presented in Table 1.

In addition, the parameters of the improved genetic algorithm are set as follows: the maximum number of iterations is considered 200, the number of populations is 150, the crossover probability is 0.6, and the initial mutation probability is 0.01.

4.1. Effect of Mutation Probability on Algorithm Performance. The set of parameters of the improved genetic algorithm have a great influence on the performance of the proposed strategy, such as initial population size, mutation probability, and the number of genetic iterations. In order to obtain the best performance of the proposed strategy, the effect of mutation probability and the number of genetic iterations on the total cost is studied using the control variable method. The effect of mutation probability on the total cost is illustrated in Figure 4. At this time, the parameters of the improved genetic algorithm are set as follows: the initial population size is 150, the number of iterations is 200, the mutation probability varies from [0, 0.2], and the time delay and energy consumption weight in the fitness function are both 0.5.

As can be seen from Figure 4, when the mutation probability is within the interval [0, 0.08], the total cost initially shows a downward trend and then rises with the increase of mutation probability, and when the mutation probability exceeds 0.15, the total cost begins to fluctuate. With the increase of mutation probability, a better solution in the population may mutate into a worse solution, resulting in the degradation of performance. Therefore, the mutation probability is set at about 0.08 to achieve the optimal performance of the proposed strategy.

4.2. Effect of the Number of Genetic Iterations on the Algorithm Performance. Likewise, the effect of the number of genetic iterations on the total cost of the algorithm is illustrated in Figure 5. The specific parameters are set as follows: the initial population size 150, the mutation probability 0.08, the number of iterations increases from 1 to 200, and the time delay and energy consumption weight in the fitness function are set to 0.5.

As Figure 5 shows, although the total cost fluctuates with the increasing number of iterations, it still shows a decreasing trend. When the number of iterations reaches 80, the total cost curve is basically stable, reaching 0.66. Based on the above experiments, the genetic algorithm parameters are set as follows: the number of populations is 150, mutation probability is 0.08, and the number of genetic iterations is 200.

4.3. Relationship between the Number of Computation Tasks and the Average Time Delay. Figure 6 illustrates the performance of time delay with the increasing number of tasks arriving in the system when there are different numbers of UAVs.

It can be clearly shown in Figure 6 that no matter how many UAVs are running in the system, the overall average delay of the system is positively correlated with the arrival amount of ground terminal tasks. For any curve in Figure 6, it can be found that the growth rate of the curve rises with the increasing number of ground terminal tasks. As the M/M/1 queuing model is used to calculate the queuing time delay, a larger amount of tasks arriving in unit time causes the rapid rise of the overall delay of the system due to the limited computing capacity of the ground terminals and UAVs.

In addition, when the number of UAVs is increasing in the system, the growth of system delay becomes gentler, and the system delay is lower with more UAVs under the same number of arriving tasks. This means that the increasing number of UAVs reduces the load pressure of a single UAV as the tasks from the ground terminal can be offloaded partially to multiple UAVs. Each UAV does not need to deal with a large amount of tasks. Under the same computing capacity, it is obvious that the fewer the tasks offloaded to each UAV can lead to the lower the average delay and the lower the overall delay of the system.

4.4. Effect of the Number of UAVs on Total Utility under Different Strategies. Figure 7 depicts the comparison of the total utility of strategies proposed in this paper and in References [11, 12, 14] with the increasing number of UAVs when there are 20 mobile terminals in the system.

As can be seen in Figure 7, compared with other strategies, the total system utility of the proposed strategy is largest, and when the number of UAVs is 12, the time efficiency is highest, reaching about 1.83. When the number of UAVs is less than 12, the load on UAVS is larger, and consequently more tasks are transmitted to the base station, resulting in a lower utility. On the contrary, when the number of UAVs is greater than 12, the flight energy consumption of UAVs accounts for a big proportion leading to a lower utility. The proposed strategy uses the improved genetic algorithm to find the optimal solution for the system object and offload the computation task efficiently, jointly considering the energy consumption and time delay. The study in [14] selects the target MEC server for task offloading by using the fuzzy logic method. Although it reduces the dependence on the remote cloud, it only considers the fixed MEC servers and the applicability in the UAV cluster needs to be improved. The study in [12] proposed a flow splitting and aggregation scheme to decrease the offloading delay of the cloud server in edge computing and realize the task offloading. However, this scheme does not introduce any selection mechanism for the optimization algorithms, thus the overall utility is low, which is less than 0.95. The study in [11] formulated the task offloading problem into a nonlinear equation by using the Lagrange multiplier method and proposed the bisection search algorithm to achieve the optimal solution, which can effectively reduce the energy consumption. Nevertheless, the communication delay affects the overall utility.

Table 1: Main simulation parameters.

Parameter description	Value
Hovering height, m	10
Computation task data size, bits	$[10, 20, \ldots, 200] \times 10^3$
CPU cycle number, cycles·bit^{-1}	1000
Channel gain	1/2
Network bandwidth, MHz	5
Transmission power of MT, W	1.5
Noise power, W	2×10^{-13}
Transmission power of UAV, W	4
CPU frequency of MT, GHz	1
CPU frequency of UAV, GHz	10
Effective capacity coefficient	10^{-26}
Flight speed of UAV, m·s^{-1}	5

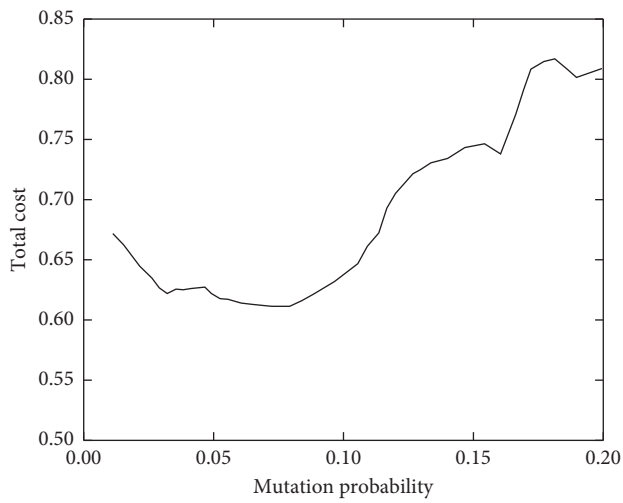


Figure 4: The effect of mutation probability on algorithm performance.

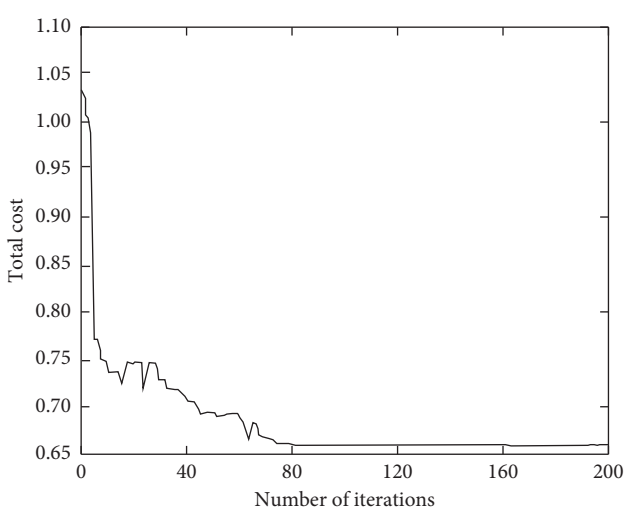


Figure 5: Effect of the number of genetic iterations on the algorithm performance.

4.5. Relationship between Application Size and Task Completion Time. In Figure 8, the MT task completion time of four offloading strategies under different task sizes is shown, in which the application in MTs is divided into four task groups and the data volume of each task group is randomly allocated.

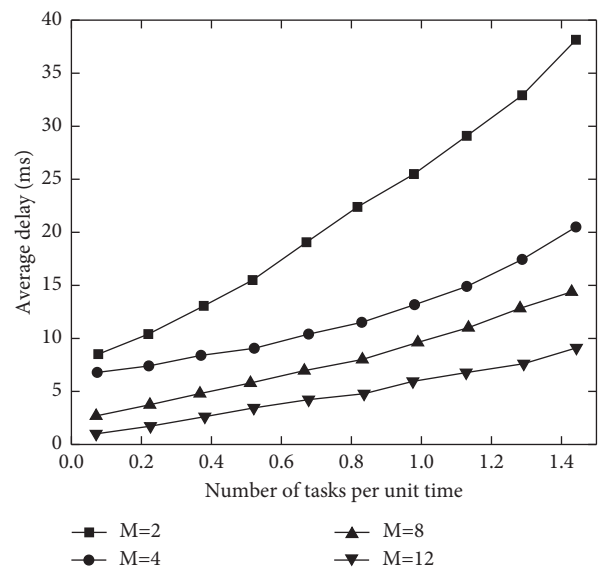


Figure 6: The relationship between the number of tasks and the average time delay under different numbers of UAVs.

As Figure 8 depicts, with the increase of data size, the task completion time also becomes higher, but the performance of the proposed strategy is always better than the other three strategies, and the maximum task completion time is about 110 ms. It should be noted that the performance gap between the proposed strategy and the strategy in Reference [14] becomes smaller and smaller as the application data size increases. In the proposed strategy, when the application data size is large, the UAV will choose to relay the task to the BS for execution and the UAV trends to fly to the places closer to the MT (or BS) to collect data (or transmit offloaded data), thus the communication channel conditions can be improved, thus reducing the task completion time. The proposed offloading scheme is similar to that in [14], which uses the fuzzy logic method, so the performance gap with task completion time is gradually narrowed. Similarly, the study in [11] proposed the bisection search algorithm to solve the offloading and resource allocation optimization problem. Although an ideal offloading scheme was selected, the algorithm is very complex and the computation and communication time are very long, taking more than 200 ms. The study in [12] uses the flow splitting and aggregation scheme to improve the offloading delay of the cloud server in edge computing, but it lacks an efficient optimization algorithm. Therefore, with the increase of application data volume, the task completion time soars, reaching more than 330 ms. Due to the complex task dependency between multiple sites of the UAV cluster, it is much more difficult to find the optimal task offloading scheme. Therefore, the proposed strategy adopts the improved genetic algorithm to solve this problem, taking energy consumption and delay into consideration

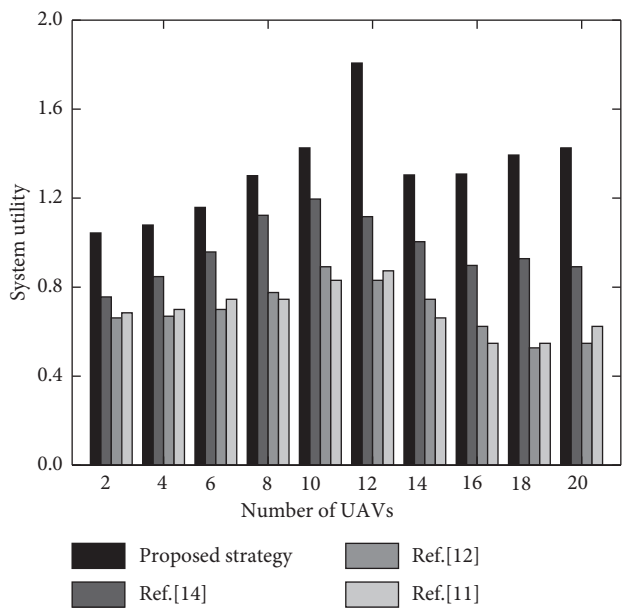


Figure 7: Effect of the number of UAVs on total utility under different strategies.

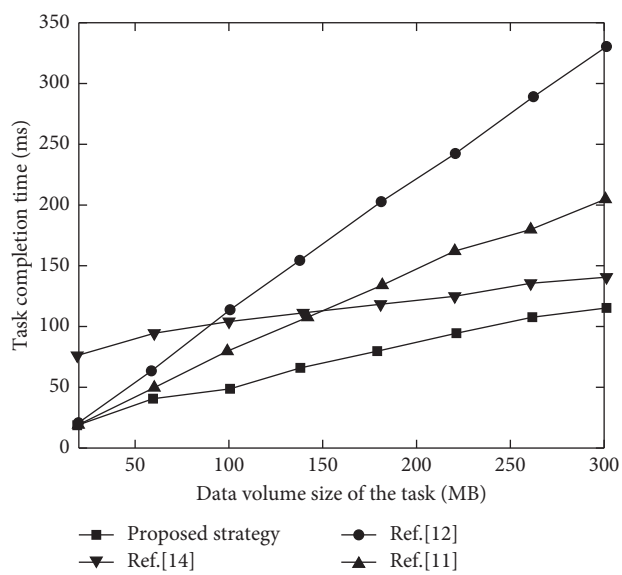


Figure 8: Relationship between application size and task completion time.

simultaneously, and finds the optimal offloading scheme with minimum target, which can reduce the system delay and energy consumption.

5. Conclusion

The emergence of edge computing has solved the problem of high network delay in the traditional cloud computing, but it needs to invest a lot of infrastructure construction. Therefore, a UAV cluster collaborate task offloading strategy using improved genetic algorithm in MEC is proposed, taking advantages of low cost and high mobility of UAVs. The improved genetic algorithm is used to solve the optimization function of minimizing the delay and energy consumption in the UAV cluster collaborate task offloading model, so as to obtain the optimal task offloading scheme. The simulation results conducted on MATLAB show that the improved genetic algorithm can improve the performance of offloading strategy. When the mutation probability is 0.08 and the number of genetic iterations is 200, the total cost of the system tends to be 0.66, and its energy consumption and delay reach the optimal value. The total utility is highest, reaching about 1.83 when the number of UAVs is 12, and the maximum task completion time is about 110 ms, which are both better than other comparison methods.

The proposed strategy does not consider that the tasks of different terminals or tasks in different periods of the same terminal may not be the same type and their requirements for computing resources or storage capacity of edge nodes are generally different. Therefore, we intend to classify the tasks into different types to meet the rapid growth of various application scenarios in the future work.

Data Availability

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

Conflicts of Interest

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

References

[1] Z. Song, Y. Liu, and X. Sun, "Joint task offloading and resource allocation for NOMA-enabled multi-access mobile edge computing," *IEEE Transactions on Communications*, vol. 69, no. 3, pp. 1548–1564, 2021.

[2] F. Wang, J. Xu, and S. Cui, "Optimal energy allocation and task offloading policy for wireless powered mobile edge computing systems," *IEEE Transactions on Wireless Communications*, vol. 19, no. 4, pp. 2443–2459, 2020.

[3] Q.-V. Pham, L. B. Le, S.-H. Chung, and W.-J. Hwang, "Mobile edge computing with wireless backhaul: joint task offloading and resource allocation," *IEEE Access*, vol. 7, no. 99, pp. 16444–16459, 2019.

[4] T. Zhang, Y. Xu, J. Loo, D. Yang, and L. Xiao, "Joint computation and communication design for UAV-assisted mobile edge computing in IoT," *IEEE Transactions on Industrial Informatics*, vol. 16, no. 8, pp. 5505–5516, 2020.

[5] Y. He, D. Zhai, F. Huang, D. Wang, X. Tang, and R. Zhang, "Joint task offloading, resource allocation, and security assurance for mobile edge computing-enabled UAV-assisted VANETs," *Remote Sensing*, vol. 13, no. 8, pp. 1547–1557, 2021.

[6] J. He, D. Zhang, Y. Zhou, and Y. Zhang, "A truthful online mechanism for collaborative computation offloading in mobile edge computing," *IEEE Transactions on Industrial Informatics*, vol. 16, no. 7, pp. 4832–4841, 2020.

[7] G. S. S. Chalapathi, V. Chamola, C.-K. Tham, G. Gurunarayanan, and N. Ansari, "An optimal delay aware task assignment scheme for wireless SDN networked edge cloudlets," *Future Generation Computer Systems*, vol. 102, no. 3, pp. 862–875, 2020.

[8] M. Mukherjee, S. Kumar, C. X. Mavromoustakis et al., "Latency-driven parallel task data offloading in fog computing networks for industrial applications," *IEEE Transactions on Industrial Informatics*, vol. 16, no. 9, pp. 6050–6058, 2020.

[9] U. Saleem, Y. Liu, S. Jangsher, X. Tao, and Y. Li, "Latency minimization for D2D-enabled partial computation offloading in mobile edge computing," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 99, pp. 4472–4486, 2020.

[10] X. He, R. Jin, and H. Dai, "Peace: privacy-preserving and cost-efficient task offloading for mobile-edge computing," *IEEE Transactions on Wireless Communications*, vol. 19, no. 3, pp. 1814–1824, 2020.

[11] L. Yang, G. Xu, J. Ge, P. Liu, and X. Fu, "Energy-efficient resource allocation for application including dependent tasks in mobile edge computing," *KSII Transactions on Internet and Information Systems*, vol. 14, no. 6, pp. 2422–2443, 2020.

[12] Y. Ito and H. Koga, "Improving offload delay using flow splitting and aggregation in edge computing," *IEICE Communications Express*, vol. 8, no. 12, pp. 468–473, 2019.

[13] B. Gu and Z. Zhou, "Task offloading in vehicular mobile edge computing: a matching-theoretic framework," *IEEE Vehicular Technology Magazine*, vol. 14, no. 3, pp. 100–106, 2019.

[14] M. D. Hossain, T. Sultana, V. Nguyen et al., "Fuzzy based collaborative task offloading scheme in the densely deployed small-cell networks with multi-access edge computing," *Applied Sciences*, vol. 10, no. 9, pp. 3115–3126, 2020.

[15] X. Chen, Z. Liu, Y. Chen, and Z. Li, "Mobile edge computing based task offloading and resource allocation in 5G ultra-dense networks," *IEEE Access*, vol. 7, no. 3, pp. 184172–184182, 2019.

[16] C. Yang, Y. Liu, X. Chen, W. Zhong, and S. Xie, "Efficient mobility-aware task offloading for vehicular edge computing networks," *IEEE Access*, vol. 7, no. 8, pp. 26652–26664, 2019.

[17] H. Wu, K. Wolter, P. Jiao, Y. Deng, Y. Zhao, and M. Xu, "EEDTO: an energy-efficient dynamic task offloading algorithm for blockchain-enabled IoT-edge-cloud orchestrated computing," *IEEE Internet of Things Journal*, vol. 8, no. 4, pp. 2163–2176, 2021.

[18] C. Li, W. Chen, J. Tang, and Y. Luo, "Radio and computing resource allocation with energy harvesting devices in mobile edge computing environment," *Computer Communications*, vol. 145, no. 9, pp. 193–202, 2019.

[19] A. M. Maia, Y. Ghamri-Doudane, D. Vieira, and M. Franklin de Castro, "An improved multi-objective genetic algorithm with heuristic initialization for service placement and load distribution in edge computing," *Computer Networks*, vol. 194, no. 4, pp. 108146–108156, 2021.

[20] X. Xu, Q. Liu, Y. Luo et al., "A computation offloading method over big data for IoT-enabled cloud-edge computing," *Future Generation Computer Systems*, vol. 95, no. 6, pp. 522–533, 2019.

[21] L. Tang, B. Tang, L. Kang, and L. Zhang, "A novel task caching and migration strategy in multi-access edge computing based on the genetic algorithm," *Future Internet*, vol. 11, no. 8, pp. 181–193, 2019.

[22] R. Ezhilarasie, A. Umamakeswari, M. S. Reddy, and P. Balakrishnan, "Grefenstette bias based genetic algorithm for multi-site offloading using docker container in edge computing," *Journal of Intelligent and Fuzzy Systems*, vol. 36, no. 3, pp. 2419–2429, 2019.

[23] H. A. Almashhadani, X. Deng, S. N. A. Latif, M. M. Ibrahim, and A. H. Alshammari, "An edge-computing based task-unloading technique with privacy protection for Internet of connected vehicles," *Wireless Personal Communications*, vol. 6, no. 1, pp. 1–22, 2021.

[24] L. Bing, 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.

[25] B. Huang, Z. Li, P. Tang et al., "Security modeling and efficient computation offloading for service workflow in mobile edge computing," *Future Generation Computer Systems*, vol. 97, no. 8, pp. 755–774, 2019.

[26] L. Ruiz, R. J. D. Barroso, I. De Miguel et al., "Genetic algorithm for holistic VNF-mapping and virtual topology design," *IEEE Access*, vol. 8, no. 7, pp. 55893–55904, 2020.

Hindawi
Journal of Robotics
Volume 2021, Article ID 3776440, 10 pages
https://doi.org/10.1155/2021/3776440

Research Article

Optimization of Computer Communication Monitoring System for Wind Turbine Speed

Zuoshan Li

School of Information Engineering, Suihua University, Suihua, Heilongjiang 152000, China

Correspondence should be addressed to Zuoshan Li; lixxgc@163.com

Received 12 November 2021; Revised 6 December 2021; Accepted 7 December 2021; Published 21 December 2021

Academic Editor: Shan Zhong

This article first studies the operating principles of wind turbines, focusing on the analysis of the structure and working principles of permanent magnet direct-drive wind turbines. According to the actual needs of the wind power system, the monitoring objects of the monitoring system are determined, and the overall monitoring plan for wind power generation is proposed to realize real-time analysis of the operating characteristics of the wind power system. At the same time, it pointed out the great significance of the wind power generation simulation experiment system and focused on the wind speed modeling. In terms of hardware research and analysis, relevant sensors, high-speed data acquisition cards, etc., were selected, and relevant signal conditioning circuits were designed, and a permanent magnet direct-drive wind power generation system simulation monitoring platform was constructed. In terms of software, LabVIEW was chosen as the design language of the monitoring system, and it pointed out the advantages of using LabVIEW in this monitoring system. Finally, the system uses the laboratory permanent magnet direct-drive wind turbine as the monitoring object. The practicality and accuracy of the system are verified through experiments such as permanent magnet motor power test, motor speed test, database system test, and remote monitoring test. The experimental results show that the monitoring system has a friendly interface and perfect functions and has important practicability and reference in the field of wind power monitoring.

1. Introduction

In the current society, with the continuous deterioration of the human living environment and the ever-increasing energy crisis, the development and utilization of renewable energy has attracted more and more attention from all countries in the world [1]. Wind energy is a clean renewable resource, and the development of wind energy is conducive to alleviating the current tension of energy shortage. Now that the installed capacity and power generation capacity of wind power are getting larger and larger, wind power is developing in the direction of expanding single-unit capacity, developing offshore wind power, and intelligent monitoring of wind farms [2]. There are two main measurement methods for speed measurement in the current industry: one is to convert the speed into an analog signal and measure the analog signal [3]. For example, a tachogenerator directly converts the speed into a voltage signal and then measures its voltage. The other is to use a sensor to detect periodic signals related to the rotational speed, convert this electrical signal into a pulse signal, and obtain the rotational speed pupil 1 by calculating the number of pulses within a specific time or calculating the time required for a fixed number of pulses. Finally, the speed is displayed, analyzed, and judged [4].

In rotating machinery, the measurement of engine speed is often involved, and the real-time and correctness of the speed measurement is required. Rotation speed detection and control also occupies a large proportion in the real-time control of industrial processes [5]. It has a vital influence on the stability of the system. In the application of generators, the motor speed is one of the important signs to judge the operation status of the motor. In the measurement and control of many motion systems, it is necessary to detect and control the speed of the motor to improve the performance of the control system and increase the accuracy of the system [6]. The accuracy of measurement directly affects the control status of the system, and only the high accuracy of the speed

can obtain a high-precision control system. Traditional methods for measuring speed include mechanical speed measurement method, magnetoelectric method, vibration speed measurement method, and photoelectric method. However, they are all based on single-chip microcomputer and sensors, which makes the user interface not friendly enough. Moreover, the speed measuring instruments using these measurement methods are not easy to communicate with the computer and are limited by the number of computer communication ports [7].

This article uses LabVIEW to structure the monitoring system and designs the system startup module, user login module, parameter configuration module, acquisition and display module, file I/O module, database system module, alarm system module, temperature system module, and signal analysis system module. Different from the conventional monitoring system, this system has better functions. It uses TDMS files and SQL server mixed programming to centrally manage system login, data storage, and data playback, which can realize real-time analysis of large amounts of data. From a theoretical point of view, because the CUSUM control chart considers the data of historical observation points, the detection effect of small drift with a trend is obviously better than that of the chart, which is the result of the further development of the control chart. This paper presents a wind turbine status monitoring method based on the CUSUM control chart. For the research on the overall condition monitoring of wind turbines, firstly, the method is used to obtain the power characteristics of the unit, and then the output power is monitored by the CUSUM control chart, so as to realize the condition monitoring of the whole machine. For the monitoring research of wind turbine operating conditions, firstly, the robust least-squares regression is used to establish the wheel speed prediction model under the normal operating mode. Finally, the standardized residual is used as the test statistic, and a CUSUM control chart based on the residual is established to monitor the operation status of the wind turbine, so as to realize the failure prediction. At the same time, the multitechnology integration of Data Socket and Remote Panels is used to realize remote monitoring of the system. Finally, the system uses the laboratory permanent magnet direct-drive wind turbine as the monitoring object. The practicality and accuracy of the system are verified through experiments such as permanent magnet motor power test, motor speed test, database system test, and remote monitoring test. PLC communication realizes real-time control of the motor. The experimental results show that the monitoring system has a friendly interface and perfect functions and has important practicability and reference in the field of wind power monitoring.

2. Related Work

Since the country's wind power industry started relatively late, the existing research results are mainly focused on the design of wind power generation units and large-scale power generation. The research in the field of wind turbine equipment status monitoring and fault diagnosis is still in its infancy, and the existing research results are relatively limited, mainly based on data-driven methods.

By analyzing the recorded index data, such as pressure, temperature, vibration, speed, power, etc., we predict the operating status of the wind turbine. For the evaluation of the operating status of the whole machine and the prediction of faults, Ali et al. [8] proposed a Gaussian mixture model parameter estimation to evaluate the health status of the operating wind turbine in real time. Herbert [9] applied fuzzy mathematics theory to establish a comprehensive evaluation model of unit operating status and used matter-element analysis to establish an evaluation method of grid-connected wind turbine operating status. Literature [10] uses the unit power curve obtained by the method to study unit performance. For units of the same configuration, according to the characteristics of installation and fault data, they construct random ending data of failure shutdown and use the Kaplan–Meier method to estimate the reliability of wind turbines by further fitting the w bull distribution to get the mean time between failures. Neyja et al. [11] used the maintenance times of the unit to obey the nonhomogeneous Poison process to obtain the future failure time distribution, thereby predicting the failure time. Based on the alum control chart of SPC, the wind turbine was researched on the failure of the wind turbine, which can detect the early failure of the equipment, but it is only effective for large drift detection, and is not suitable for the detection of small and medium drifts. It is difficult to find the failure trend of equipment earlier. For the research on the failure prediction of the key components of wind turbines, the existing research mainly includes support vector machines, neural networks, and multiple linear regression. Most of the basic ideas are to analyze the trend of residuals, so as to achieve the purpose of predicting failures. The online monitoring technology for judging whether the wind wheel is damaged mainly includes ultrasonic, vibration analysis, and acoustic emission. Among them, acoustic emission has good sensitivity to blade aging and damage, and you can find the location of abnormal areas. Yang et al. [12] extracted the sound source signal to obtain the initial crack characteristics and used the optimized wavelet redistribution scale spectrum analysis method to identify the crack initiation and re-expansion. The monitoring research on the balance of the wind wheel is mainly based on the electrical data of the generator. Dalala et al. [13] studied the impact of unbalanced faults on the power of a wind turbine, established a simulation model, and predicted the operating state of the wind turbine through spectrum analysis of the effective output power.

Developed countries in wind power generation, the United States, Denmark, Germany, etc., have achieved many results in unit design, condition monitoring, and failure prediction and maintenance. Merabet et al. [14] proposed using the Copula function to obtain the unit power curve probability model to effectively detect early faults and use the Hotelling statistical method to identify the operating status of the whole unit. Some scholars monitor and evaluate the status of wind turbines based on the artificial neural network of multiparameter fusion. For the key components of wind

turbines, there are mainly vibration analysis, lubrication analysis, thermal imaging, temperature, pressure, flow, oil, etc. [15]. Some researchers have proposed to use acoustic emission to monitor the position of the weakened area of the blade, monitor the running state of the wind wheel, use the spectral kurtosis to analyze the impact part of the vibration signal, and diagnose the fault [16]. Other researchers use generator power to predict faults and use continuous wavelet transform to extract the effective output power of generators for further condition monitoring and fault diagnosis. Many developed countries have accumulated a certain amount in the field of wind power monitoring [17]. In contrast, the country's wind power monitoring system is still in the exploratory stage at the technical level. Although the current imported monitoring system can also monitor the wind power production process, because domestic technicians have not mastered its core technology, once a problem occurs in the monitoring system, it cannot be solved in time and must rely on foreign technicians to troubleshoot the problem [18]. This not only delays production but also requires high costs for hiring foreign maintenance personnel, which will increase the burden on wind farms. In the event of a more serious situation, if foreign technical support is not available, the failure of the wind power equipment or system will directly lead to the paralysis of the wind farm [19]. The backwardness of domestic monitoring technology will constrain the development of China's wind power and make the entire wind power industry vulnerable. Nowadays, the management of wind power is gradually evolving towards digital operation, management information, and data centralization [20]. Many wind farms have multiple types of wind turbines. Due to different wind power management indicators, wind farms must be equipped with multiple sets of wind power monitoring systems, and staff must learn different systems, which directly increases the difficulty of maintenance and increases the cost of operation and maintenance [21].

3. Model Construction of Computer Monitoring System for Wind Power Generator Rotation Speed

3.1. Wind Energy Power Generation System Architecture. Wind turbine is a complex electromechanical equipment, which converts the wind energy passing through the turbine into electrical energy, and implements grid-connected power generation. The power generation process is mainly composed of the following three links: 1. primary energy conversion. A certain wind speed corresponds to the corresponding average power of wind energy. According to the knowledge of physics, the relationship between average power of wind energy and wind speed is as follows:

$$\frac{v(x)}{r(i)} + \frac{v(y)}{r(i)} + \frac{v(z)}{r(i)} = u(i). \tag{1}$$

The wind blown into the fan drives the wind wheel to rotate, which converts part of the wind energy into mechanical energy to generate mechanical power, where v is the wind energy utilization coefficient. 2. Mechanical energy transfer: the low-speed shaft drives the high-speed shaft to rotate, which is converted into the mechanical power of the high-speed shaft.

$$\overline{Y} \sim N\left(\mu - u\left(1 - \frac{\alpha}{2}\right), \frac{\sigma}{\sqrt{n}}\right). \tag{2}$$

For the needs of subsequent research, this article is based on the real-time measurement data of the SCADA system, according to the measurement method of the power characteristics of wind turbines provided by the International Electrotechnical Commission's IEC 61400-12 standard, using the method to estimate the mean and standard deviation.

The electrical power output by the wind turbine is closely related to the wind speed cut in during the operation of the wind turbine. Figure 1 shows the wind energy power generation system architecture. Therefore, by detecting and analyzing the output effective power and wind speed, the overall operating state of the wind turbine can be effectively monitored for potential faults. We select the SCADA monitoring data under the normal operation of the wind turbine, find the average value V of the environmental wind speed in each time t, where M is the amount of data monitored by SCADA in the time t,

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} r1 & 0 \\ 0 & rn \end{bmatrix} \times \begin{bmatrix} i \\ j \end{bmatrix} + p \times \begin{bmatrix} \alpha \\ \beta \end{bmatrix}, \tag{3}$$

$$u(x) = p \cdot (i, j) \cdot \sin(x) \cdot \cos(x), \tag{4}$$

$$\begin{cases} s(i) & = r(i) + p * \sin(i), \\ t(j) & = r(j) + p * \cos(j). \end{cases} \tag{5}$$

Since the effective output power of a wind turbine under normal operating conditions changes with the wind speed, the normal fluctuation range of the output power will be different under different wind speeds. Therefore, the traditional statistical process control technology is used. Simple single fixed value control can not effectively identify the moment of abnormal working state.

$$C(n) = \min(0, C(N-1) + k - \mu), \tag{6}$$

$$\begin{cases} C(n, 0) & = 0, \\ C(n) & = \min(0, p - \mu + k + C(n-1)) < 0. \end{cases} \tag{7}$$

Since the wind speed in each bin has little difference, the influence of wind speed in the same interval on the output power is ignored. When the wind turbine is in normal operation, its output electric power will fluctuate within a certain range. If there is an abnormal state or failure, the output power change will exceed the control line. Here, we use the gradient descent method to perform linear regression on the 123 groups of rotation speed sample curves in the normal startup state. Through observation, the curve of speed and time satisfies a polynomial.

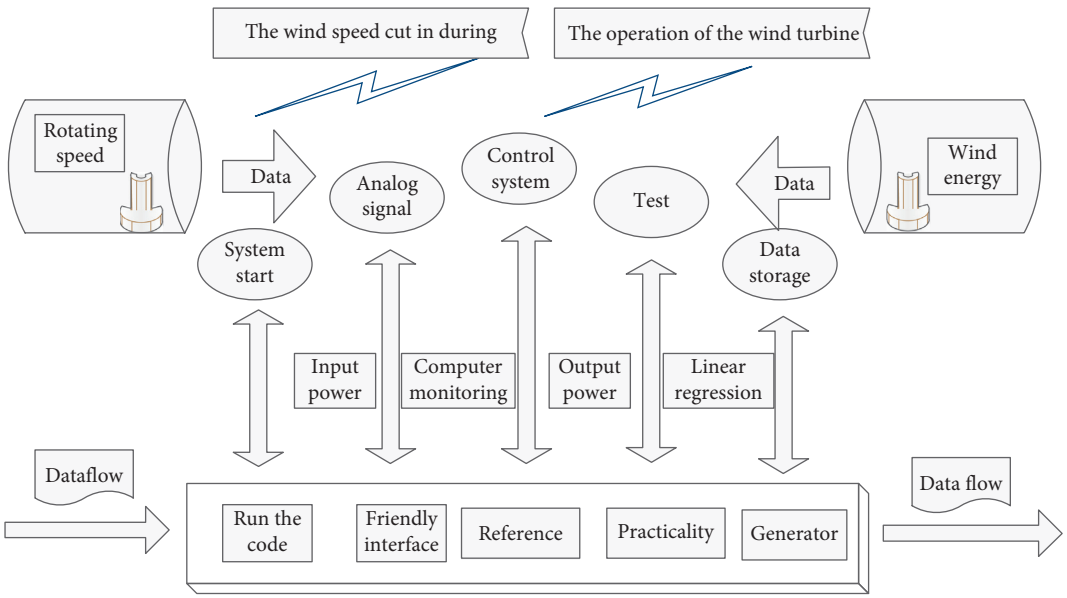


Figure 1: Wind energy power generation system architecture.

$$f(i, j) = \frac{1}{N} \times \sum_{j=1}^{N} \overline{p(j)} = \frac{1}{N(i, j)} \times p. \tag{8}$$

In the formula, v and p are the estimated speed and time, respectively; f is the coefficient of the power of i; N is the highest order of the polynomial. According to the analysis of sample data, a better fitting effect can be obtained when N is at least 5, and we will fit the speed-time curve with a fifth-order polynomial. It can be seen that the speed tends to stabilize after 50 s, so we focus on the range where the guide vane opening starts to increase until the unit speed reaches a stable value, that is, 5 to 50 s. We convert these samples into (t, v) data sample points, a total of $M = 123 * (50 - 5 + 1)$. Using the principle of minimum mean square error, the objective function is set to the following:

$$\frac{\partial g(x)}{\partial x} = -\sum \left\{ 2 \times (v - \overline{v(x)}) \times \sum_{i=0}^{n} t\hat{i} \right\}. \tag{9}$$

Among them, x and $g(x)$ are the actual value and estimated value of the unit speed, respectively. Our goal is to find a set of f so that L takes the minimum value. Using the principle of the gradient descent method, L calculates the partial derivative of $a0, \ldots, a5$, respectively, we can find the direction of the gradient descent, and use the gradient to iterate out the coefficient of the next cycle.

3.2. Computer Speed Monitoring Algorithm. This paper adopts robust regression, that is, adopts a robust estimation method in the regression model, which is intended to fit the structure of most of the data in the training data and accurately grasp the basic trend of the data. At the same time, we can also identify potential strong influence points, outliers, and structures that deviate from the model assumptions. In the range of use, robust estimation is more extensive. When the error of the variable obeys the normal distribution, the effect of the robust estimation is as good as the least square estimation. When there is a non-normal distribution and the least square estimation condition is not satisfied, the result of robust estimation will be far better than the least square estimation for defining different objective functions corresponding to different robust regression methods. The LAR method finds a curve that minimizes the absolute difference of the residuals, rather than the squared difference. Therefore, outliers have less influence on the fitting.

Then the Bellman equation shown can be used to obtain the optimal solution of the problem P1 through the classical value iteration algorithm or the strategy iteration algorithm. The bisquare weighting method is to minimize the weighted sum of squares, in which the weight assigned to each data point depends on the distance of the point from the fitted line. The points near the fitted line get all the weights. The weights of points that are farther offline gradually decrease. If the distance from the fitting line exceeds the preset range, its weight coefficient is zero. In most cases, the bisquare weighting method is better than LAR because it can identify outliers and minimize the influence of outliers while using most of the data to fit the curve. In fact, it is a robust least square method. Therefore, this paper adopts a robust regression based on bisquare to establish a model for predicting the rotation speed of the wheel. Figure 2 shows the computer speed monitoring algorithm flow.

Therefore, the study of whether the wind wheel fails or there is a potential failure is converted to the determination of the relationship between the wind speed and the hub speed, and the prediction model of the hub speed is established. By calculating the residual error between the predicted value and the actual measured value, and using it as a statistic, the CUSUM control chart is constructed.

These are perfect prior knowledge about renewable energy, computing tasks, channel status, etc. By judging whether the rotation speed of the hub is within the normal range under the wind speed of the environment where the wind turbine is located, the operation status of the wind turbine can be

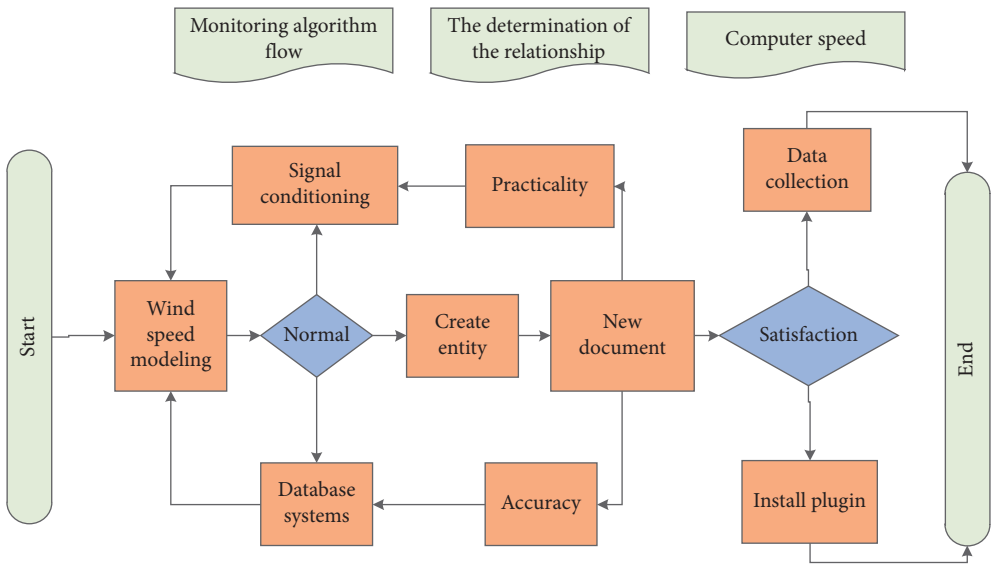


Figure 2: Computer speed monitoring algorithm flow.

detected, and then potential faults can be found. It can be seen that there is a functional relationship between wind speed and hub speed. According to the expression, the relationship should be a polynomial. We consider using the SCADA data of wind speed and hub speed under normal operation of wind turbines to conduct regression analysis and establish a prediction model for hub speed.

3.3. System Model Index Optimization. For analysis, this paper collected 125 sets of power generation start-up data for Unit 4 of Storage Power Station, including 2 sets of abnormal speed increase data caused by the damage of the guide vane opening feedback sensor and 123 sets of normal start-up data (including 2 sets of unit boot data under low head and 2 sets of boot data under high head). Through numerical analysis of the change law of unit speed with guide vane opening, time, and wind head, the curve fitting method is finally used to find the speed fitting curve with time as the independent variable and further use the idea of probability statistics to set a reasonable confidence interval to get the upper and lower thresholds.

However, in the real network environment, it is difficult to obtain as the main user of the agent. When the unit speed exceeds the threshold, it is determined that the guide vane opening feedback is faulty. This solution can detect the abnormal speed-up process of the unit caused by the damage of the guide vane opening feedback sensor in time, reduce the abnormal speed-up time of the unit, avoid the overspeed of the unit, and ensure the safe operation of the unit.

In normal operation, the speed of the hub of the same type of wind wheel is mainly affected by the intensity of the surrounding wind. The wheel speed value will be different under the same wind speed, and this fluctuation is mainly affected by random factors. If the wind wheel does not fail, the speed of the hub should fluctuate within the normal range. Once the wind wheel is icing, damaged, loose, and other faults, it will definitely affect the balance, the rotation of the wind wheel is affected, and the speed of the hub will become abnormal. Figure 3 shows the bar graph of residual speed of wind turbine generator. Therefore, it is only necessary to judge whether the fluctuation of the hub speed is within the normal range.

In order to facilitate the construction of CUSUM control charts, the statistical variables are standardized. Due to standardization, the wind wheel begs under normal conditions. Figure 4 shows the model framework of the wind power generator speed computer monitoring system. It should obey the standard normal distribution $t\,(0, 1)$. For the wind turbine, when its operation is abnormal or malfunctions, the speed of the hub will slow down, which will affect the performance of the entire wind turbine. According to the central limit theorem, the residual error between the predicted value of the wheel speed under normal conditions and the real measured value should obey the normal distribution under the influence of random factors.

According to the classic sample estimation method, the residual mean s and standard deviation of the sample data when the wind turbine is operating normally are obtained. We use MATLAB's Distribution Fitting T00l to test the normal distribution of the sample residuals and estimate that the mean value and the standard deviation as the overall mean and standard deviation account for them to predict the speed of the hub at the ambient wind speed at time f to get it. Therefore, the standardized residuals are used as statistics to construct a CUSUM control chart, and only the upward drift situation needs to be considered to construct a CUSUM control chart for judging upward drift.

4. Application and Analysis of Wind Power Generator Speed Computer Monitoring System Model

4.1. Wind Energy Generator Data Collection. In order to facilitate the understanding of the internal logic relationship, first simply qualitatively analyze the influence of the guide

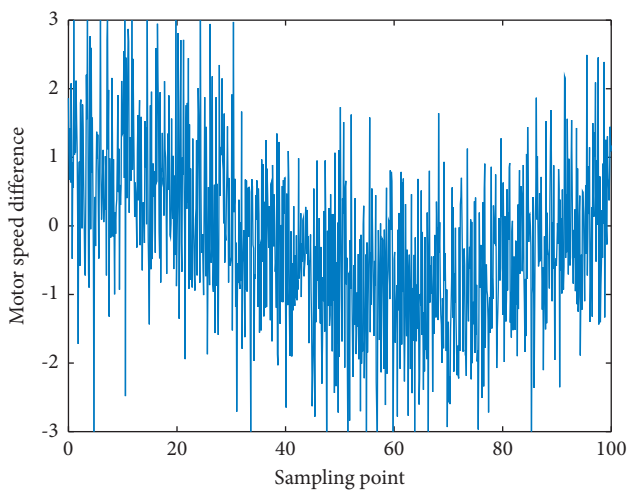


Figure 3: Bar graph of residual speed of wind turbine generator.

vane opening, wind head, and time on the speed of the unit. First of all, from the perspective of wind head, we compared 4 sets of engine speed data under different heads, which are the two highest wind heads and the two lowest wind heads in the 125 sets of data. It can be seen from the comparison results that the high head reaches the rated speed faster than the low head, and the arrival time is about 48 s and 50 s, respectively. However, the difference between the two sets of curves is small, even lower than the amount of sudden change caused by random disturbance. Therefore, under the high and low heads reached in our actual production, the change trend of the unit speed is basically the same. Next, we analyze the opening of the guide vanes. For the convenience of observation, we averaged the speed data of the 123 groups of units under the normal start-up state. Figure 5 shows the histogram of the average value of unit speed data.

In order to make the system run more stably and prevent contact malfunctions, the intermediate relay control method is used in the PLC program processing instead of directly controlling the PLC output, and then the output of QB0 is judged by the judgment instruction. Through the analysis of the system, we establish the switch value contact action truth table. According to the requirements of fan control, the switch will be performed only when the output of QB0 meets the following truth table, so that the program has a strong software protection capability, and the system is more safe and secure.

They include the system's state transition probability matrix and system cost function. We first analyze the variance of the sample data at each moment. Observation shows that the overall trend of variance is negatively correlated with time, and it will increase instantaneously when the guide vane opening changes suddenly. Because the variance reflects the uncertainty of the system, the greater the variance is, the greater the uncertainty of the unit's speed is. Therefore, the reasonable variance values should all be greater than the sample variance. In order to facilitate the calculation of the threshold, we select a linear function of time t above all standard deviations as the standard deviation. After successful log-in, nine fans and their numbers will be displayed to click on different fan components to enter the corresponding monitoring interface. The navigation bar at the bottom of the monitoring interface includes three buttons for status monitoring, dynamic performance analysis, and evaluation indicators. Among them, state monitoring mainly monitors some key parameter values during the operation of some wind turbines, such as power factor, generator speed, reactive power, active power, availability, failure time, total active power, wind speed, and other variables.

The real-time curve interface takes rotor speed as the research object, including wind speed curve, speed tracking curve, speed tracking error curve, and speed set value curve. Figure 6 shows the three-dimensional scatter fitting of motor wind speed curve. We click the corresponding button to view the corresponding real-time curve. The refresh interval is 1 second.

This curve shows the wind speed curve at low wind speeds to display the situation of actual speed tracking the set value. The curve shows that the actual value can track the set value well; the speed tracking error curve shows the curve of the difference between the actual speed and the set value; the speed set value curve displays the size of the speed setting value. The change of real-time data exceeds the interval variation range of the production data ±2, so it is necessary to use the data transmission module to write the program to release the production data with the serial number 05; when the serial number ID = 02, the identification is 01, 02, the value has not changed, the serial number is 06, the amount of data change does not exceed the set change range of ±0.01%, so there is no need to update, according to this method, the other obtained production whose data are judged and released in sequence to complete the real-time monitoring of the operation status of the wind turbine.

4.2. Computer Monitoring Performance Simulation. This paper selects a wind power generator set of a wind farm as the research object. Its SCADA system records the parameter data (ambient wind speed, wheel speed, effective output power, generator temperature, etc.) of the generator every 1 s and records that the wind turbine is in abnormal operation status information. On the one hand, in order to increase the time span and improve the accuracy, on the other hand to avoid unnecessary tedious calculations, the SCADA data of this type of unit with a sampling time of 1 s in one month in 2019 were selected in a total of 2592,000 groups. According to the wind turbine fault record table of the SCADA system, the typhoon motor operated abnormally due to heavy rain in the early morning on September 13 of a certain year, and the wind leaves were not covered by a large amount of water droplets. The time when the company's monitoring system detects the fault is t. From the above comparison, we can get the alarm by using the CUSUM control chart, which is 20 minutes earlier than the control chart. It can be seen that in the event of a wind turbine operating failure, the CUSUM control diagram can detect abnormal operating conditions earlier than the diagram, predict potential failures earlier, and facilitate further

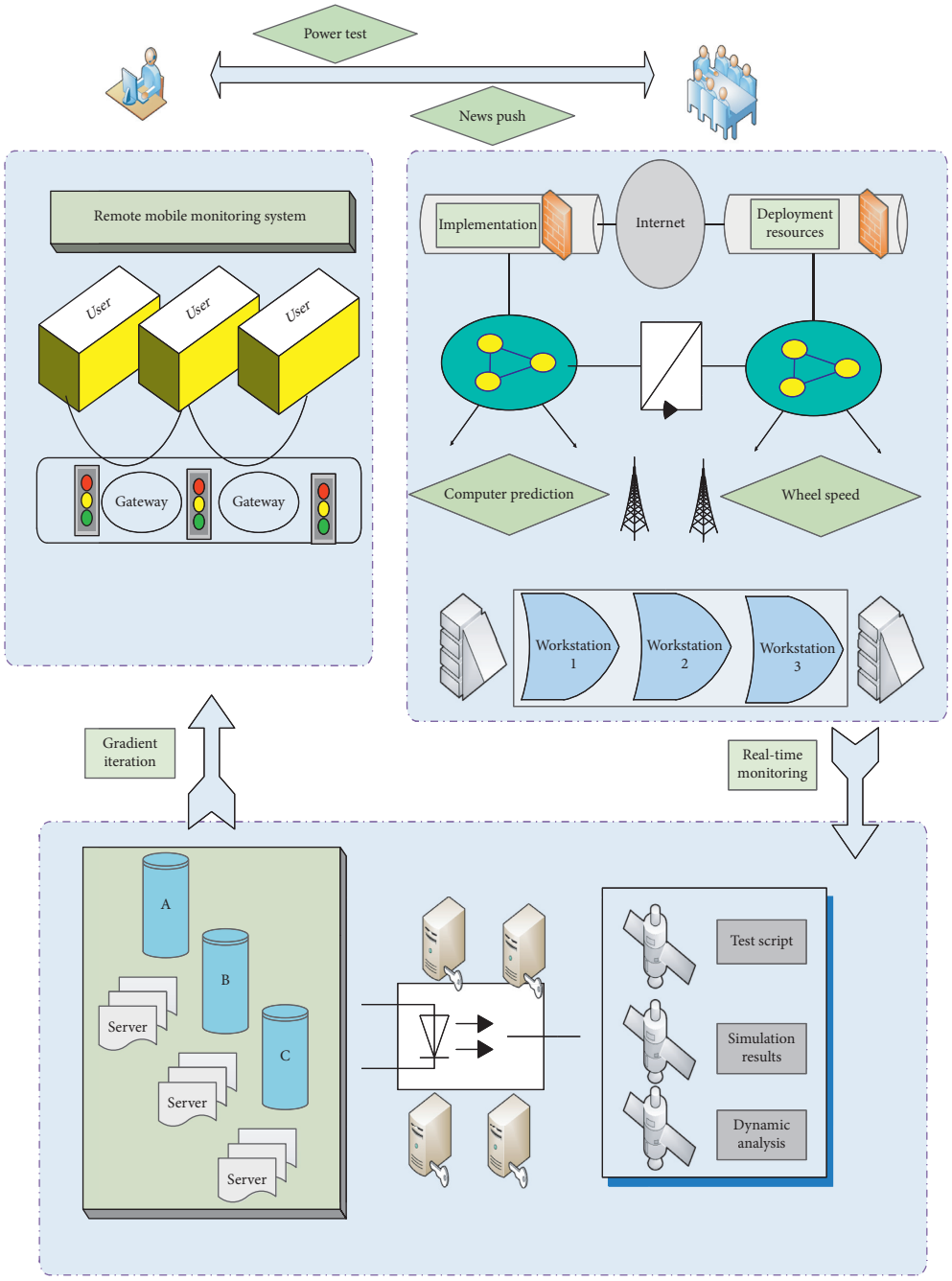


Figure 4: The model framework of the wind power generator speed computer monitoring system.

maintenance. In order to avoid accidental calculation of a single set of data, the SCADA data of this type of wind turbine before two failures were selected for calculation and obtained separately.

According to the unit setting, the guide vane opening change process is divided into four stages. In the first stage, the guide vane opening rapidly opens to 14%, which corresponds to between 0 and 7 s. The unit speed increases and the acceleration gradually increases. Then it enters the holding phase of about 30 s. The speed of the unit continues to increase, but the acceleration gradually decreases. At about 40 s, the guide vane opening increased to about 21% again in a short time.

Finally, the opening of the guide vane dropped back and remained at 17%, and the speed of the unit no longer increased and remained at about 100%. Figure 7 shows the line graph of change of correlation coefficient of unit speed with time. Therefore, the guide vane opening degree determines the final stable value of the speed. It can be seen that the change curve of the speed of 125 sets of units with time includes 2 sets of fault data. It can be seen from the figure that there is a continuous functional relationship between the speed and time, and the downward trend of different heads is consistent. The guide vane opening of the system is set fixed, and the actual applied head range has little

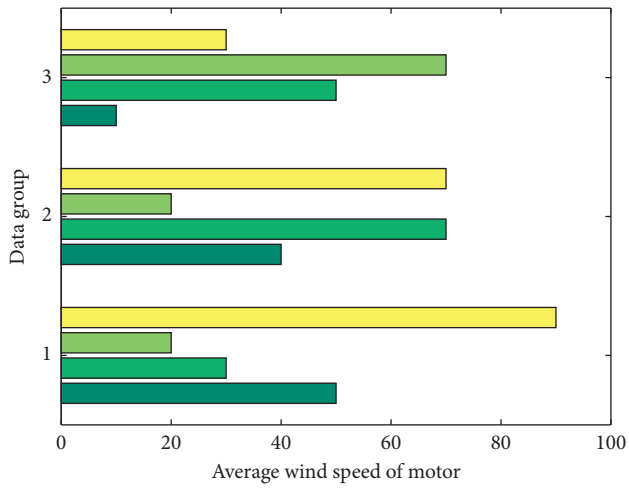


Figure 5: Histogram of the average value of unit speed data.

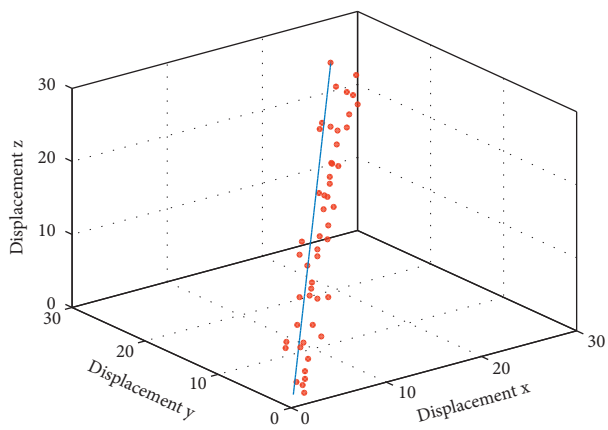


Figure 6: Three-dimensional scatter fitting of motor wind speed curve.

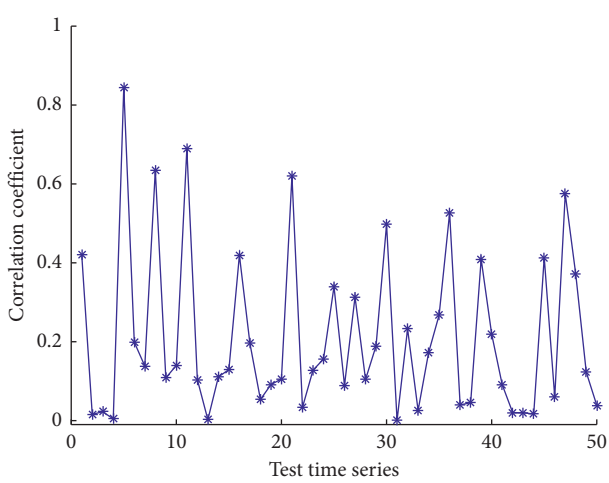


Figure 7: Line graph of change of correlation coefficient of unit speed with time.

influence on the trend of the speed curve. Therefore, fitting the function of the speed with time can effectively reflect the trend of the current system speed. For the SCADA data of the selected one-month wind turbine in normal operation, after the first step of the process, the wind speed is classified by 0.5 m/s, and the average value of the wind speed and output power in each interval is calculated.

4.3. Example Results and Analysis. We select the data recorded by the SCADA system of the same type of wind turbine for two consecutive weeks, and process the data stream according to the aforementioned process to obtain the value, and plot the last calculated 2019 data in the paper to obtain the CUSUM control chart. The upper and lower thresholds can completely cover the speed curves of 123 units under normal conditions without misjudgment.

Based on the above analysis, the main user has prior knowledge of the network environment. At the same time, we can also observe two sets of fault curves, which exceed the upper threshold and can be distinguished. And the speed is less than 60% of the rated speed when the boundary is exceeded for the first time. Figure 8 shows the deviation threshold curve for abnormal changes in unit speed. Therefore, real-time comparison between the current speed and the upper and lower speed threshold curves given by equations can quickly identify the abnormal speed increase of the unit in 99.99% of the cases.

In practical applications, a set of antioverspeed logic can be set in the monitoring system: during the speed-up phase of the unit, when the speed of the unit exceeds the range of the upper and lower speed threshold curves given by formulas, the output of the monitoring system increases abnormally. Speed alarm: this alarm can trigger an accident shutdown when the alarm is maintained for 2 s without reset, and it can identify and prevent the abnormal speed increase of the unit in time to avoid overspeed.

In order to prove that the CUSUM diagram can find that the wind turbine is in an abnormal operating state before the failure occurs, we select the data recorded by the SCADA system in the previous period of time when the failure occurred, calculate the CUSUM control chart, respectively, and compare it with the commonly used control chart and company monitoring system. Figure 9 shows the histogram of the fluctuation interval of the fan operating state. According to that, it can be seen that there was a small range of downward fluctuations before, but it quickly returned to the positive direction, indicating that the fan is operating in good condition, and it was always below the control line. According to the corresponding judgment criteria, the wind turbine was abnormal at 10 times. It can be seen that the CUSUM control chart is 5.12, and the exceeding control line is 5. It is at 50 minutes and keeps drifting high; it means that the wind turbine is malfunctioning. According to the aforementioned judgment principle, the moment when the wind turbine fails is the point closest to C; that is, the wind turbine starts to be abnormal.

During the collection process, according to the Kafka and incremental data transmission scheme, production data needs to be collected at regular intervals. Figure 10 shows the simulation results of the feedback signal value of the fan with the number of data acquisitions. First, we set the time interval for data collection.

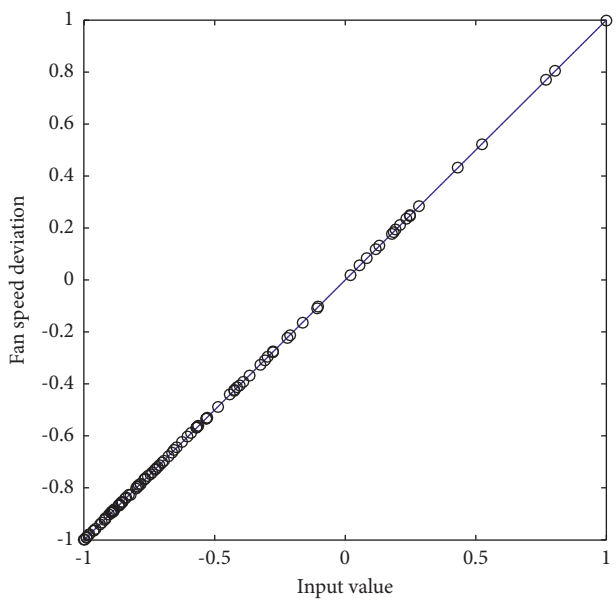


Figure 8: Deviation threshold curve for abnormal changes in unit speed.

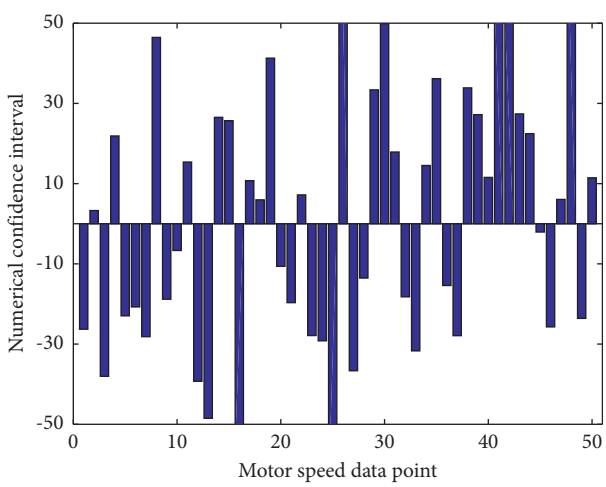


Figure 9: Histogram of the fluctuation interval of the fan operating state.

Due to the consideration of the number of digits in the number, the number is encoded in a cyclic processing method, and the threshold of the number of acquisitions is set. If the number of acquisitions reaches the threshold, the number of acquisitions will be cleared and counted again; according to the number of production data collected each time, the data items of each production data are numbered. The numbered group production data format is shown in paper. By comparing the difference between the set value and the feedback signal, the error is obtained, and then the error is transmitted to the controller. The controller uses a certain algorithm for calculation and analysis, and outputs control commands, adjusts the controlled object, and realizes the control of the speed. The speed can track the set value well, work in the best performance area, and realize the maximum wind energy capture by monitoring the speed.

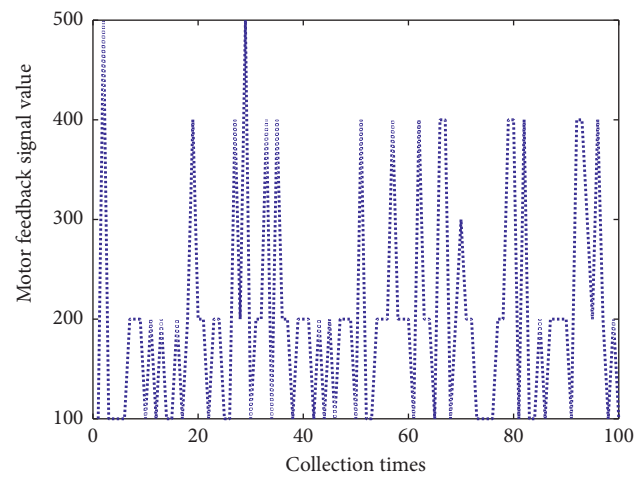


Figure 10: The simulation results of the feedback signal value of the fan with the number of data acquisitions.

5. Conclusion

This article focuses on the method of turning angle measurement. In order to measure the rotation angle, the relationship of the rotating body in the space coordinate system was studied, and a rotating machine speed measurement model was established. It is no longer a problem in modern science to measure small angles, but this is all at the expense of complex and precise mechanical structures, and these angles have no directionality at all. This article uses a USB camera and a PC to collect and process images and use the geometric vector method to obtain the rotation angle between two frames. The noncontact detection method in this paper avoids the errors caused by the traditional contact method due to mechanical wear, slippage, and vibration. At the advantage of no pollution or damage, they can be used in electronic detection and industrial control for rapid and accurate measurement of speed and rotation angle. With the further development of CCD technology, the influence of this limiting factor will inevitably become smaller and smaller. The upper limit of the measurement speed is 1200 r/min, and they can obtain accurate measurement results. Experiments show that the method used to measure the rotational speed of the analysis object has high accuracy, strong anti-interference ability, high cost performance, and simple system structure. Finally, this paper takes the actual operation monitoring data of a certain wind farm as an example for empirical analysis and compares it with the control chart. The empirical analysis results show that the CUSUM control chart based on SCADA data is an improvement of the alum chart, which can realize the detection of the operation status of the wind turbine and find the abnormality of the whole machine and the wind turbine. Through the research in this paper, the process of maintenance can be accelerated, equipment damage accidents can be avoided, and the safe and efficient operation of wind farms is beneficial.

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 Innovation and Entrepreneurship Training Program for Undergraduate in Heilongjiang Province (no. 202110236029) and Basic Scientific Research Business Expenses Projects of Provincial Universities in Heilongjiang Province (no. KYYWF10236190204).

References

[1] A. Kusiak, Z. Zhang, and A. Verma, "Prediction, operations, and condition monitoring in wind energy," *Energy*, vol. 60, pp. 1–12, 2013.

[2] C. L. Kana, M. Thamodharan, and A. Wolf, "System management of a wind-energy converter," *IEEE Transactions on Power Electronics*, vol. 16, no. 3, pp. 375–381, 2001.

[3] W. Qiao, P. Zhang, and M. Y. Chow, "Condition monitoring, diagnosis, prognosis, and health management for wind energy conversion systems," *IEEE Transactions on Industrial Electronics*, vol. 62, no. 10, pp. 6533–6535, 2015.

[4] F. P. G. Márquez, A. M. Tobias, J. M. P. Pérez, and M. Papaelias, "Condition monitoring of wind turbines: techniques and methods," *Renewable Energy*, vol. 46, pp. 169–178, 2012.

[5] L. Qihui, H. Yikang, and Z. Rende, "The maximal wind-energy tracing control of avariable-speed constant-frequency wind-power generation system," *Automation of Electric Power Systems*, vol. 20, no. 27, pp. 62–67, 2018.

[6] S. M. Dehghan, M. Mohamadian, and A. Y. Varjani, "A new variable-speed wind energy conversion system using permanent-magnet synchronous generator and $ Z $-source inverter," *IEEE Transactions on Energy Conversion*, vol. 24, no. 3, pp. 714–724, 2009.

[7] E. Koutroulis and K. Kalaitzakis, "Design of a maximum power tracking system for wind-energy-conversion applications," *IEEE Transactions on Industrial Electronics*, vol. 53, no. 2, pp. 486–494, 2006.

[8] M. Ali, S. Qaisar, M. Naeem, and S. Mumtaz, "Energy efficient resource allocation in D2D-assisted heterogeneous networks with relays," *IEEE Access*, vol. 4, pp. 4902–4911, 2016.

[9] G. M. J. Herbert, S. Iniyan, and E. Sreevalsan, "A review of wind energy technologies," *Renewable and Sustainable Energy Reviews*, vol. 11, no. 6, pp. 1117–1145, 2017.

[10] Y. Zhao, C. Wei, Z. Zhang, and W. Qiao, "A review on position/speed sensorless control for permanent-magnet synchronous machine-based wind energy conversion systems," *IEEE Journal of Emerging and Selected Topics in Power Electronics*, vol. 1, no. 4, pp. 203–216, 2018.

[11] M. Neyja, S. Mumtaz, K. M. S. Huq, S. A. Busari, J. Rodriguez, and Z. Zhou, "An IoT-based E-health monitoring system using ECG signal," in *Proceedings of the GLOBECOM 2017 - 2017 IEEE Global Communications Conference*, pp. 1–6, Singapore, 2017.

[12] W. Yang, P. J. Tavner, C. J. Crabtree, and M. Wilkinson, "Cost-effective condition monitoring for wind turbines," *IEEE Transactions on Industrial Electronics*, vol. 57, no. 1, pp. 263–271, 2019.

[13] Z. M. Dalala, Z. U. Zahid, W. Yu, Y. Cho, and J.-S. Lai, "Design and analysis of an MPPT technique for small-scale wind energy conversion systems," *IEEE Transactions on Energy Conversion*, vol. 28, no. 3, pp. 756–767, 2013.

[14] A. Merabet, K. T. Ahmed, H. Ibrahim, R. Beguenane, and A. M. Y. M. Ghias, "Energy management and control system for laboratory scale microgrid based wind-PV-battery," *IEEE transactions on sustainable energy*, vol. 8, no. 1, pp. 145–154, 2017.

[15] P. Tchakoua, R. Wamkeue, M. Ouhrouche, F. Slaoui-Hasnaoui, T. Tameghe, and G. Ekemb, "Wind turbine condition monitoring: state-of-the-art review, new trends, and future challenges," *Energies*, vol. 7, no. 4, pp. 2595–2630, 2019.

[16] A. Salem, A. Abu-Siada, and S. Islam, "A review of condition monitoring techniques of the wind turbines gearbox and rotor," *Computer and Electrical Engineering*, vol. 3, pp. 11–19, 2019.

[17] K. Protsenko and D. Xu, "Modeling and control of brushless doubly-fed induction generators in wind energy applications," *IEEE Transactions on Power Electronics*, vol. 23, no. 3, pp. 1191–1197, 2008.

[18] Z. Zhou, C. Zhang, J. Wang et al., "Energy-efficient resource allocation for energy harvesting-based cognitive machine-to-machine communications," *IEEE Transactions on Cognitive Communications and Networking*, vol. 5, no. 3, pp. 595–607, 2019.

[19] J. A. Barrado-Rodrigo, J. I. Talpone, and L. Martinez-Salamero, "Variable-speed wind energy conversion system based on a dual stator-winding induction generator," *IET Renewable Power Generation*, vol. 11, no. 1, pp. 73–80, 2017.

[20] A. Pliego Marugan, F. P. Garcia Marquez, and J. M. Pinar Perez, "Optimal maintenance management of offshore wind farms," *Energies*, vol. 9, no. 1, p. 46, 2019.

[21] D. J. Willis, C. Niezrecki, D. Kuchma et al., "Wind energy research: state-of-the-art and future research directions," *Renewable Energy*, vol. 125, pp. 133–154, 2018.

Hindawi
Journal of Robotics
Volume 2021, Article ID 9326695, 9 pages
https://doi.org/10.1155/2021/9326695

Research Article

Expression Recognition Method Using Improved VGG16 Network Model in Robot Interaction

Shengbin Wu

School of Information Engineering, Changsha Medical University, Changsha, Hunan 410219, China

Correspondence should be addressed to Shengbin Wu; shengbinwu123456@163.com

Received 12 November 2021; Accepted 4 December 2021; Published 20 December 2021

Academic Editor: Shan Zhong

Aiming at the problems of poor representation ability and less feature data when traditional expression recognition methods are applied to intelligent applications, an expression recognition method based on improved VGG16 network is proposed. Firstly, the VGG16 network is improved by using large convolution kernel instead of small convolution kernel and reducing some fully connected layers to reduce the complexity and parameters of the model. Then, the high-dimensional abstract feature data output by the improved VGG16 is input into the convolution neural network (CNN) for training, so as to output the expression types with high accuracy. Finally, the expression recognition method combined with the improved VGG16 and CNN model is applied to the human-computer interaction of the NAO robot. The robot makes different interactive actions according to different expressions. The experimental results based on CK + dataset show that the improved VGG16 network has strong supervised learning ability. It can extract features well for different expression types, and its overall recognition accuracy is close to 90%. Through multiple tests, the interactive results show that the robot can stably recognize emotions and make corresponding action interactions.

1. Introduction

Different facial expressions can reflect people's emotional and psychological changes in different situations. Therefore, expression recognition has very important research significance and practical application value for the study of human behavior and psychological activities. In recent years, with the rapid development of computer vision, the rise of deep learning, the improvement of machine learning, and other related theoretical systems, facial expression, as a bridge of human-computer interaction, has attracted the attention of researchers at home and abroad. Facial expression recognition (FER) technology has developed rapidly [1, 2]. Different methods are used to extract facial expression feature information, and different methods are used to recognize and classify facial expressions, so as to make better use of emotional information and apply expression recognition technology to actual production, work, and life. With the development of deep learning, FER has gradually become one of the hot technologies in the field of human-computer interaction technology. FER has been widely used in human-computer interaction, business, medical treatment, entertainment, psychology, fatigue driving, and other fields [3, 4].

With the prevalence of artificial intelligence and pattern recognition, FER has encountered many practical problems in the application process, such as the sharp increase of data leads to the reduction of recognition effect. Therefore, researchers need to constantly improve algorithm to improve its recognition rate and robustness. There are two traditional FER methods: geometry-based method and whole-based recognition method [5]. These two methods rely on the advantages and disadvantages of the previous manual feature extraction, are subject to many interference factors in feature extraction, and are not realistic for massive image data processing [6]. With the continuous upgrading of computer technology, the improvement of computing power and communication ability has promoted the development of machine learning, especially, deep learning has made outstanding contributions to facial expression recognition [7, 8]. However, at present, most expression recognition methods based on deep learning still have great room for improvement in training samples and recognition

efficiency. Therefore, this paper proposes an expression recognition method based on the improved VGG16 network model.

2. Related Works

Researchers at home and abroad have carried out research studies on facial expression recognition. Traditional expression recognition methods mainly include geometry-based and overall FER methods. Its main advantage is that it is simple and easy to implement, but there is still room for improvement in recognition accuracy and efficiency [9]. For example, Madzin et al. [10] proposed a FER method based on geometric skin color to accurately detect faces in real-world images, which can adapt to image face recognition under different indoor and outdoor lighting environments and complex background conditions. However, the extraction and recognition of facial expression features cannot be carried out at the same time, which greatly limits the recognition efficiency.

Different from the traditional recognition methods, deep learning breaks the fixed pattern of pattern recognition after traditional feature extraction, and can carry out feature extraction and expression classification at the same time. Moreover, the feature extraction of deep learning iteratively updates the weights through back-propagation algorithm and error optimization, so it can extract key points and features unexpected to human beings [11, 12]. Chen et al. [13] proposed a simplified face clipping and rotation strategy combined with the image recognition method of convolutional neural network (CNN), which ensures the richness of data while extracting facial features and considers the accuracy and richness of expression recognition, but the training time for some complex expressions is long. Xu et al. [14] proposed a microexpression recognition method based on the Wasserstein generative adversarial network, established facial expression recognition network and facial identity recognition network, and improved the accuracy and robustness of facial expression recognition by suppressing intraclass changes. Lee et al. [15] converted the expression image into a local binary pattern (LBP) feature map and then used the LBP feature map as the input of CNN for training, which had achieved good results. However, it will lead to low accuracy and insufficient robustness in unknown environment. Bunyak et al. [16] proposed an automatic screen expression recognition system based on deep learning. On the basis of extracting the basic features of facial expressions, the original facial images are used to classify facial expressions and icon facial images to realize real-time recognition of facial expressions, but the overall accuracy needs to be improved. For facial expression recognition, Shao and Qian [17] proposed three facial recognition frameworks: shallow CNN, dual branch CNN, and pretrained CNN. The results show that they have good practicability and effectiveness. Unwala et al. [18] proposed a facial expression recognition system based on depth CNN. This paper constructs a dual channel CNN model and gathers the channel feature information into the full-connection layer for expression classification. The results show that it obtains high expression recognition accuracy, but it needs large training dataset and long training time in the early training process, so the recognition efficiency is not high.

In view of the low recognition accuracy of most existing expression recognition methods, this paper proposes an expression recognition method using the improved VGG16 network model, which improves the overall effect of expression recognition on the basis of taking into account the accuracy and recognition efficiency.

3. Proposed Method

3.1. Overall Network Structure. Deep learning perfectly solves the problem of automatic feature extraction and has a certain effect on face recognition and prediction. However, in the field of face expression recognition, the existing implementation effect does not reach the level required by the production environment [19]. With the continuous development of deep learning theory and structure research, this paper proposes an improved VGG16 model for facial expression recognition and makes in-depth research on expression feature extraction to effectively solve the problems of weak feature representation ability of machine learning extraction. The overall network structure is mainly composed of VGG16 and CNN. VGG16 is used for feature extraction, and CNN is used for classification and recognition. The specific structure is shown in Figure 1.

Compared with other deep learning methods for feature extraction, the proposed method uses the VGG16 network for feature extraction, which uses a large number of small-size convolution and pooling operations, so that VGG16 network can extract more implicit features, then uses CNN network for facial expression model training, and finally uses Softmax classifier for classification and recognition.

3.2. Data Preprocessing. In order to prevent the overfitting phenomenon of the network model, the data are randomly turned horizontally and cut or rotated at an angle during the training process. This method is called data enhancement. The data enhancement methods of horizontal, clipping, and angular rotation are adopted to enable the network model to learn the expression features from multiple angles, so as to enhance the robustness of the network model [20]. Data reduction can expand the number of images in the dataset, so that the neural network can learn more comprehensive feature information. Therefore, the dataset is usually randomly reduced in the experiment.

The size of the facial expression image after the facial expression data set CK + alignment is moderate, so the image is not enlarged in the experiment to retain the original information of the image to the greatest extent. When training, the original 100×100 images are randomly reduced to 90×90 size sub-image, then the image is randomly mirrored, and the processed image is sent to the network model for training. At the same time, in the test stage of the experiment, ten times reduction is used to expand the test image set to improve the expression recognition effect of the

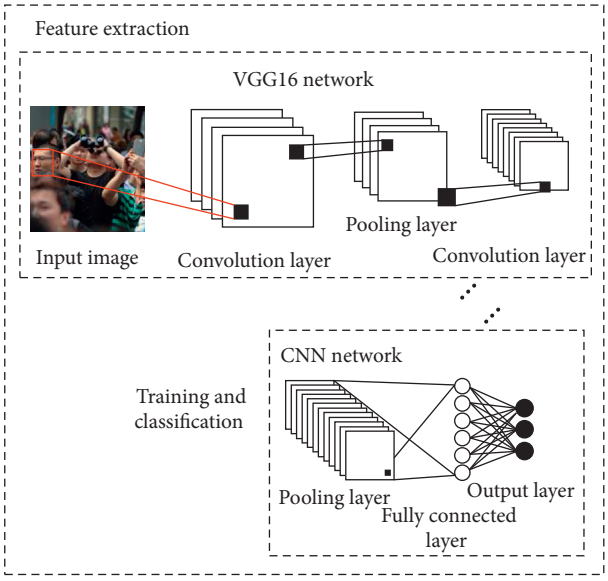


Figure 1: Overall framework of the proposed method.

network model, and experiments are used to verify the impact of ten times reduction on the expression recognition rate of the network model. Tenfold reduction refers to the reduction of the image in the upper left corner, lower left corner, upper right corner, lower right corner, and center, and the mirror operation is used at the same time. The tenfold reduction operation expands the number of input images by 10 times. The reduced image input model is tested, and the probability of 10 images output by the network model is calculated to obtain a mean value, which is used as the final classification result of the input image. This operation mode can effectively reduce the classification error rate of the network model.

The original input size of VGG16 network model is 224×224. If the image resolution is too small, VGG16 will not be able to fully extract feature information in feature extraction process, and if the image resolution is too large, it will increase the computational burden of computer memory during network model training [21, 22]. Therefore, in order to maintain the image resolution, the computer memory load is too heavy, resulting in the failure of the experiment. The proposed method enlarges the facial expression image of the dataset to 128×128, and then randomly reduce the image to 112×112. Finally, the image is randomly mirrored and input into the network model for training.

3.3. Improved VGG16 Model Feature Extraction. In the VGG16 network, due to the continuous use of small convolution cores for many times and the multiple growth of convolution cores in each layer, the number of corresponding output feature maps becomes more and occupies more graphics card space [23]. Experiments on the original model VGG16 show that a very large amount of parameters will be generated on the first fully connected layer, which makes the amount of calculation huge and consumes more computing resources. In addition, due to the size constraints of the dataset, the small- and medium-sized data samples do not perform well in the deep network, and the final experimental results are far lower than expected. Part of the reason is the overfitting problem caused by the small data scale, which leads to the insufficient generalization ability of the model and fails to reflect the original excellent performance of the depth network VGG16. Reducing the depth of the neural network in different ways to reduce the amount of parameters is helpful to prevent overfitting to a certain extent [24].

Inspired by GoogleNet and AlexNet, large convolution is used to reduce the dimension directly on the high-level feature map, which does not produce too much calculation. Moreover, the continuous large convolution kernel replaces the small convolution to reduce the complexity of the model, further compress the number of parameters, and reduce part of the full-connection layer, which will not affect the expression of the feature layer, but reduce the number of parameters [25]. Therefore, VGG16 structure is improved from two aspects: (1) The 5×5 convolution kernel is used on the larger feature map of the initial layer, and the 3×3 convolution kernel is used on the convolution layer stacked in the later three layers, which effectively reduces the space occupied by the feature map and maintains the feature extraction ability of the model; (2) Delete the fully connected layer 1 is deleted and directly connected with the fully connected layer 2. Secondly, the number of neurons in the remaining two fully connected layers is reduced to 1024 and 256. While reducing the number of parameters, it can make the features obtained by the last convolution layer more distinctive, which is helpful to improve the fusion effect. The improved VGG16 network constructed by aforementioned method is shown in Figure 2.

The input data dimension of the network is $224 \times 224 \times 3$. In the network, the rectified linear unit (ReLU) is used as the activation function, and dropout is used to solve the overfitting problem. Softmax is selected as the classifier for the classification task to estimate the probability of each class label in class K. Usually, the convolutional neural network needs to connect the low-dimensional full-connection layer after the convolution layer as a new feature layer to reduce the feature size, and the features obtained by the convolution layer usually contain rich image detail information. Therefore, the features obtained from the convolution are used as the features to be fused. Due to the large amount of VGG16 network parameters and easy overfitting, the proposed method uses large convolution kernel instead of small ones to reduce the complexity of the model, and reduces some fully connected layers to reduce the amount of parameters, so as to ensure the recognition efficiency on the basis of ensuring the accuracy.

3.4. Network Training. The whole network training takes the feature map extracted by the improved VGG16 network as a black box operation. The input layer data of CNN are the high-dimensional convolution abstract feature data output from the improved VGG16 network. The convolution neural

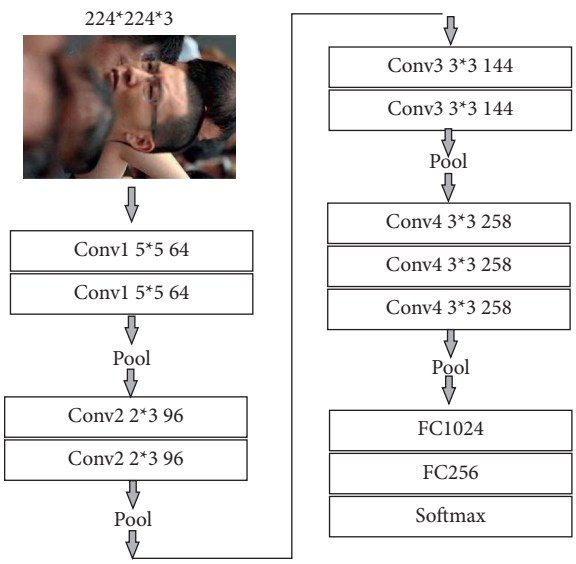


Figure 2: Improved VGG16 network.

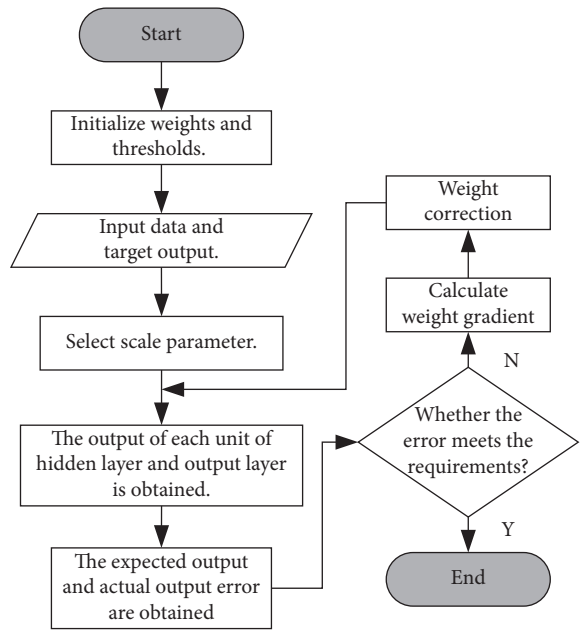


Figure 3: Overall training process of facial expression recognition networks.

network is used to train the characteristic data to obtain the training model [26]. The full-connection layer of CNN is a feedforward neural network. Its main responsibility is to undertake expression classification, but Softmax classifier is usually used to further improve the accuracy of expression classification. The overall training process of facial expression recognition network model is shown in Figure 3.

(1) Initialization and thresholds: There are seven categories of facial expressions. Therefore, compared with the traditional binary classification, the model parameters will increase significantly. The prediction model is a polynomial with parameters, and the prediction effect depends on the parameters of the model [27]. At present, the initialization of network parameters is mostly directly set to 0, which will lead to the same change trend of network parameters in the process of training, and even the parameter values are the same. Therefore, the proposed method uses random function assignment for the parameter initialization of the network model, which can avoid the influence of the initial parameter setting on the training model.

(2) Scale parameters: In machine learning, if the training data is less than the test data or the whole dataset itself is less, the whole model will have the phenomenon of overfitting or underfitting, resulting in poor prediction ability and failure to achieve the preset effect. In order to solve this imbalance, the corresponding proportion is selected by means of cross validation, retention method, and so on. In the training process, for the selection of network parameters, dropout is used to discard the parameters in proportion, so that the whole network becomes sparse, so as to enhance the generalization ability of the model. However, in the testing phase, dropout will be set to 0, that is, the parameters during training the model will be retained and used for model testing.

(3) Parameter iterative optimization is an important part of deep learning. When the gradient descent algorithm is used alone and when the initial value is far away from the minimum value, its step size will increase, and when it is close to the minimum value, its step size will become very slow [28]. When the gradient descent algorithm is used alone, regardless of the position where the initial value is far from or close to the minimum value, the step size will be large, and it is possible to skip directly for the next iteration. The proposed method combines the two methods and combines the gradient descent method and Gauss–Newton method through a linear combination. Its advantage is that it can quickly solve the problem of falling into local optimization and speed up the training speed.

3.5. Robot Interaction Design. The NAO robot uses NAOqi system. During program execution, a broker is used to load the required library files. Each library file contains one or more modules, such as "ALVideoDevice" vision module, "ALAudioDevice" sound module, and "ALMtion" action module.

3.5.1. Visual Module. The NAO robot has two 2D cameras, which are distributed on the forehead and mouth of the head. The camera can obtain pictures and video streams with a maximum resolution of 1280×960, which can be widely used in education, auxiliary medical treatment, and other fields. The camera adopts MT9M114 image sensor, and the photo format directly output by this device is YUV. In facial expression recognition, RGB color space is used. Therefore, it is necessary to subscribe to the "ALVideoDevice" module to change the YUV format into RGB. Considering the

limited CPU computing power of the NAO robot, the resolution of the proposed method is 640 × 480. This resolution will not cause too much loss of face image information, nor overload the processor of the robot.

3.5.2. Voice Module. The head of the NAO robot is equipped with four microphones and two speakers. The receiving frequency range of the microphone is 150 Hz–12 kHz, and the saved file format is WAV or OGG. In the NAOqi system, the "ALAudioDevice" sound module manages the input and output of audio. Therefore, when you need to read the data of the microphone, you must subscribe to the "ALAudioDevice" sound module.

3.5.3. Action Module. In the process of robot interaction, the face images collected by the NAO camera are output to the NAO robot after model prediction at the computer ends. The robot makes different interactive actions according to different expressions. The "ALMtion" action module includes methods related to the robot's actions. The robot action is written with the official Choregraphe graphical programming software. Using the time axis command box to design the formulated action, it is not necessary to use complex robot kinematics knowledge and program programming. The time axis is shown in Figure 4.

4. Experiment and Analysis

In the experiment, the hardware environment is one desktop computer, the CPU is Intel (R) core (TM) i7, 16 GB memory, and the GPU is GTX 1060. The development environment is TensorFlow framework based on Python language. In the training process, all learning rates are set to 0.01, the batch size is 500, the momentum is 0.99, the weight attenuation is 0.0001, the number of iterations is set to 100, and the maximum limit number is 10000. At the same time, the random gradient descent optimization method is used to train the network.

In addition, the data set used in the experiment is CK + data set. CK + dataset is an extension of Cohn-Kanade dataset, which is commonly used in facial expression recognition research. This dataset was released in 2010 and can be obtained free of charge. It is often used in facial expression recognition research. This dataset collected the frontal facial expressions of 123 people, a total of 593 image sequences. There are 7 facial expressions in total, including anger, disgust, fear, happiness, sadness, surprise, and contempt. The picture example of CK + dataset is shown in Figure 5.

988 pictures were selected, including 136 angry, 178 disgusted, 76 afraid, 208 happy, 85 sad, 250 surprised, and 55 neutral. 14 anger, 10 disgust, 11 fear, 23 happiness, 11 sadness, 26 surprise, and 8 neutral vision were selected as the test set, and the rest as the training set.

4.1. Training and Testing Accuracy versus Cycle Curve. The facial expression recognition accuracy curve of VGG16 network model and the proposed improved VGG16 network model during training and testing on the facial expression dataset CK + is shown in Figure 6.

It can be seen from Figure 6 that the improved VGG16 achieves 100% training accuracy faster during training, which proves that the fitting speed of the proposed network is faster. The improved VGG16 network has higher expression recognition accuracy than VGG16 network. Taking the test set as an example, the recognition accuracy of the improved VGG16 network is about 90%, which proves that its fitting effect is also better. In order to further improve the recognition accuracy, the proposed method combines the feature extraction advantages of improved VGG16 network and the classification characteristics of convolutional neural network to realize the high-precision human-computer interaction of the NAO robot system.

4.2. Influence of Different Feature Dimensions on Recognition Effect. In order to verify the influence of different networks and different feature dimensions on the recognition effect, the VGG16 network and the improved VGG16 network are used to conduct experiments on CK + dataset, and the features of different dimensions are selected to compare the recognition accuracy. The results are shown in Table 1.

As can be seen from Table 1, with the increase of feature dimension, the recognition accuracy of the two networks does not increase, but decreases slightly at 256 dimensions, indicating that the feature dimension is sufficient to characterize the CK + dataset at 256 dimensions, which contains most of the effective information in the dataset. The texture and detail features extracted by the proposed improved VGG16 in the shallow network are richer than those in the VGG16, especially some key features, such as eye feature information. In the deeper network, the improved VGG16 extracts more features such as contour and shape, especially the abstract features obtained in the last convolution layer, which are relatively more representative. At the same time, the improved VGG16 network does not reduce the ability of feature extraction due to the simplified structure. Therefore, this paper adopts the 256 dimensional feature dimension.

4.3. The Confusion Matrix. In the facial expression dataset CK + , the confusion matrix of 8 kinds of expression recognition results by the VGG16 network before and after improvement is shown in Figure 7.

As can be seen from Figure 7, compared with the VGG16 network, the improved VGG16 network improves the facial expression recognition rate of facial expression. In the original model, the highest accuracy is the recognition accuracy of happy expression, which reaches 92.83%. The lowest accuracy rate was neutral expression, only 71.16%; the average recognition accuracy is 84.03%. The improved VGG16 network optimizes the fully connected layer, so that the features obtained by the last layer of convolution are more differentiated, so the recognition effect is better. In addition, for the seven kinds of facial expressions, the recognition accuracy of disgust, neutral, and other expressions is low and easy to be confused. However, the improved VGG16 network uses a continuous large convolution kernel

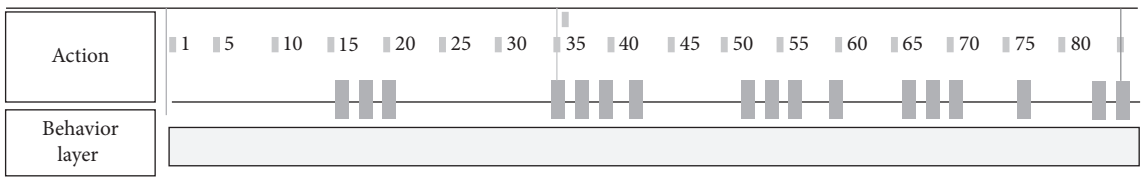


Figure 4: Time axis.

Figure 5: Picture example of CK + dataset.

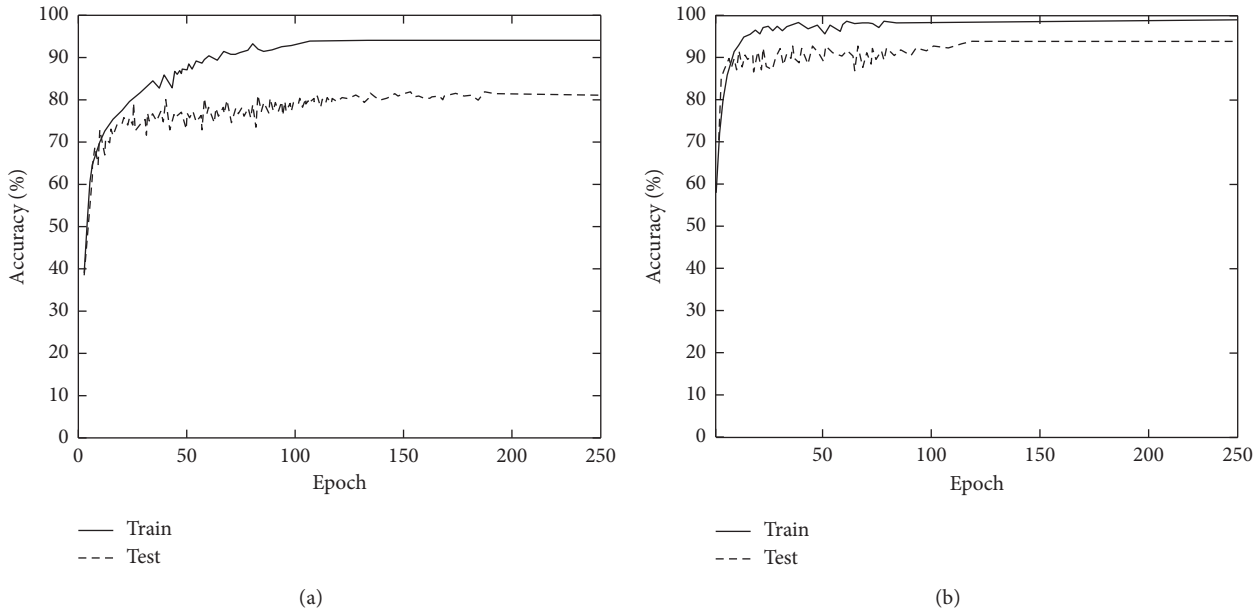


Figure 6: Expression recognition accuracy during training and testing. (a) VGG16 network training accuracy. (b) Improved VGG16 network training accuracy.

instead of a small convolution kernel and reduces part of the full-connection layer, which can better extract the expression detail features. Therefore, its recognition accuracy is better than the VGG16 network. The highest accuracy of the improved model is the recognition of happy expression, which is 98.78%; the lowest was disgusting expression, and

Table 1: Recognition accuracy of different feature dimensions.

Network structure	Feature dimension (%)			
	64	256	512	1024
VGG16	86.73	85.94	81.59	73.28
Improved VGG16	90.51	89.27	85.12	80.35

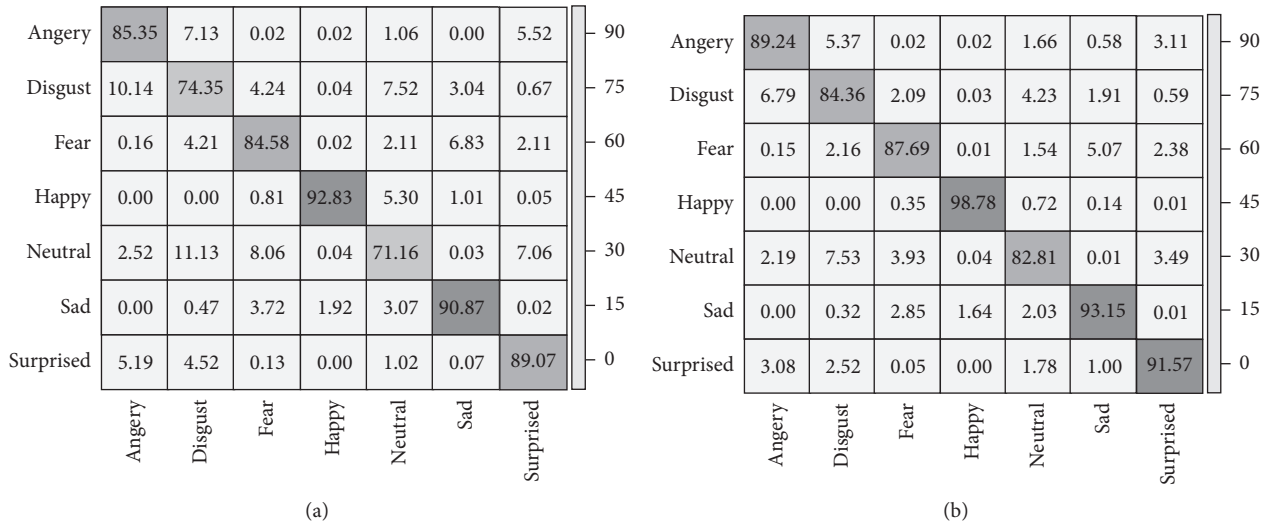


Figure 7: Confusion matrix of expression recognition results of different networks. (a) The VGG16 network. (b) The improved VGG16 network.

the recognition accuracy was 84.36. The average recognition accuracy is 89.66%, which is 5.63% higher than that of the original model. Therefore, it can be explained that the performance of the improved VGG16 network model is better than that of the traditional VGG16 network.

4.4. Human-Computer Interaction Test. The trained facial expression model is completed in Python 3.6, while the NAO robot can only support Python 2.7, so it cannot be deployed directly. Therefore, this interactive experiment is conducted through wireless connection. During the connection process, it should be noted that when the robot connects with the computer for the first time, it needs to insert the network cable to obtain the corresponding IP address, then log in to the web page through the IP address and set up the router for wireless connection. After the robot is connected to the network normally, the program is downloaded to the robot wirelessly. The PC obtains the face image by calling the robot's camera and sends it to the trained model. The predicted results of the model are then sent back to the NAO robot. The robot makes corresponding voice and actions according to the recognized emotions. This paper selects two representative expressions: happy and sad. The interaction results are shown in Figure 8.

As can be seen from Figure 8, the picture on the left is the expression result predicted by the model after the camera acquisition. The picture on the right shows the voice and corresponding actions made by the robot. In order to visualize the voice effect, the virtual robot was connected and the voice results were directly visualized. When the happy expression is detected, the robot will interact with you and say "You look happy. Can you share it with me?" When the sad expression is detected, the robot will make a stand up action and say "I'll be unhappy if you're sad." Through many tests, the interaction results show that the robot can stably recognize emotions and can make corresponding action interaction.

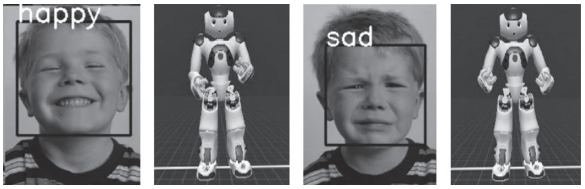


Figure 8: The emotional interaction results.

5. Conclusions

Face recognition technology has been relatively mature and has been widely used in many fields. However, face expression recognition is still in an exploratory stage. Aiming at the problems of insufficient feature data and a low recognition rate in face expression recognition, an expression recognition method based on the improved VGG16 network model is proposed. The improved VGG16 network is used to extract the features of facial expression gray image, and the high-level features are input into the CNN model. The classification of facial expression is realized by Softmax classifier. The comprehensive experimental results show that

the improved VGG16 network model is effective in training and testing expression recognition. With the expansion of the depth and structure of the network, the application of certain networks in some fields has shown better results. The next step of in-depth research can also consider the use of deeper network structures, structural innovations, and the parallel connection of multiple networks. And other methods are applied to the facial expression recognition system.

Data Availability

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

Conflicts of Interest

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

Acknowledgments

This work was supported by the scientific research project from Hunan Provincial Department of Education (no. 20C0218).

References

[1] D. K. Jain, P. Shamsolmoali, and P. Sehdev, "Extended deep neural network for facial emotion recognition," *Pattern Recognition Letters*, vol. 120, no. 8, pp. 69–74, 2019.

[2] Y. Chen, Z. Zhang, L. Zhong, T. Chen, J. Chen, and Y. Yu, "Three-stream convolutional neural network with squeeze-and-excitation block for near-infrared facial expression recognition," *Electronics*, vol. 8, no. 4, pp. 385–397, 2019.

[3] Y. Gan, J. Chen, and L. Xu, "Facial expression recognition boosted by soft label with a diverse ensemble," *Pattern Recognition Letters*, vol. 125, no. 7, pp. 105–112, 2019.

[4] X. Fan and T. Tjahjadi, "Fusing dynamic deep learned features and handcrafted features for facial expression recognition," *Journal of Visual Communication & Image Representation*, vol. 65, no. 11, pp. 102659.1–102659.6, 2019.

[5] D. Li, X. Zhao, G. Yuan, and X. Zhao, "Robustness comparison between the capsule network and the convolutional network for facial expression recognition," *Applied Intelligence*, vol. 51, no. 4, pp. 1–10, 2021.

[6] S. Cao, Y. Yao, and G. An, "E2-capsule neural networks for facial expression recognition using AU-aware attention," *IET Image Processing*, vol. 14, no. 11, pp. 2417–2424, 2020.

[7] A. Renda, M. Barsacchi, A. Bechini, and F. Marcelloni, "Comparing ensemble strategies for deep learning: an application to facial expression recognition," *Expert Systems with Applications*, vol. 136, no. 11, pp. 1–11, 2019.

[8] F. Wang, J. Lv, G. Ying, S. Chen, and C. Zhang, "Facial expression recognition from image based on hybrid features understanding," *Journal of Visual Communication and Image Representation*, vol. 59, no. 3, pp. 84–88, 2019.

[9] W. Cao, Z. Feng, D. Zhang, and Y. Huang, "Facial expression recognition via a CBAM embedded network," *Procedia Computer Science*, vol. 174, no. 4, pp. 463–477, 2020.

[10] N. Amjed, F. Khalid, R. W. O. K. Rahmat, and H. B. Madzin, "A robust geometric skin colour face detection method under unconstrained environment of smartphone database," *Applied Mechanics and Materials*, vol. 892, no. 3, pp. 31–37, 2019.

[11] 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.

[12] M. Sajjad, S. Zahir, A. Ullah, Z. Akhtar, and K. Muhammad, "Human behavior understanding in big multimedia data using CNN based facial expression recognition," *Mobile Networks and Applications*, vol. 25, no. 4, pp. 1611–1621, 2020.

[13] K. Li, Y. Jin, M. W. Akram, R. Han, and J. Chen, "Facial expression recognition with convolutional neural networks via a new face cropping and rotation strategy," *The Visual Computer*, vol. 36, no. 2, pp. 391–404, 2020.

[14] C. Xu, Y. Cui, Y. Zhang, P. Gao, and J. Xu, "Person-independent facial expression recognition method based on improved Wasserstein generative adversarial networks in combination with identity aware," *Multimedia Systems*, vol. 26, no. 1, pp. 53–61, 2020.

[15] H. Zhang, Z. Qu, L. Yuan, and G. Li, "A Face Recognition Method Based on LBP Feature for CNN[C]," in *Proceedings of the 2017 IEEE 2nd Advanced Information Technology, Electronic and Automation Control Conference*, pp. 544–547, IEEE, Chongqing, China, March 2017.

[16] G. Yolcu, I. Oztel, S. Kazan et al., "Facial expression recognition for monitoring neurological disorders based on convolutional neural network," *Multimedia Tools and Applications*, vol. 78, no. 22, pp. 31581–31603, 2019.

[17] J. Shao and Y. Qian, "Three convolutional neural network models for facial expression recognition in the wild," *Neurocomputing*, vol. 355, no. 8, pp. 82–92, 2019.

[18] H. Wang, J. Lu, and L. Nwosu, "Two-channel convolutional neural network for facial expression recognition using facial parts," *International Journal of Big Data Intelligence*, vol. 6, no. 3/4, pp. 259–268, 2019.

[19] M. Aamir, T. Ali, A. Shaf, M. Irfan, and M. Q. Saleem, "ML-DCNNet: multi-level deep convolutional neural network for facial expression recognition and intensity estimation," *Arabian Journal for Science and Engineering*, vol. 45, no. 12, pp. 10605–10620, 2020.

[20] Y. Ye, X. Zhang, Y. Lin, and H. Wang, "Facial expression recognition via region-based convolutional fusion network," *Journal of Visual Communication and Image Representation*, vol. 62, no. 7, pp. 1–11, 2019.

[21] J. Li, Y. Mi, and G. Li, "CNN-based facial expression recognition from annotated RGB-D images for human–robot interaction," *International Journal of Humanoid Robotics*, vol. 16, no. 4, pp. 504-505, 2019.

[22] D. K. Jain, Z. Zhang, and K. Huang, "Multi angle optimal pattern-based deep learning for automatic facial expression recognition," *Pattern Recognition Letters*, vol. 139, no. 3, pp. 157–165, 2020.

[23] U. B. Chavan and D. Kulkarni, "Optimizing deep convolutional neural network for facial expression recognition," *European Journal of Engineering Research and Science*, vol. 5, no. 2, pp. 192–195, 2020.

[24] R. Ramya, K. Mala, and S. Selva Nidhyananthan, "3D facial expression recognition using multi-channel deep learning framework," *Circuits, Systems, and Signal Processing*, vol. 39, no. 2, pp. 789–804, 2020.

[25] W. Wei, Q. Jia, Y. Feng, G. Chen, and M. Chu, "Multi-modal facial expression feature based on deep-neural networks," *Journal on Multimodal User Interfaces*, vol. 14, no. 1, pp. 17–23, 2020.

[26] Y. Cai, J. Gao, G. Zhang, and Y. Liu, "Efficient facial expression recognition based on convolutional neural network," *Intelligent Data Analysis*, vol. 25, no. 1, pp. 139–154, 2021.

[27] S. Suchitra, P. S. Sathya, P. Balachandran, and M. Faustina, "Intelligent driver warning system using deep learning-based facial expression recognition," *Scopus*, vol. 8, no. 3, pp. 831–838, 2019.

[28] A. Caroppo, A. Leone, and P. Siciliano, "Comparison between deep learning models and traditional machine learning approaches for facial expression recognition in ageing adults," *Journal of Computer Science and Technology*, vol. 35, no. 5, pp. 1127–1146, 2020.

Hindawi
Journal of Robotics
Volume 2021, Article ID 8363940, 8 pages
https://doi.org/10.1155/2021/8363940

Research Article

Multistage Feature Complimentary Network for Single-Image Deraining

Kangying Wang and Minghui Wang

College of Computer Science, Sichuan University, Chengdu, Sichuan 610065, China

Correspondence should be addressed to Kangying Wang; 2019223049279@stu.scu.edu.cn

Received 12 November 2021; Accepted 29 November 2021; Published 16 December 2021

Academic Editor: Kaijian Xia

Rain will cause the occlusion and blur of background and target objects and affect the image visual effect and subsequent image analysis. Aiming at the problem of insufficient rain removal in the current rain removal algorithm, in order to improve the accuracy of computer vision algorithm in the process of rain removal, this paper proposes a multistage framework based on progressive restoration combined with recurrent neural network and feature complementarity technology to remove rain streak from single images. Firstly, the encoder-decoder subnetwork is adapted to learn multiscale information and extract richer rain features. Secondly, the original resolution image restored by decoder is used to preserve refined image details. Finally, we use the effective information of the previous stage to guide the rain removal of the next stage by the recurrent neural network. The final experimental results show that a multistage feature complementarity network performs well on both synthetic rainy data sets and real-world rainy data sets can remove rain more completely, preserve more background details, and achieve better visual effects compared with some popular single-image deraining methods.

1. Introduction

When the computer vision systems work outdoors, they often suffer the effects of poor weather. Especially in rainy days, rain generally results in collecting images quality degradation and important information damage. This will degrade the performance of the computer vision systems or even break down and make it difficult to fully play a role in tasks such as object tracking [1] and object detection [2]. It is thus important to restore clean background images from degraded and rainy images.

In recent decades, many researchers have been paying attention to the deraining research and proposing a large number of methods, which mainly can be divided into two major directions: on the one hand, removing the rain in a video; on the other hand, deraining in a single image. Removing the rain in the video can use the temporal information between the frames of the video. Single rainy images not only lack the temporal information of frames but also have complex rain streaks with different sizes and directions in a single rainy image. So it is more challenging to research single-image deraining.

Single-image deraining algorithms mainly include model-based methods and data-driven methods. Kang et al. [3] firstly extract the high-frequency features from the rainy images by a bilateral filter and then decompose the extracted high-frequency features into a rainy part and a nonrainy part by dictionary learning and sparse coding. Finally, the high-frequency features of the nonrainy part are combined with the low-frequency features to obtain the final deraining result. Kim et al. [4] research on the typical characteristics of rain streaks and design nonlocal means (NL-means) to remove rain streaks. Zhang et al. [5] are inspired by sparse coding and low-rank representation and propose a general convolution filter, a single-image deraining algorithm based on convolution coding.

In recent years, with the continuously improving computing capacity of computer hardware, deep learning technology has been popular with the researchers of computer vision gradually. In the current research on single-

image deraining, the algorithms by deep learning dominate the single-image deraining method research. Representative methods mainly include the convolutional neural network by combining rain streaks detection and deraining architecture by Yang et al. [6], which can effectively handle the scenario of heavy rain, overlapping rain streaks, and rain accumulation. Fu et al. [7] decompose each input image into low-frequency background features and high-frequency detail features by a low-pass filter and then input the high-frequency detail features to the CNN to restore the clean background image. About the training of high-frequency detail features, Fu et al. [8] are inspired by the deep residual network and introduce residual learning to predict residual, which improves the efficiency of network training greatly. Although the deraining algorithms based on deep learning have achieved great success, there is still plenty of room to improve.

By studying a large number of proposed methods, most of the methods they proposed adapt a one-stage design, which idea of architecture design comes from high-level visual tasks, such as residual learning [9] and dense skip connection [10]. But there are some general problems: blur background texture, inadequate deraining resulting in some rain remains, and others. On multistage methods, they have usually broken down the whole process into multiple stages to remove rain progressively. Since the rain removal information in the earlier stage is instructive for the rain removal in the later stage, it is necessary for these different stages to cooperate to complete the rain removal, which is exactly the problem not taken into account by the representative multistage rain removal methods [6, 11].

To remove rain streaks and preserve more background details, we are inspired by deep learning and progressive restoration [12, 13] and propose a multistage framework decomposing challenging restoration tasks into smaller subtasks. In our network, each subtask corresponds to each stage, and each stage has a different design. We use an encoder-decoder network [14] to learn the features of rain streaks with different sizes and directions and use a recurrent neural network to learn the cross-stage complementary features. Finally, we combine the features with the original resolution rain image to preserve richer local details.

2. Deraining Method

Many deraining algorithms assume that rain steaks are sparse and have similar characters in falling directions and shapes. However, the rain streaks in the real world are usually very complicated and are not necessarily satisfied with this hypothesis. Even in synthetic rainy images, the rain may have different directions and sizes and overlap with each other, so it is difficult to remove all rain streaks at once. The close and large rain streaks are often removed in most methods, while the small rain streaks in the distance are ignored. Because, in reality, the deraining model is unpredictable and is difficult to remove rain by the one-stage method, even parallel subnetwork [12].

2.1. Deraining Model. The model research of deraining generally decomposes the rain image O into a linear combination of nonrain streaks background layer$B$ and rain streaks layer$R$ as follows:

$$O = B + R. \tag{1}$$

Nonrain streak background layer B can be obtained by removing the rain streaks layer$R$ from the rain image$O$.

Based on the deraining model in formula (1), in order to reduce the complexity of the model, we regard the similar rain streaks as one layer and then decompose the rain image into a combination of multilayer rain streaks layers and a clean background layer. The rain image model can be reformulated as follows:

$$O = B + \sum_{s=1}^{S} R_s, \tag{2}$$

where R_s represents the s-th rain streak layer in the stage s and S is the total number of stages.

2.2. Network Structure. In order to better adapt to the image rain removal task and model lightweight, the original basic components of the U-Net required are retained in the encoder-decoder network: there are two 3×3 conv in each layer of the encoder part (ReLU as the activation function), and 2×2 max pooling is used for downsampling. There are still two 3×3 conv in each layer of the decoder (ReLU as the activation function), and the upsampling is 2×2 deconv (transpose convolution). The encoder and decoder of the same layer use skip connection for feature concatenation.

On this basis, in order to better adapt to the multistage single-image rain removal task proposed in this paper, the original five-layer U-Net structure is reduced to three layers, which will minimize the amount of parameters in each stage of the network, and considering that the rain streak has different distributions in direction, color, and shape, the BN layer is removed (the normalization characteristics of BN layer are inconsistent with the features of rain map model proposed in this paper), which will also enhance the operation efficiency of the algorithm.

For the single-image rain removal task, the features of the rain streak are repetitive. By introducing RNN, the rain streak features extracted by ConvGRU in the encoder-decoder prediction process of the previous stage are fully utilized in the encoder or decoder prediction process of the later stage, which can make these features work together to capture the global features of the rain streak. In other words, ConvGRU is introduced to extract the rain streak feature information flow in the spatial dimension, so that these relevant context features have recurrent dependencies, so the dependent features can work together to extract the global streak features.

We propose a multistage rain removal network with feature aggregation in this paper shown in Figure 1, inspired by progressive restoration [13] to remove rain streak progressively. In the previous stage, we use encoder-decoder subnetwork [14] to encode multiscale information

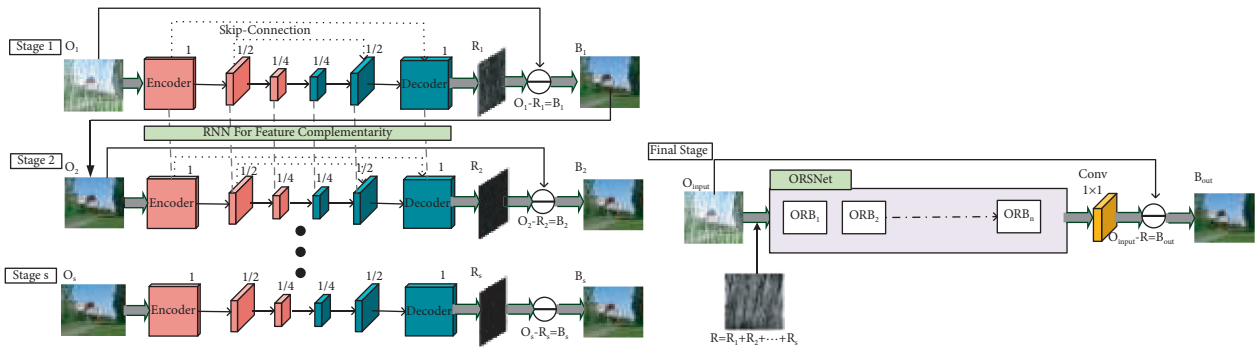


Figure 1: The structure of multistage feature aggregation network.

effectively. In the final stage, an original resolution subnetwork (ORS) is introduced to preserve the refine texture required in the final output image needed.

In the first stage of the network (RNN is not considered here), we can output a 256×256 feature map at the top layer of the decoder, which is recorded as R_1. R_1 is combined with the rain map input O_1 in this stage to output the rain removed image, which is recorded as B_1. This process completes the rain removal work in the first stage, that is, $B_1 = O_1 - R_1$.

After completing the rain removal work in the first stage, the rain removal output image B_1 is taken as the input of the next stage. For example, the input in stage s $(s > 1)$ is the output B_{S-1} of stage $s-1$, which is recorded as O_S. In stage s, the rain map to be predicted is still input to the encoder-decoder subnetwork for further rain streak feature extraction. But different from the first stage, the convolution operation of each layer in the encoder and decoder network of stage s will cooperate with the rain streak feature extracted by RNN in the previous stage encoder-decoder subnetwork to capture the global rain line feature of the image. In other words, by introducing the gated recurrent unit (ConvGRU) in RNN to capture the rain streak feature information flow in the spatial dimension, these relevant context texture features have recurrent dependencies, so the dependent features can work together to better extract the global texture features.

Finally, the rain streak features have been fully extracted by combining the encoder-decoder network and RNN in the s stages before the last stage, but considering that due to the repeated downsampling operation in the encoder, they tend to lose spatial details. In order to preserve the fine details from the input image to the output image, in the last stage, we introduce the original resolution subnetwork and input the final predicted rain streak feature map into the original resolution subnetwork to generate rich spatial high-resolution features, so as to make up for the loss of spatial information. At the end of the whole network, the high-resolution rain line features obtained from the original resolution subnetwork are combined with the original rain map to obtain the final rain removal image.

Instead of using decomposition methods with artificial prior information to solve the problem of formula (2), the model intends to learn a function f, which maps the observed rain image O to the rain streak image$R$ directly. Because the rain streak layer R is sparser than the observed rain image$O$ and has a simpler texture. Then we can obtain a nonrain streak background image$B$ by subtracting the rain streak image$R$ from the observed rain image O. The above function can be expressed as a deep neural network, which is learned by optimizing the loss function $\|f(O) - R\|_F^2$.

2.3. Multiscale Feature Extraction. For the single-image deraining task, the multiscale information from the input image has been proved important for the rain streaks recognition and removal task [14]. Because there may be more than one rain streak in the rain image, extracting the features of multiple scales and directions is more helpful for the description of rain, A module that can efficiently extract rain streak features is particularly important for the whole network. Therefore, we propose an encoder-decoder subnetwork based on U-Net to capture multiscale information of rain streaks. The encoder-decoder subnetwork is shown in Figure 2. For example, in the first stage of the network (the rain streak feature information of the previous stage supplemented by RNN in the subsequent stage is not considered here), if the input picture scale is 256×256, the first convolution operation in the encoder will extract an original scale (256×256) rain feature. After two times of pooling (downsampling) and convolution, the feature map of 256×256 will become 128×128 and 64×64. The deconvolution is applied in the decoder section to upsampling, so the previous 64×64 is upsampled to obtain a new 128×128 feature map. Then, the new 128×128 feature map is connected with another 128×128 feature map predicted by encoder, which we can get a 128×128 feature map of information fusion. After convolute the fused feature map, we get the final feature prediction map of a certain layer at the output of the decoder. Further using the same operation, we can output a 256×256 feature map at the top layer of the decoder. At the end of the decoder, through a 1×1 convolution kernel to recover the RGB channel of the color image or the gray channel of the gray image, which is recorded as R_1. R_1 combined with the rain map input O_1 in this stage to output the rain removed image, which is recorded as B_1. This process completes the rain removal work in the first stage, that is, $B_1 = O_1 - R_1$. In addition, the gray arrows in Figure 2 represent skip connections and aggregation operation.

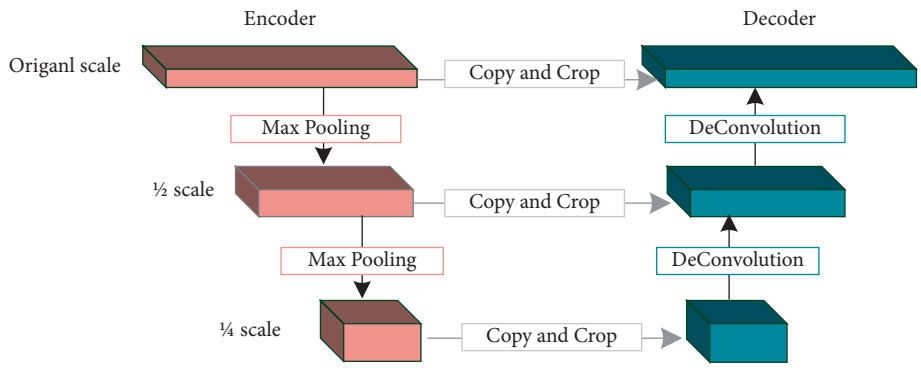


Figure 2: Encoder-decoder subnetwork.

Compared with the previous model [15], there is a big difference between our encoder-decoder network used in this module and others in that we do not adopt the batch normalization (BN) layer [9]. Although BN is widely used in deep neural network training and can reduce internal covariate shift of feature maps. Each scalar feature is normalized and has zero mean and unit variance by applying BN. However, these features are independent of each other and have the same distribution. In formula (2), the different layers of rain streak have different distributions in directions, colors, and shapes, and the same to each scalar feature of different rain streak layers. Therefore, BN contradicts the characteristics of our proposed deraining model; we remove BN from our model.

2.4. Multistage Feature Complement. Rain streaks may have different directions and overlap with each other, so it is difficult to remove them all at once. We send the preliminary results of rain streaks removal, their feature representation, and each predicted rain layer to the next stage for further refined restoration. Because the main information with obvious rain streaks (the closer and larger rain streaks) has been removed, the predictor of the next stage is better able to remove the remaining rain streaks [16]. This argument can be understood in different senses: (1) in the case of overlapping rain streaks mentioned previously, removing the nearest (and therefore brightest) rain streaks can reveal the darker rain streaks below the nearest rain streaks and (2) in the heavy rainy image, most rain streaks have relatively similar characteristics. After removing dominant rain streaks with similar direction or size, the other rain streaks will be detected and removed easily, which is inconsistent with the global pattern in size or direction.

Therefore, combined with the recurrent architecture, the process of rain streaks removal can be decomposed into multiple stages, which can be expressed as follows:

$$O_1 = O, \tag{3}$$

$$R_s = U_{\text{Code c}}(O_s), \quad 1 \le s \le S, \tag{4}$$

$$O_{s+1} = O_s - R_s, \quad 1 \le s \le S, \tag{5}$$

$$R = \sum_{s=1}^{S} R_s, \tag{6}$$

where S is the max number of the stage, R_s is the output of the s-th stage, O_{s+1} is the output image by the$s$-th stage rain streaks removal, and Codec means encoder-decoder subnetwork.

The above rain streaks removal model has been adopted in [6, 11]. But their methods only regard the recurrent structure as the same network with shared weights. Simultaneously, they only use the output feature map of the current stage as the input of the next stage and ignore the complementarity of different stages in their works. In our work, we use a multistage network structure to simulate the inverse process of rain formation to restore the nonrain streak image. With the gradual restoration of each stage, the interference degree of rain streaks to the image gradually decreases. So this is the reason why we introduce the recurrent neural network into our model. Formula (4) can be further expressed as follows:

$$R_s = U_{\text{Code c}}(O_s), s = 1 R_s = U_{\text{Code c+RNN}}(O_s, x_{s-1}), \quad 1 < s \le S, \tag{7}$$

where the O_S represents the rain map input in the stage s. x_{s-1} generally refers to the different feature maps extracted by the encoder-decoder network in the convolution layer of different scales in the previous stage of the s stage. In the downsampling stage of the encoder, we will extract the rain line features of the original size, 1/2 size, and 1/4 of the input image. Similarly, three rain streak feature maps of different scales will be extracted in the upsampling process of the decoder. We use ConvGRU to cooperate with the encoder-decoder subnetwork in the next stage to guide the same layer encoder or decoder to predict the rain streak feature.

ConvGRU is a popular current unit in sequence model. Our network uses a convolution based on ConvGRU in the model. x_s^j represents feature maps of the s-th stage and the j-th layer and can be described by x_{s-1}^j (the feature map of the upper stage) and x_s^{j-1} (the feature map of the current stage) as follows:

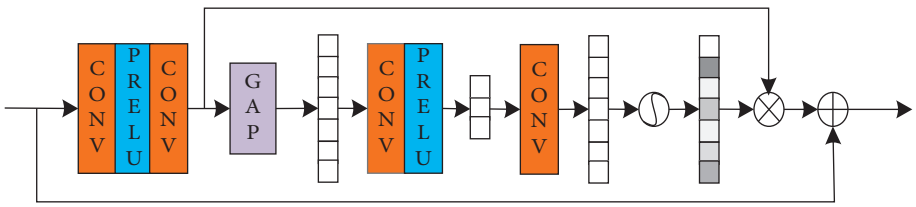


Figure 3: The structure of the original resolution model.

$$z_s^j = \sigma\left(W_z^j \# x_s^{j-1} + U_z^j \# x_{s-1}^j + b_z^j\right), \tag{8}$$

$$r_s^j = \sigma\left(\left(W_r^j \# x_s^{j-1} + U_r^j \# x_{s-1}^j + b_r^j\right), \tag{9}$$

$$n_s^j = \tanh\left(W_n^j \# x_s^{j-1} + U_n^j \# \left(r_s^j \odot x_{s-1}^j\right) + b_n^j\right), \tag{10}$$

$$x_s^j = \left(1 - z_s^j\right) \odot x_{s-1}^j + z_s^j \odot n_s^j, \tag{11}$$

where σ is the sigmoid function$\sigma(x) = 1/(1 + \exp(-x))$ and $\odot$ represent element multiplication.

2.5. Original Resolution Network. In the s stages before the implementation of the last stage, the rain streak features have been fully extracted by combining the encoder-decoder network and RNN. Although the model generates multiscale information effectively, it will tend to lose some spatial detail information by downsampling repeatedly. In order to preserve the refine details from the input image to the output image, we introduce the original resolution network (ORSNet) in the final stage.

We input the final predicted rain streak feature map into the original resolution subnetwork to generate rich spatial high-resolution features, so as to make up for the loss of spatial information. At the end of the whole network, the high-resolution rain line features obtained from the original resolution subnetwork are combined with the original rain map to obtain the final rain removal image.

ORSNet generates rich high-resolution spatial features without downsampling operations, which consist of many original resolution blocks (ORBs). The structure of the ORB is shown in Figure 3, where GAP represents global mean pooling.

3. Experiment

3.1. Training and Data Sets. To verify the effectiveness of our method, we test on many data sets, such as Rain100L, Rain100H, and Rain12 data sets. The three data sets are synthetic rainy image data sets and are used in deraining research popular. The typical characteristics of the Rain100L data set are thinner rain streaks and smaller raindrops. The rain streaks of the Rain100H data set have large sizes and different directions. The Rain12 data set is the synthetic rainy images but is the most similar to real-world rainy images by using rendering techniques in the process of synthetic. In terms of the real rainy image data set, we adopt the data sets Rain100L and Rain100H proposed by Yang et al. [6] to verify the deraining affection of our model in the real scene.

Table 1: Comparison of SSIM quality evaluation on three data sets.

Data sets	Yang et al. [6]	Fu et al. [8]	Zhang et al. [14]	Ours
Rain100L	0.9701	0.8447	0.8516	0.9729
Rain100H	0.7629	0.6954	0.7822	0.8683
Rain12	0.9568	0.9407	0.9130	0.9592

Table 2: Comparison of PSNR quality evaluation on three data sets.

Data sets	Yang et al. [6]	Fu et al. [8]	Zhang et al. [14]	Ours
Rain100L	35.96	32.57	25.84	36.09
Rain100H	24.15	23.60	23.94	24.82

Rain100L and Rain100H respectively contain 2,000 pairs of synthetic images (synthetic image pairs are composed of rainy images and their corresponding nonrainy images), from which 1,800 pairs of synthetic images are selected as the training data set and the remaining 200 pairs of synthetic images are selected as the test data set. There are only 12 pairs of synthetic rainy images in Rain12, so Rain12 is only used as a test data set in the experiment.

In our baseline, we reduced the number of layers of the original U-Net to 3, so that the encoder can obtain the feature map of the original scale, 1/2 scale, and 1/4 scale in the downsampling process. Similarly, only two deconvolution operations are required in the upsampling process. For the nonlinear operation, we use ReLu. We use Adam optimizer with a batch size of 8 for training on the NVDIA 2080TI GPU; for optimization, the ADAM [17] algorithm is adopted with a start learning rate of 5×10^{-3}. During training, the learning rate is divided by 10 at every 20,000 steps. And the network is trained for 30 epochs with the above settings.

3.2. Comparison and Analysis of Experiment Results. To verify the deraining affection of the model we proposed, we compare the proposed method with JORDER [6], DDN [8], DIDMDN [14], and another representative single-image rain streaks removal method on synthetic data sets. Since there are corresponding nonrainy images in the synthetic data set, the index can use structural similarity index metric (SSIM) and peak signal-to-noise ratio (PSNR) to evaluate the image quality and compare the deraining effects of different algorithms objectively. However, lacking corresponding real nonrainy images, the quality of rain removal is evaluated by a subjective visual affection on the test of the real rainy image data set.

3.2.1. Synthetic Data Sets. In Tables 1 and 2, the comparison results of SSIM and PSNR between the proposed method and other three popular methods [6, 8, 14] in three synthetic

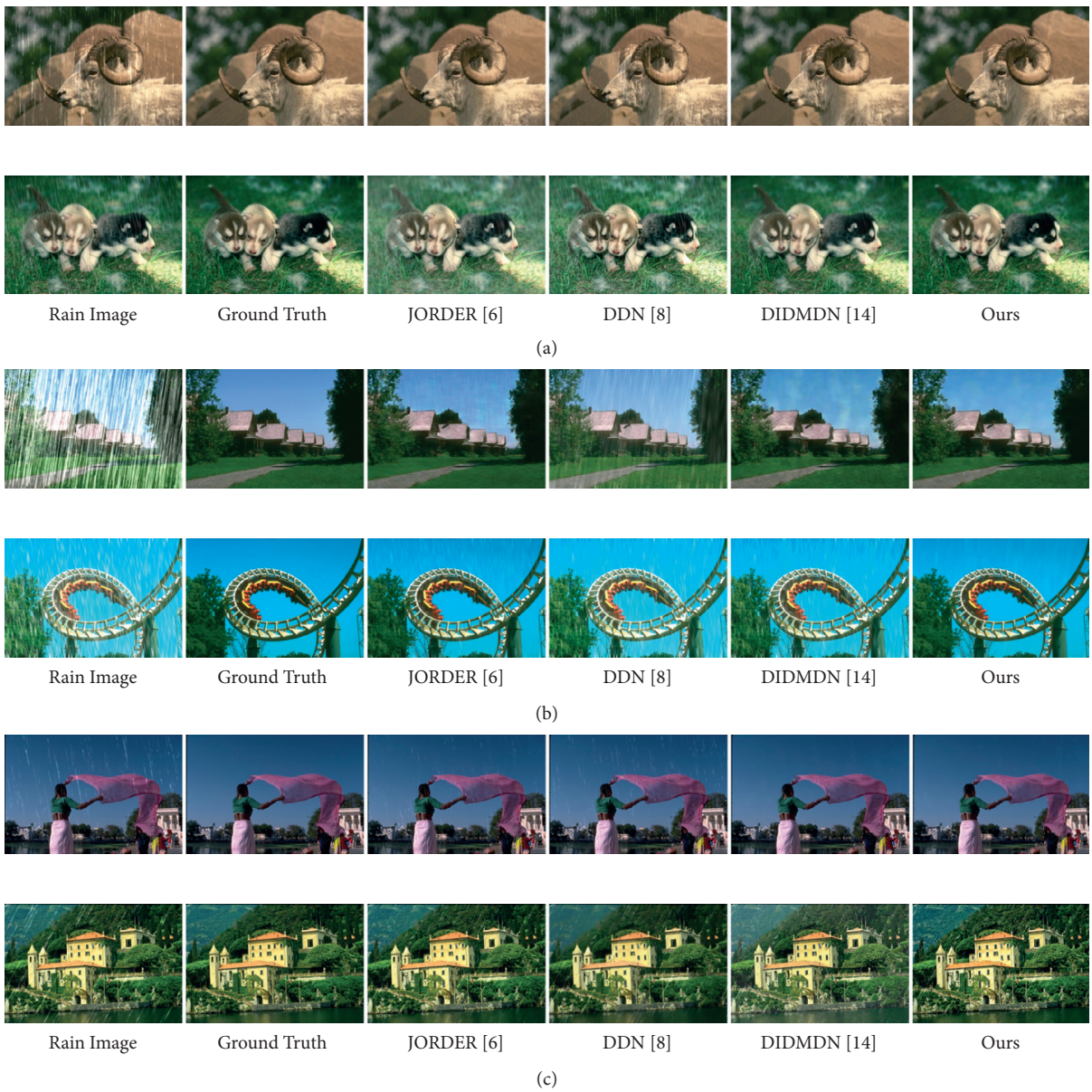


Figure 4: Contrast diagram of rain removal effect of each method in synthetic image: (a) the experiment results of Rain100L data set, (b) The experiment results of Rain100H data set, and (c) the experiment results of Rain12 data set.

data sets are shown separately. As we see from Table 1, the index SSIM of the method we proposed is superior to other methods on three data sets, especially in the Rain100H data set. It means that the method we proposed can remove rain from heavy rain, overlapping rain, and other complex conditions. In Table 2, the PSNR index of our method performs better than the method proposed in Fu et al. [8] and Zhang and Patel [14] on Rain100L and Rain100H data sets obviously but only slightly superior to the method in Yang et al. [6]. The fog removal algorithm is added in the process of rain removal in Yang et al. [6] to restore nonrainy image effectively.

Figure 4 shows the comparison of rain streaks removal effects, in which Figures 4(a)–4(c) are the experiments on Rain100L, Rain100H, and Rain12, respectively. From Figure 4, we can know that the methods in [6] and [8] both have the obvious problem of rain streaks remaining and blur images in restored images. The method in [14] is more thorough in removing rain streaks but lost some details of goat horns. Our method removes all rain streaks and preserves sufficient image details. As seen from Figure 4(b) of Rain100H experiment results: compared with [8], the method in [6] removed more rain streaks, but there were still some rain streaks remaining and lost some refine details. The method in [14] could well identify and deal with heavy rain, but the image is a blurred image after removing rain streaks. Our method has no rain streaks remaining, and the restored image is clear. From Figure 4(c) of results on Rain12, the method in [14] still has obvious rain streaks remain in the skirt part of the image, while there is no obvious rain streaks

Rain Image JORDER[6] DDN [8] DIDMDN [14] Ours

Figure 5: Comparison diagram of rain removal effect of each method on a real image.

remain in [6, 8], but there are still problems of detail loss and blur. However, the method in this paper almost removes all rain streaks and the image is clear.

Compared with single-stage rain removal networks such as DDN [8] and DIDMDN [14], although better spatial image details can be retained in the single-scale pipeline architecture, due to the limited receptive field, they cannot well detect the features of rain streak from different shapes and sizes, so that they always cannot fully remove rain streak. In contrast, our method recurrently uses the encoder-decoder subnetwork through the multistage network structure, which can better detect and remove all kinds of rain streaks. Compared with Jordan [6], which is a multistage rain removal network, Jordan [6] does not consider the relationship between rain streak features in each stage. Furthermore, our network saves the extracted rain streak multiscale information in the encoder-decoder subnetwork through RNN and uses it to guide the rain streak detection and removal in the later stage, which is very effective for fully removing rain streaks. In addition, in order to make up for the image information lost in the upsampling process in the encoder, the rain removal image output by the final stage through the original resolution subnetwork to ensure the information integrity and clarity of the image, which makes the method comparable to the single-scale pipeline architecture in spatial image details.

3.2.2. Real Scene Data Sets. To verify the practicability of the proposed method, three other popular rain removal methods and the method we proposed are tested on realistic rain image data sets. Since there are no nonrainy images for comparison in real scenes, the evaluation index of real rainy images is subjective visual effects. Figure 5 shows the rain removal effect of each method on the real rainy image. Compared with other methods, the method we proposed removes rain streaks more thoroughly and retains more details, resulting in the optimal visual effect after removing, which further verifies the practicability of the method proposed in this paper.

4. Conclusion and Discussion

A new single-image deraining method named multistage feature complimentary network based on progressive restoration has been proposed. The result shows that the method we proposed can effectively deal with different rainfall scenarios; especially in the case of large rainfall and overlapping rain streaks, it can give full play to the performance of the network. Although the network has done the lightweight operation suitable for the rain removal task on the basic U-Net architecture, the network still appears redundant when the rain streaks almost are small scale or the difficulty of rain line removal is not high. The data sets used in the training of the network are pictures in sunny weather, so the difficulty of extracting rain streak features will be reduced. If the rain information is no longer the highlighted white rain bar in the common rain removal data set at night, fog, or under the reflection of light, the rain removal effect of the algorithm will not be ideal.

Because the number of improved network layers is still large, and the calculation speed is not well suitable for real-time applications, there are still many problems to be solved from the perfect real-time rain removal. Whether it is video image or single-image rain removal, there is not a large number of successful precedents in practical application. Due to the complexity of rainfall, it is currently impossible to use a network framework to carry out rain removal under various circumstances. Therefore, for different tasks, how to effectively integrate rain removal algorithms to improve the accuracy of computer vision algorithms is worth studying. For the landing application of rain removal algorithms, it involves the compression and acceleration of network models, which is also a hot research topic in image rain removal in recent years.

Data Availability

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

Conflicts of Interest

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

References

[1] Y. Sun and Y. Shi, "Adaptive decision multi-layer features fusion for object tracking," *Journal of Electronics & Information Technology*, vol. 41, no. 10, pp. 2464–2470, 2019.

[2] Z. Guo, P. Song, and Y. Zhang, "Remote sensing image aircraft detection based on deep convolutional neural Network," *Journal of Electronics & Information Technology*, vol. 40, no. 11, pp. 2684–2690, 2018.

[3] L. W. Kang, C. W. Lin, and Y. H. Fu, "Automatic single image-based rain streaks removal via image decomposition," *IEEE Transactions on Image Processing*, vol. 21, no. 4, pp. 1742–1755, 2012.

[4] J.-H. Kim, J.-Y. Sim, and C.-S. Kim, "Video deraining and desnowing using temporal correlation and low-rank matrix completion," *IEEE Transactions on Image Processing*, vol. 24, no. 9, pp. 2658–2670, 2015.

[5] X. Zhang, H. Li, Y. Qi, W. K. Leow, and T. K. Ng, "Rain removal in video by combining temporal and chromatic properties," in *Proceedings of the IEEE International Conference on Multimedia and Expo*, pp. 461–464, IEEE Press, Toronto, Ontario, Canada, July 2006.

[6] W. Yang, R. T. Tan, J. Feng, J. Liu, Z. Guo, and S. Yan, "Deep joint rain detection and removal from a single image," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1357–1366, IEEE Press, Honolulu, HI, USA, July 2017.

[7] X. Fu, J. Huang, X. Ding, Y. Liao, and J. Paisley, "Clearing the skies: a deep network architecture for single-image rain removal," *IEEE Transactions on Image Processing*, vol. 26, no. 6, pp. 2944–2956, 2016.

[8] X. Fu, J. Huang, D. Zeng, Y. Huang, X. Ding, and J. Paisley, "Removing rain from single images via a deep detail network," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1715–1723, IEEE Press, Honolulu, HI, USA, July 2017.

[9] K. Jiang, Z. Wang, Yi Peng et al., "Multi-scale progressive fusion network for single image deraining," in *Proceedings of the CVPR*, Seattle, WA, USA, June 2020.

[10] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, "Densel connected convolutional networks," in *Proceedings of the CVPR*, Honolulu, Hawaii, July 2017.

[11] R. Li, L. F. Cheong, and R. T. Tan, "Single image deraining using scale-aware multistage recurrent network," 2019, https://arxiv.org/abs/1712.06830.

[12] X. Li, J. Wu, Z. Lin, H. Liu, and H. Zha, "Recurrent squeeze-and-excitation context aggregation net for single image deraining," in *Proceedings of the IEEE European Conference Computer Vision*, Munich, Germany, September 2018.

[13] S. Waqas Zamir, A. Arora, S. Khan et al., "CycleISP: real image restoration via improved data synthesis," in *Proceedings of the CVPR*, Seattle, WA, USA, June 2020.

[14] H. Zhang and V. M. Patel, "Density-aware single image deraining using a multi-stream dense network," in *Proceedings of the IEEE Int'l Conference Computer Vision and Pattern Recognition*, Salt Lake City, Utah, USA, June 2018.

[15] M. Lan and C. Li, "A hybrid neural network for image restoration," *In Computer Engineering and Applications*, vol. 54, no. 9, pp. 201–206, 2018.

[16] Y. Zheng, X. Yu, M. Liu, and S. Zhang, "Residual multiscale based single image deraining," in *Proceedings of the BMVC*, Cardiff, UK, September 2019.

[17] D. P. Kingma and B. Jimmy, "Adam: a method for stochastic optimization," 2014, https://arxiv.org/abs/1412.6980.

Hindawi
Journal of Robotics
Volume 2021, Article ID 6551159, 11 pages
https://doi.org/10.1155/2021/6551159

Research Article

Design and Performance Simulation of Computer Control System for Automatic Monitoring of Upper Computer Communication Operation State

Li Yang [1] **and Huitao Zhang** [2]

[1] *School of Electrical and Mechanical, Shijiazhuang University, Hebei, Shijiazhuang 050035, China*
[2] *School of Computer Science, Shijiazhuang University, Hebei, Shijiazhuang 050035, China*

Correspondence should be addressed to Huitao Zhang; 12116317@bjtu.edu.cn

Received 22 October 2021; Revised 8 November 2021; Accepted 10 November 2021; Published 6 December 2021

Academic Editor: Shan Zhong

The upper computer communication operation state automatic monitoring system is mainly used to remotely monitor the equipment, obtain various parameter indexes in the operation process of remote equipment, realize remote monitoring and fault diagnosis, and improve the management efficiency of decentralized equipment. This paper completes the design of communication, data storage, query, and other subsystems of upper and lower computers. The lower computer establishes a data channel with the OPC server through the MPI protocol and uploads the collected data to the OPC server in real time. The upper computer reads the data through the OPC server and displays the changes of monitored parameters in real time through the monitoring interface, so as to give an alarm under abnormal conditions. In addition, since the default database of Kingview is access, considering that the Microsoft Access database can store up to 2G of content, in order to upgrade and expand the subsequent system, SQL Server database is selected for data query, backup, and saving. The parameter setting method of communication control system is analyzed, the simulation model of industrial boiler control system is established by using Matlab/Simulink, and the interface between host computer software (IBCCS-e) and the model is provided. This paper analyzes the results of communication parameter adjustment. The simulation results show that the industrial boiler computer control system (IBCCS) has stable performance, low cost, convenient operation, and good maintainability. After further improvement, it has certain application value in the operation transformation of new small- and medium-sized boilers and original boilers.

1. Introduction

With the continuous expansion of the operating scale of industrial enterprises, only realizing the control of various field devices can no longer meet higher requirements, and the real-time response to the working conditions of the field devices has more practical significance [1]. As the company's on-site environment is becoming more and more complex, and the targets to be monitored are more diverse, the structure of the monitoring system becomes more and more decentralized. The traditional programmable logic device to directly control field devices can no longer meet the decentralized requirements of the system topology [2]. However, in the application that combines PLC and fieldbus technology after development, only PLC programming is used to realize the control of the device, so industrial monitoring systems began to tend to use computer monitoring systems [3].

The computer remote monitoring system is a monitoring network system that combines multimedia technology, network technology, and industrial automation technology with a computer as the core [4]. Its obvious advantage lies in remote monitoring and effectiveness. The remote monitoring system can obtain on-site information and send it to the monitoring center and make it reach the personal desktop PC through the computer network, so that it can be easily integrated with the information management system to better serve management and improve management level

and efficiency. The use of computers to monitor and control industrial production processes is widely used in modern society and plays an extremely important role in the safety and economy of industrial production operations. The production situation on site can be reproduced in time in front of decision-makers and managers away from the site, and the real-time data on site can be provided to functional departments at all levels to form charts to integrate production control and modern management [5]. The computer monitoring system can easily turn the scattered, large-area console-type monitoring into a centralized screen monitoring, which greatly reduces the work intensity of the operators [6].

This paper studies the overall architecture of the host computer communication operation state automatic monitoring system, provides a reasonable solution according to the specific situation of this paper, analyzes the network topology and architecture of the wireless monitoring system, determines the selection of measurement parameters, and identifies the radio wave propagation characteristics and signal maintenance methods of the communication system. The host computer system of PLC monitoring system of node boiler, the remote monitoring system of centralized control center, and the networking mode of 19 monitoring nodes are designed. The storage, calling, and permission setting of boiler operation state parameters are clarified. The remote monitoring is described. The working method of PID parameter setting module in host computer communication software (IBCCS-e) is introduced in detail. Then, the simulation model of industrial boiler computer control system is established by using Matlab/Simulink, and the data interface between Delphi and the simulation model is realized. Based on the simulation model, the parameter setting simulation experiments of water supply communication control system and combustion control system are carried out, and the experimental results are analyzed.

2. Related Work

In the centralized computer monitoring structure, the host computer in the central control room processes through the host computer configuration software system the various analog signals and switch signals that are introduced into the central control room host interface via the relevant cables, and then the form is issued to realize the background related data processing and real-time control of on-site equipment [7]. The monitoring system uses a single host; once the host fails, the entire system will be paralyzed; all signals are processed by a CPU, and the real-time performance is difficult to guarantee. Related scholars pointed out that, for systems with scattered signal collection points such as small hydropower stations, a large amount of cables must be used for signal transmission, which increases the cost of the entire system design; the existence of electromagnetic interference seriously affects the reliability and measurement of the system [8].

There are relatively few researches on the real-time database and cloud platform of APF operation and management, and the structure analysis is not clear. In the design of database-related strategies, existing or backward strategies are mostly used without reasonable analysis [9]. The data of the APF running real-time monitoring system has its own characteristics, but the existing real-time database and relational database are not specific to the data processing generated by the APF operation. The real-time database and the relational database are part of the APF running real-time monitoring platform, and the APF operation real-time monitoring platform is consistent to be integrated into the APF operation real-time monitoring platform [10]. Therefore, traditional industrial real-time databases and relational databases are not suitable for real-time data processing of APF running real-time monitoring systems. In terms of APF equipment monitoring, power electronic technology and cloud computing technology are widely used to realize real-time data collection and intelligent processing [11]. The APF monitoring system integrates database technology and computer technology, effectively stores and manages data, and provides the system with corresponding basic data for secondary calculation and processing, thereby improving the automation of the APF real-time monitoring system.

With the gradual expansion of the network scale, the demand for network monitoring models is gradually changing [12]. Currently, commonly used network monitoring models are centralized network monitoring model based on C/S, network monitoring model based on Web mode, network monitoring model based on distributed mode, and network monitoring model based on multiagent mode [13]. The network monitoring model has gradually changed from a traditional centralized model to a distributed model, and from a single model to a hybrid model.

At present, domestic and foreign monitoring software can be divided into three parts. The first part is foreign professional monitoring software represented by HP and IBM. Its management products focus on the server monitoring of their own products, mainly through agent monitoring. The performance of the monitoring software is relatively stable. The second part is the domestic companies represented by MXsoft, Youlong, and North Tower. These companies rely on secondary development and localized technical support and services to obtain more and more applications, and their product technologies are gradually becoming more popular. The third part is composed of other related software vendors. The network monitoring software analyzes from a technical point of view and realizes the comprehensive management mode from the product's own tool management to the efficient integration, and distributed combination is the future development trend of information system management technology [14, 15]. These commercial monitoring software tools are easy to operate and have stable performance, but they also have shortcomings such as high closure, strong pertinence, poor scalability, and high cost.

At present, the domestic condition monitoring system is based on a microcomputer and centered on software development [16, 17]. It has multiple functions, great flexibility, wide application, and low system price, which can better complete the purpose of condition monitoring of hydropower units. However, there are still some problems such as weak hardware functions, mainly roving monitoring,

and poor real-time performance. Field wiring workload is large, signal anti-interference issues need to be carefully considered, it is difficult to adapt to the harsh environment of the industrial site, and because each manufacturer uses different technologies, it is difficult to achieve interchange and interoperability with each other, and compatibility is not good [18]. How to solve these problems and further improve and perfect the performance of the condition monitoring system has become an urgent problem for system designers [19, 20].

3. The Overall Architecture of the Automatic Monitoring System for the Running Status of the Upper Computer

3.1. The Network Topology of the Host Computer Wireless Monitoring System. This paper proposes a remote monitoring system for equipment based on wireless communication. Its network topology is shown in Figure 1. It can be seen from the figure that the overall structure of the monitoring system is mainly composed of three parts, including information collection and transmission equipment, data storage and analysis equipment, and remote automatic monitoring system software.

The monitoring terminal collects and analyzes the information during the operation of the equipment in real time and periodically transmits the collected data wirelessly to the server; the data storage and analysis equipment stores and analyzes the data in the server. Data display and remote automatic monitoring system software processing analyses are convenient for managers to carry out real-time monitoring and fault trend prediction; at the same time, the data in the server can also be published to various remote terminals (such as mobile phones and computers) through the Internet. With software, supervisors can obtain real-time monitoring equipment operating behavior data information from all over the world, which will greatly benefit the future integration with smart home systems and realize the ubiquitous "Internet of Things" communication era [21].

3.2. Host Computer Wireless Monitoring System Architecture. From the system architecture of the monitoring system, it can be divided into the bottom field data collection layer, the wireless data transmission layer, and the top application service layer, as shown in Figure 2.

The data acquisition layer relies on various sensors for real-time acquisition and conversion of data from various measuring points of field equipment. The data transmission layer is composed of wireless terminals and concentrators, without on-site wiring and has high flexibility and scalability. The concentrator and the upper computer are relatively fixed. The upper computer is generally located in the monitoring room, and the concentrator is located above the monitoring room to receive data. The concentrator directly transmits on-site information to the upper computer for data storage and processing through a cable connection. The application service layer can realize the database management of the monitoring system and respond to client requests. It can provide authorized data services to corporate intranet users via the local area network and can provide corresponding data services to authorized personnel in various places via the Internet [22].

3.3. Analysis of System Radio Wave Propagation Characteristics. The signal received by the concentrator is composed of direct waves, ground reflection waves, and obstacle diffraction waves. The height of the concentrator is hb, and the height of the wireless transmitter is hm. The propagation distances of the radio waves are d, d_1, and d_2, respectively. Then, the signal field strength E_r at the receiving point of the concentrator is the combination of the above three waves:

$$E_r = E_0 \cdot \left|1 - e^{-2\pi j/\lambda} \cdot \Delta d_1 - e^{-2\pi j/\lambda} \cdot \Delta d_2\right|. \quad (1)$$

Among them, $\Delta d_1 = d_1 - d$, $\Delta d_2 = d_2 - d$, E_0 is the direct wave field strength, λ is the working wavelength, a_1 is the attenuation coefficient of the ground reflected wave relative to the direct wave, and a_2 is the obstacle diffraction wave relative to the direct wave.

In on-site communication, the propagation of radio waves will encounter various reflections, scattering, and absorption. The received signal is not only composed of direct waves, but a combination of multiple radio waves, and the received point field strength is an interference field. Suppose that the carrier signal transmitted by the wireless transmitter is

$$S_0(t) = a_0 \cdot \lambda \cdot e^{j\phi_0 - w_0 t}. \quad (2)$$

After reflection, scattering, and absorption, the carrier signal of the electric wave reaching the receiving point is

$$S(t) = \prod_{i=0}^{n-1} \Phi_i \cdot S_i(t). \quad (3)$$

Among them, $S_i(t)$ is the i-th carrier signal arriving at the receiving point, a_i is the amplitude of the i-th signal, and Φ_i is the phase change value of the i-th signal. Therefore,

$$S(t) = \prod_{i=0}^{n-1} a_i \cdot e^{\phi_i - kvt \sin\theta_i + kw_0 t}. \quad (4)$$

Among them, k is the phase shift constant, $k = 2\pi/\lambda$; v is the moving speed of the device; θ_i is the incident angle of the radio wave; $kv\cos\theta_i$ is the angular frequency shift caused by the Doppler effect.

Since there are many factors that cause reflection in the environment, $S(t)$ is a combination of multiple radio waves. In the above formula, n is infinite, so it can be rewritten as the following formula:

$$S(t) = A_t jS \cdot e^{j\Psi_t} \cdot e^{jw_0 t - \Psi_i}. \quad (5)$$

Among them,

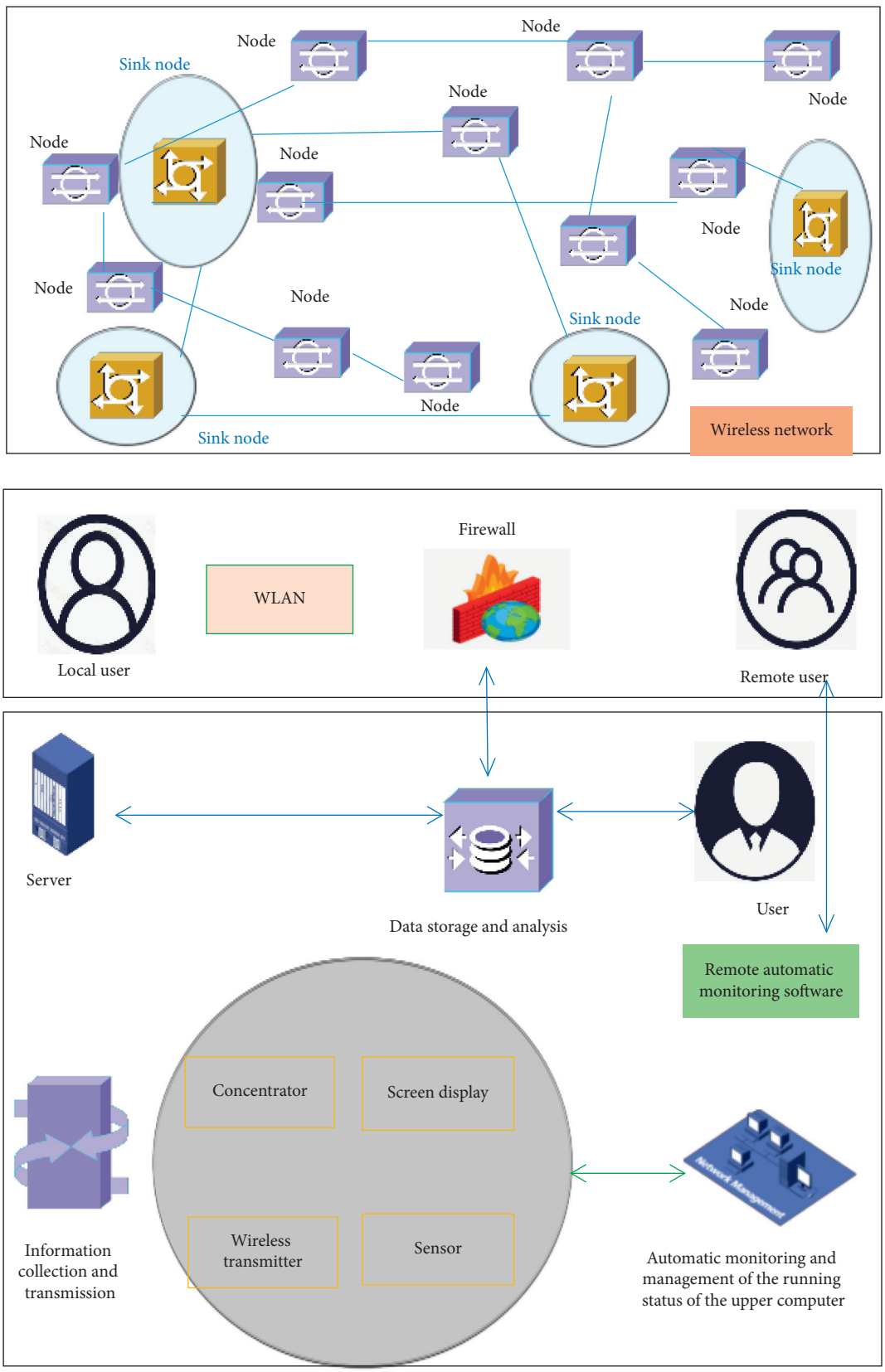


Figure 1: Wireless monitoring system network architecture diagram.

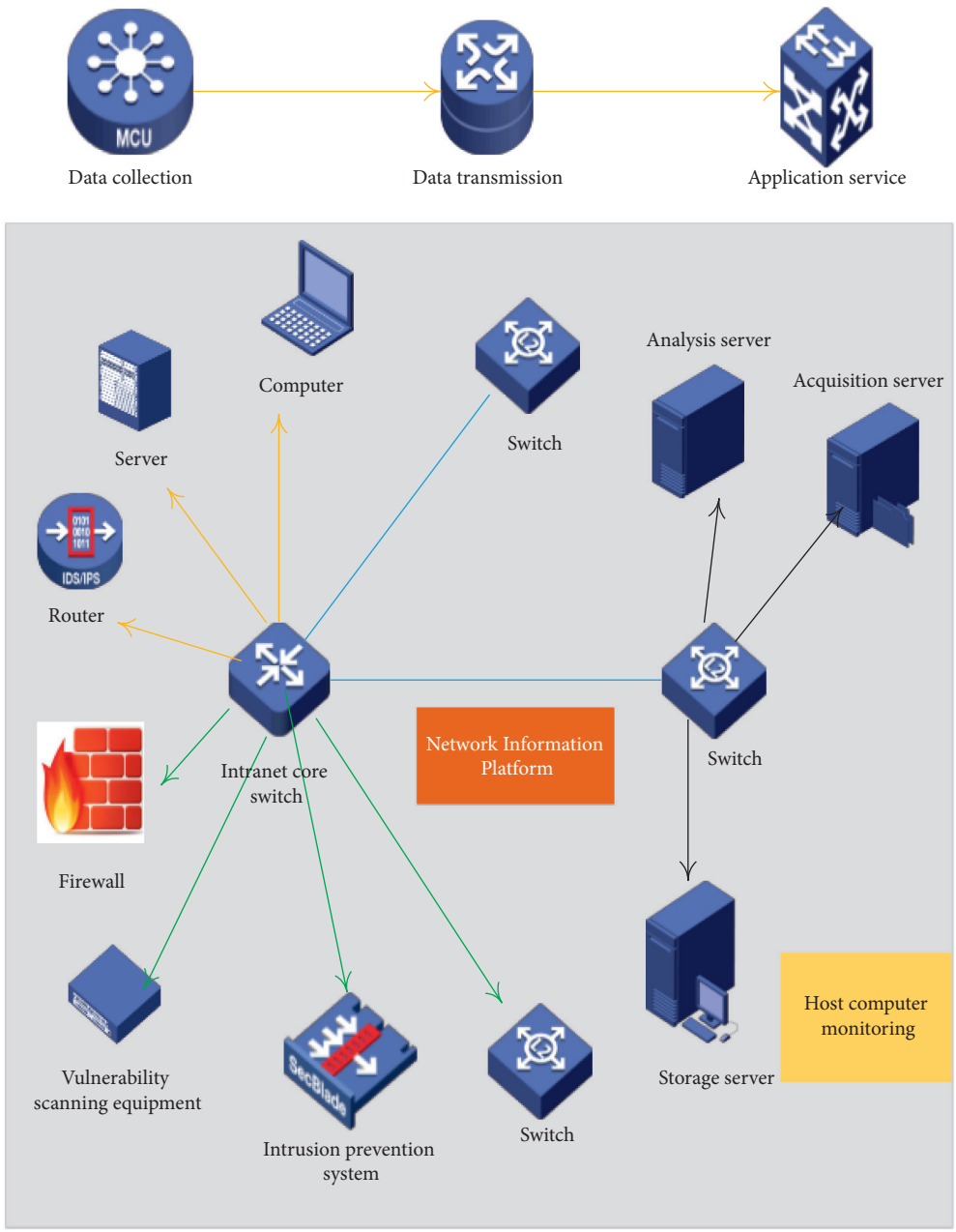


Figure 2: Wireless monitoring system architecture.

$$A_t = \prod_{i=0}^{n-1} a_i \cdot \sin \Psi_i,$$

$$S = \prod_{i=0}^{n-1} a_i \cdot \tan \Psi_i, \tag{6}$$

$$\Psi_i = kvt \sin \theta_i - \phi_i,$$

$$\Psi_t = \arccos\left(\frac{S}{R}\right).$$

Among them, R and S are independent variables composed of the sum of two infinitely large numbers of random variables, which obey a normal distribution, respectively. Their composite wave (that is, the amplitude A_t of the carrier signal at the receiving point) obeys the Rayleigh distribution, and the phase Ψ_t obeys the uniform distribution. In the field environment, if the monitoring point on the device is constantly moving, then as the monitoring point moves, the relative position between the wireless transmitter and the obstacle will continue to change, making the amplitude of the carrier signal received by the concentrator, causing the fading of the propagating signal.

4. The Design of the Host Computer Monitoring System

4.1. Monitoring System Design Scheme. In view of the actual monitoring requirements of the hot water boiler of the group company, each node boiler is monitored by a host computer during the system design, and its monitoring system is also developed by Kingview software. On this basis, the main control computers of all nodes are networked, and the information is integrated and displayed in real time on the screen of the group company's hot water boiler operation status monitoring center. It should be pointed out that the system development of the monitoring center is also completed in the Kingview environment. This section describes the design and description of the development of the host computer monitoring system for a certain node's boiler operating status [23].

The design goal of the host computer monitoring system is to dynamically display the changes of important operating parameters of the boiler in the monitoring screen in real-time. You use Kingview's rich library and library development tools to monitor the operating status of a group company's boiler and create it according to the actual needs. The new screen is configured using the library, and then the corresponding relationship between the screen object and the system variable parameters is established through the animation connection and the command language. In addition, the system functions such as the control and command language provided by the Kingview are used in the boiler running state, the data is stored, and data reports are generated. Since Kingview is a real client-server model supported by the network as a platform, the monitoring node can be designed as an I/O server according to the network size. In the design of the boiler operating status monitoring system of a group company, using this feature of Kingview, the host computer is used as a client. The status parameters collected by the PLC of each node can realize real-time monitoring of the operation of all hot water boilers.

The host computer monitoring system adopts a modular design concept, which is mainly composed of a login interface, a main monitoring interface, a real-time trend curve interface, a real-time report interface, and an alarm window interface. The structure diagram is shown in Figure 3. In the main interface management system, the real-time data of the main parameters can be inquired in real time.

We select the configuration menu in the Kingview's project browser and set according to the following process: configuration⟶user configuration⟶user and security configuration. Three security levels have been set up: system administrators, management staff, and boiler workers. The security zone of the system administrator is configured as A, which is the highest level, followed by management staff and boiler workers. We set different system login passwords for personnel with three security levels. For example, when turning on the boiler control system, you can select the corresponding user type through the mouse or directly click on the touch screen according to the type of user.

You can also modify the password on the boot interface. You click the "Modify Password" button to pop up a subinterface to implement the password modification function. After entering the password on the boot interface, you press the "Enter" key on the keyboard. If the password is correct, the system enters the boiler system control interface. If the password is incorrect, a window will pop up to prompt that the password is incorrect, and the input can be continued after confirmation operation. If it is incorrect three times, it will take half an hour to reenter the password to be effective.

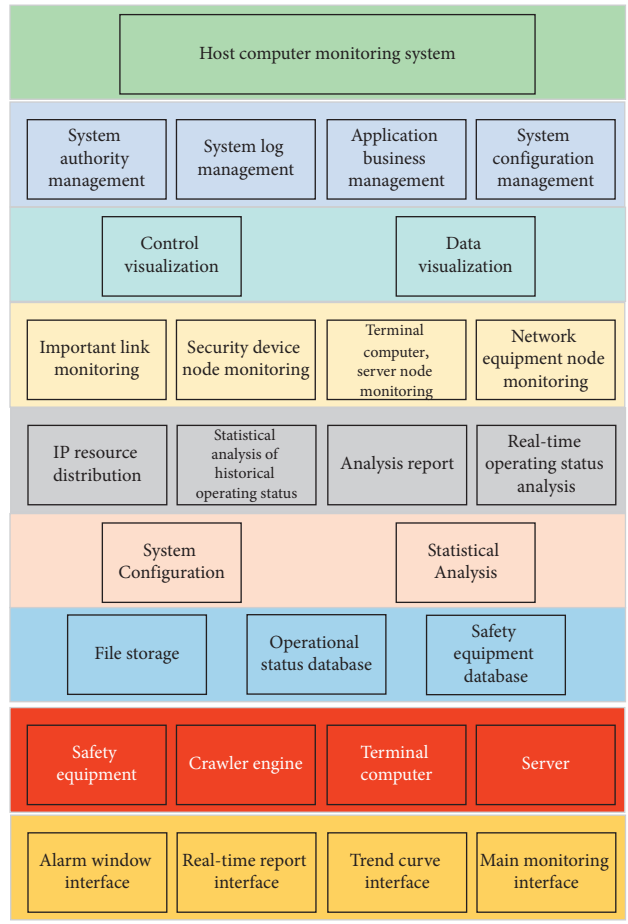


Figure 3: The structure of the host computer monitoring system.

4.2. Monitoring and Alarm Design. In the operation of the monitoring system, if there is an alarm, the alarm light turns red. After the staff finds an alarm event, they can use the alarm elimination button to eliminate the alarm and perform inspection and troubleshooting. The water inlet and outlet pipes can dynamically display the direction of water flow. At the same time, the user can directly enter the window screen by clicking each window button to view the current data display status. The host computer establishes communication with the PLC through Kingview and uploads the data collected by the PLC to the Kingview database through OPCServer in real time. The upper computer displays the collected data in real time. The alarm screen records the event date, event time, alarm date, alarm time,

alarm variable name, alarm type, alarm value/old value, and restored value/new value through the alarm window; the real-time trend screen is used to display the current working parameters of the hot water boiler. The historical data record of each monitoring equipment parameter can be called up in the historical curve screen and displayed in the form of a curve, so that the staff can check the historical data record at any time at any time; the report screen can query and export historical reports and real-time reports.

On one hand, the function of the alarm display is to drive the signal light for the duty personnel to check; on the other hand, the alarm is recorded in the data table for the administrator and supervisor to check. The alarm confirmation button is mainly to check whether the staff on duty found and dealt with the alarm. It is used to distinguish responsible accidents. The alarm system is convenient, flexible, reliable, and easy to expand. Distributed alarm management provides multiple management functions. Alarms and events have multiple output methods: files, databases, printers, and alarm event alarms including switch alarms and analog alarms. According to the data types collected by the field equipment, there are two types of alarms: one is a digital alarm; the other is an analog alarm. Through the communication channel uploaded to the alarm window of the monitoring host, the user can intuitively monitor and view the alarm information of the system and operate and control the running status of each subsystem. You select a button box in the KingView toolbar, determine the size, then select the animation connection realized by clicking the button, and select the required function in the animation connection.

When an alarm is generated, the alarm window automatically pops up on the desktop, and the report of the alarm information also pops up on the desktop. The alarm value exceeding the upper and lower limits can be seen in the alarm information report, and the sound, light, and voice alarms are automatically activated. The upper computer system will record the alarm time, and the lower computer will transmit the alarm record to the upper computer database in real time and display it in the form of a report. When the number of records in the buffer reaches the upper limit preset by the system, the system will automatically transfer the alarm information to the alarm database, and the printer can print the alarm records in the form of reports in real time.

4.3. Monitoring Data Storage and Query Design. Based on the overall architecture of a hot water boiler monitoring system of a group company, this section gives a further explanation of the database design. Because this monitoring system adopts a distributed structure, at each monitoring node, the real-time operating parameters of the corresponding boiler are first stored in the database of the host computer of the PLC monitoring subsystem of the node. This feature provides data acquisition for the monitoring center of the group company. There are two ways. One is that the real-time data obtained from the sensor system is stored in the node database and displayed on the monitoring interface of the node PLC system in real time and transmitted to the company's centralized control center through the network for storage and displayed on the monitoring main interface in real time. In the other way, the centralized control center directly retrieves the data from the upper computer database of the node PLC monitoring subsystem and displays it on the monitoring interface in real time. This data storage method of the system not only provides a data backup method, but also provides a means to verify the validity of the data. For example, when an accident occurs during the operation of a node's boiler, the data stored in the PLC host of the node can be compared with the data stored in the centralized control center to prevent data tampering and help determine the responsibility for the accident. In the specific design, this article uses the first data storage method.

In addition to the real-time upload of boiler operating parameters to the host computer system during the operation of the monitoring system, the control commands of the monitoring center to the nodes are also transmitted to the on-site node PLC controller in real time. When the upper computer and the lower computer are transmitting data, they complete the whole work according to the software flow in Figure 4.

4.4. Host Computer Network Release. The main function of network publishing is to realize the network browsing of the Kingview project, that is, to realize Internet/Intranet access. You configure the WEB publishing project in the configuration software and first configure the network for the data server in the project browser.

You select the "Networking" mode in the network parameter property page and enter the node name 219.224.14.78 corresponding to this machine in the local node. Then, you check each option when configuring the node type, because we use a Kingview project to act as all the servers required by the data server. Secondly, we set up the WEB publishing screen project and then configure the WEB server, including network parameters, node types, and client configuration. On the Kingview server, users can publish real-time information, historical information, and database information in real time. On the browser client, users can view real-time information, historical information, database information, and user management information by logging in to the main WEB interface.

In this article, the system's network publishing is through the use of a fixed IP (219.248.11.73) for WEB public network publishing. Therefore, you must go to the telecommunications department to apply for opening port 80. Then, you connect the public network directly to the data server (the data server and the WEB server are on the same server), and you can start network publishing. According to PLC's actual demand for on-site data collection of the 19 subnodes of the hot water boiler, the information of each collection point is uploaded to the designated server IP. You store the collected information in the corresponding IP database server, use the configuration software client to load, and then use the Kingview Web publishing function to publish the relevant screens, reports, and databases on the network.

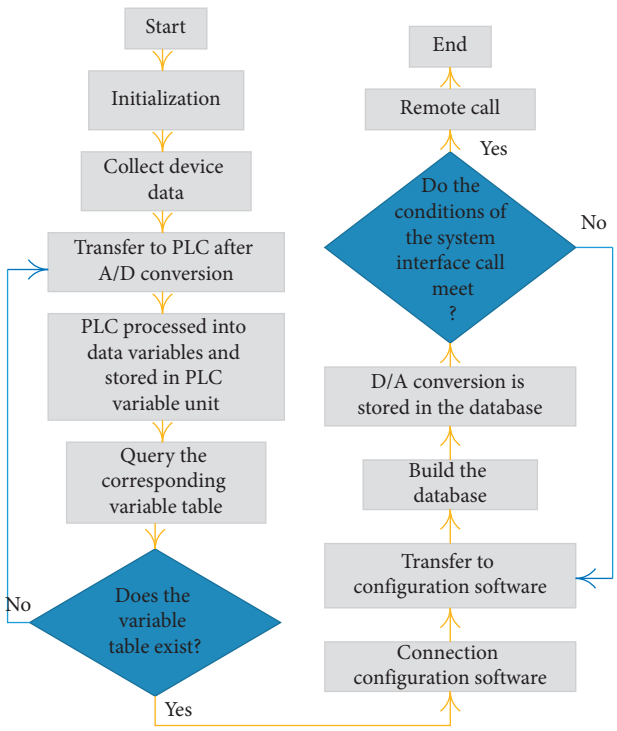


Figure 4: Storage of boiler operating status data.

Once the amount of data is relatively large, we can also store real-time images, daily reports, and real-time curves on different servers in the centralized control center. Each server has a fixed IP address. These servers are networked through a local area network to form the main server. The remote monitoring center can view the content you want to view by accessing different IP addresses. At the same time, we can also browse the project text of the hot water boiler that we have built in the IE browser client by adding trusted sites.

5. Control System Parameter Tuning and Performance Simulation

5.1. Control System Parameter Tuning. If we want to evaluate the control quality of a regulation system, we must observe its transition process under the same step disturbance. When composing the control system, we must determine the static and dynamic characteristics of each channel of the object and, at the same time, determine the control plan. Therefore, the control quality mainly depends on the tuning of the controller parameters. The task of parameter tuning is to appropriately adjust the proportional coefficient KP, integral coefficient KI, and differential coefficient KD (or proportional gain KC, integral time TI, and derivative time TD) of the controller according to a certain control loop scheme to obtain a quality production process. Generally speaking, we always hope that the system adjustment process can be "stable, accurate, and fast." But according to automatic control theory, it is difficult to meet, so it can only be based on the quality requirements of a specific production process.

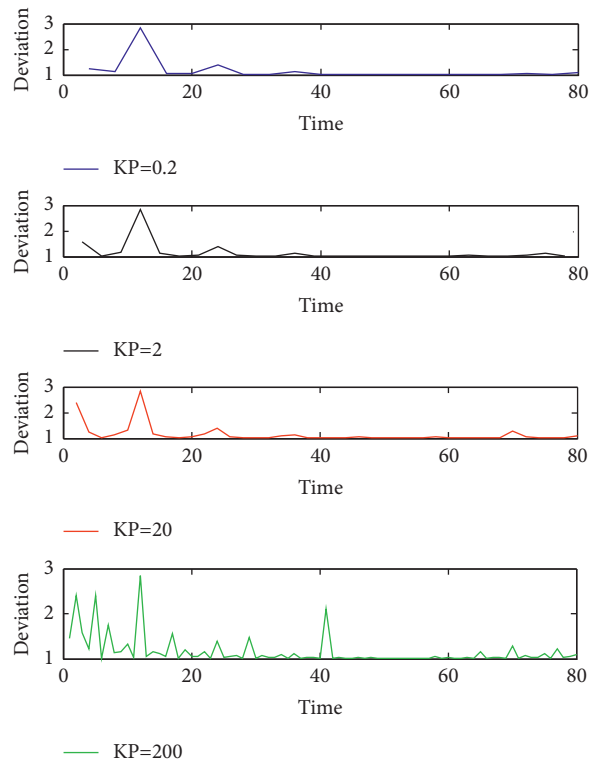


Figure 5: The influence of KP on the transition process (KI = KD = 0).

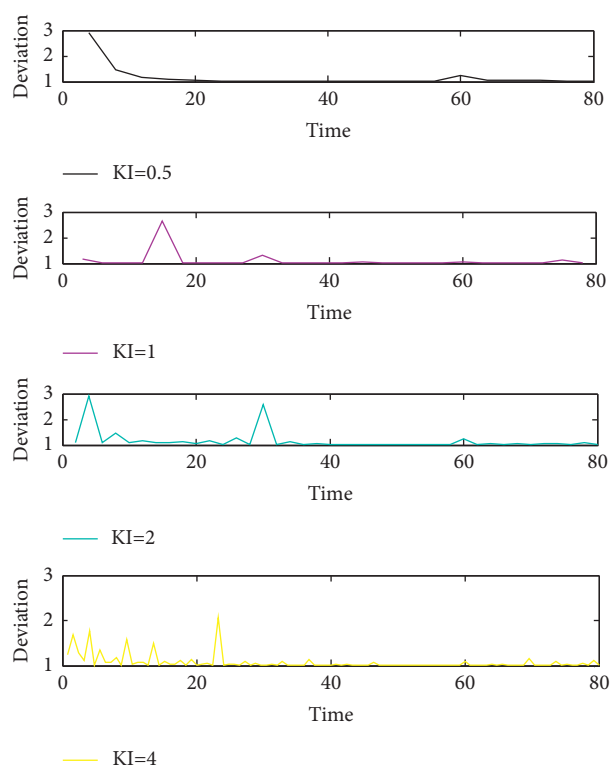


Figure 6: The impact of KI on the transition process (KP = 2, KD = 0).

The idea of the engineering tuning method is "look at the curve, adjust the parameters," which is based on the experiment, record the transition process curve and obtain the data, then look up the table to obtain the PID parameters of the regulator, and finally "look at the curve" in the system to obtain the "best" transition process that meets the process quality requirements. Because engineering tuning is an experimental trial and error method, in order to quickly make a trial, you must master the law of the controller's proportional coefficient KP, integral coefficient KI, and differential coefficient KD to the transition process curve to reduce the blindness of trial and error.

Figure 5 shows that as KP changes from small to large, the dynamic deviation of the system becomes smaller, the stability becomes worse, and the system duty cycle becomes smaller. Figure 6 shows that as KI changes from small to large, the integral action increases, and the transition process of the system changes from stable to oscillation, and the overshoot increases. The integral action can eliminate the deviation, which is the most important feature of the integral action. Figure 7 shows that, with the increase of KD, the differential effect increases, the overshoot of the closed-loop system decreases, and the system response speed becomes slower.

In addition, before the control system is put into operation, in addition to the preparation of the control system equipment itself, the production and operating conditions of the controlled boiler, the operating status of the process equipment, and the actuators directly related to the control function should also be grasped. Therefore, it is necessary to conduct field experiments within the scope of the boiler production process permit to obtain the characteristic parameters of each control object and provide a basis for parameter setting during the commissioning process.

5.2. Simulation and Parameter Tuning of Water Supply Control System. At $T = 0$ s, a step disturbance of −2.2 t/h of steam flow is added, and a step disturbance of −2.2 t/h of feedwater flow is added at $T = 300$ s. In the PID parameter tuning module of IBCCS-E, multiple parameter settings and comparisons can be performed on the same controller until the result is satisfied. The principle of tuning is to tune the inner loop first, and then the outer loop. Here, the tuning result of the inner loop is directly given. We select the PID1 of the feedwater cascade main controller, its PID parameters before tuning are KP = 3, KI = 0.1, KD = 0, and set its new PID parameters as KP = 15, KI = 0.2, KD = 0. The system will give the response curve of the drum water level and feedwater flow before and after the setting, as shown in Figure 8.

It can be seen from Figure 8 that after a step disturbance of −2.2 t/h occurs in the steam flow at $T = 0$ s, the feedwater flow tends to 2~6 t/h after an oscillation process; the drum water level still recovers after an oscillation process. Compared with the status before and after adjustment, the adjustment time of the system after adjustment is significantly shorter, the adjustment speed is accelerated, the overshoot of the drum water level is significantly reduced, the overshoot of the feedwater flow is not much different from that before the setting, and the adjustment process is greatly improved.

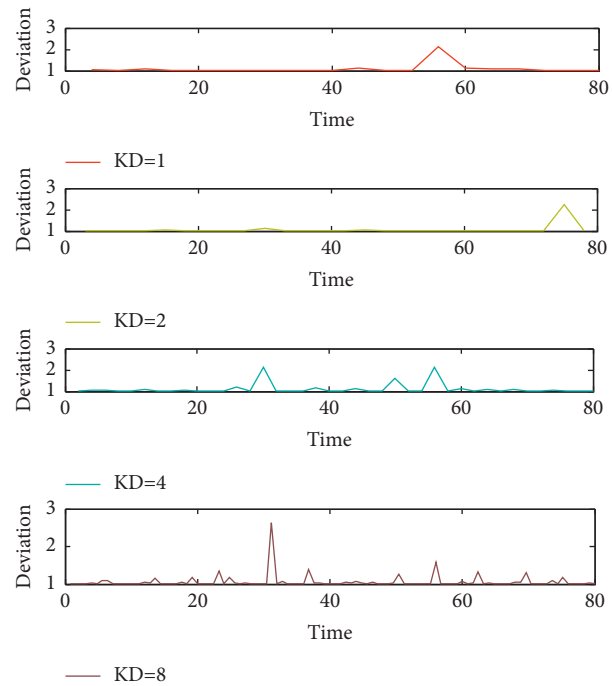


Figure 7: The influence of KD on the transition process (KP = 2, KI = 0.5).

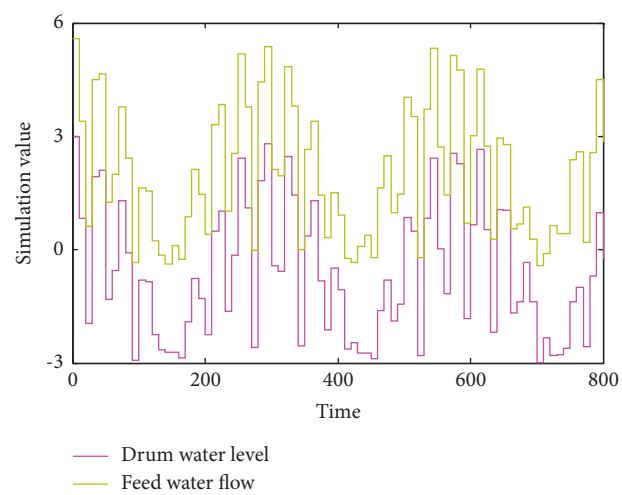


Figure 8: Simulation results of the main controller of the water supply control system.

After a step disturbance of −2.2 t/h occurs in the feedwater flow at $T = 300$ s, the secondary circuit acts as a follow-up system, and its given value is the output of PID1 of the main controller. Before parameter setting, the output of PID1 cannot be stabilized immediately However, the feedwater flow rate still fluctuates slightly for a long time; after the parameters are set, the output of PID1 stabilizes immediately, and finally, the feedwater flow rate and drum water level hardly move, and the feedwater flow disturbance is quickly eliminated.

It can be seen that the control quality after tuning is significantly better than the control quality before parameter

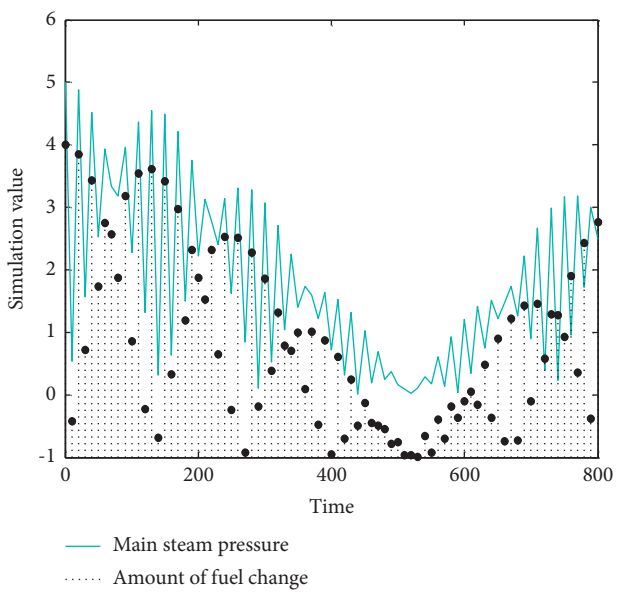


Figure 9: Parameter tuning and simulation results of the main steam pressure controller of the combustion control system.

tuning. Therefore, the parameters of PID1 are finally determined as KP = 2, KI = 0.2, and KD = 0.

5.3. Simulation and Parameter Tuning of Combustion Control System. At $T = 0$ s, a step disturbance of 30 kg/h of fuel quantity is added, and a step disturbance of steam load of -1 t/h is added at $T = 530$ s. In the PID parameter tuning module of IBCCS-E, multiple parameter settings and comparisons can be performed on the same controller until the result is satisfied. The setting sequence is to set the induced air regulation subsystem first, then the air supply regulation subsystem, then the fuel quantity control subsystem, and finally the main steam pressure control loop. We select PID3, its parameters before tuning are KP = 20, KI = 0.5, KD = 0, and set its new parameters as KP = 5, KI = 0.1, KD = 0. Figure 9 shows the response curve of main steam pressure and fuel quantity.

It can be seen from Figure 9 that after a step disturbance occurs in the fuel quantity at $T = 0$ s, the fuel quantity feedback signal will change, causing the signal imbalance at the inlet of the fuel regulator PID4 to change its output signal and actuate the fuel actuator. The fuel quantity is quickly restored to the value before the disturbance. Before and after tuning, although the main steam pressure overshoot is small before tuning, the main steam pressure and fuel volume will both produce continuous small oscillations; the overshoot of the system after tuning is not much different from that before tuning, and the adjustment trend slows down. No oscillations will occur, and the adjustment process will be improved.

After a step disturbance of −1 t/h occurs in the steam flow at $T = 900$ s, the main steam pressure Pm changes immediately and deviates from the set value, and after another oscillation process, it eventually returns to the set value; the fuel quantity goes through an oscillation process. After that, it finally stabilizes at a new stable value that is compatible with the steam flow rate. Compared with before and after tuning, although the main steam pressure overshoot is small before tuning, the main steam pressure and fuel quantity will both produce continuous small amplitude oscillations; the overshoot of the system after tuning is slightly larger than before tuning, and the adjustment trend slows down. But in the end, it reaches a stable state, no oscillation occurs, and the adjustment process is improved.

6. Conclusion

This paper designs the main monitoring screens, daily reports, real-time trend graphs of main parameters, alarm windows, and other screens of industrial hot water boilers and realizes the remote monitoring of hot water boiler data and real-time release of images using the network release provided by KingView. The focus is on the data collection work of the hot water boiler monitoring system and the design of the upper computer man-machine interface. This article introduces the method of parameter tuning of the control system, establishes a simulation model of the industrial boiler computer control system, provides an interface between the host computer control software and the simulation model, conducts a controller parameter tuning simulation experiment, and carries out the experimental results analysis. The upper computer monitors the status of the two lower computers in real time and provides the fault information of the lower computers in time. Once the faulty lower-level computer is repaired, it will automatically start to the standby state once it is powered on. This structure decentralizes the functions of the system and improves reliability. The system can also flexibly increase or decrease input and output modules to realize one machine one control (one lower machine controls one boiler) or one machine multi-control (one lower machine controls multiple boilers). The other parts have no effect. The system configuration is flexible, and the expandability is good. Because the collection signal type of the hot water boiler monitoring system is relatively simple, it is limited to the signal collection related to the control element, such as the collection of the input and output point status of the PLC, the collection of the status of the internal auxiliary relay, and the collection of the data in the data register. However, some parameters such as self-starting and manual, PID control constant temperature and pressure are not considered in the design. If you can study the PID control and participate in the control, it will be a new breakthrough in the future.

Data Availability

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

Conflicts of Interest

There are no conflicts of interest in this article.

Acknowledgments

The study was supported by Shijiazhuang University Development and Construction Projects of Applied Coursework (YYKC-202003); Shijiazhuang University Teaching Reform Research and Practical Projects (JGXM-202155G); and Hebei Province Applied Technology University Seminar Project (JY2021023).

References

[1] Y. Jiang, S. Yin, and O. Kaynak, "Data-driven monitoring and safety control of industrial cyber-physical systems: basics and beyond," *IEEE Access*, vol. 6, pp. 47374–47384, 2018.

[2] H. M. La, T. H. Dinh, N. H. Pham, Q. P. Ha, and A. Q. Pham, "Automated robotic monitoring and inspection of steel structures and bridges," *Robotica*, vol. 37, no. 5, pp. 947–967, 2019.

[3] Y. Cao, L. Ma, and Y. Zhang, "Application of fuzzy predictive control technology in automatic train operation," *Cluster Computing*, vol. 22, no. 6, pp. 14135–14144, 2019.

[4] V. Tsakanikas and T. Dagiuklas, "Video surveillance systems-current status and future trends," *Computers & Electrical Engineering*, vol. 70, pp. 736–753, 2018.

[5] M. Ali, S. Qaisar, M. Naeem, and S. Mumtaz, "Energy efficient resource allocation in D2D-assisted heterogeneous networks with relays," *IEEE Access*, vol. 4, pp. 4902–4911, 2016.

[6] C. P. Janssen, S. F. Donker, D. P. Brumby, and A. L. Kun, "History and future of human-automation interaction," *International Journal of Human-Computer Studies*, vol. 131, pp. 99–107, 2019.

[7] O. Carsten and M. H. Martens, "How can humans understand their automated cars? HMI principles, problems and solutions," *Cognition, Technology & Work*, vol. 21, no. 1, pp. 3–20, 2019.

[8] B. Li, R. Lu, W. Wang, and K.-K. R. Choo, "Distributed host-based collaborative detection for false data injection attacks in smart grid cyber-physical system," *Journal of Parallel and Distributed Computing*, vol. 103, pp. 32–41, 2017.

[9] H. C. Lee and K. H. Ke, "Monitoring of large-area IoT sensors using a LoRa wireless mesh network system: design and evaluation," *IEEE Transactions on Instrumentation and Measurement*, vol. 67, no. 9, pp. 2177–2187, 2018.

[10] B. F. Spencer Jr., V. Hoskere, and Y. Narazaki, "Advances in computer vision-based civil infrastructure inspection and monitoring," *Engineering*, vol. 5, no. 2, pp. 199–222, 2019.

[11] M. Hossain, S. M. R. Islam, F. Ali, K.-S. Kwak, and R. Hasan, "An internet of things-based health prescription assistant and its security system design," *Future Generation Computer Systems*, vol. 82, pp. 422–439, 2018.

[12] R. Casado-Vara, P. Chamoso, F. De la Prieta, J. Prieto, and J. M. Corchado, "Non-linear adaptive closed-loop control system for improved efficiency in IoT-blockchain management," *Information Fusion*, vol. 49, pp. 227–239, 2019.

[13] A. Mosenia, S. Sur-Kolay, A. Raghunathan, and N. K. Jha, "Wearable medical sensor-based system design: a survey," *IEEE Transactions on Multi-Scale Computing Systems*, vol. 3, no. 2, pp. 124–138, 2017.

[14] R. Y. Zhong, L. Wang, and X. Xu, "An IoT-enabled real-time machine status monitoring approach for cloud manufacturing," *Procedia CIRP*, vol. 63, pp. 709–714, 2017.

[15] G. Nagarajan and R. I. Minu, "Wireless soil monitoring sensor for sprinkler irrigation automation system," *Wireless Personal Communications*, vol. 98, no. 2, pp. 1835–1851, 2018.

[16] D. Zhang, "High-speed train control system big data analysis based on the fuzzy RDF model and uncertain reasoning," *International Journal of Computers Communications & Control*, vol. 12, no. 4, pp. 577–591, 2017.

[17] M. Heimberger, J. Horgan, C. Hughes, J. McDonald, and S. Yogamani, "Computer vision in automated parking systems: design, implementation and challenges," *Image and Vision Computing*, vol. 68, pp. 88–101, 2017.

[18] I. Y. Noy, D. Shinar, and W. J. Horrey, "Automated driving: safety blind spots," *Safety Science*, vol. 102, pp. 68–78, 2018.

[19] A. F. Hussein, N. A. Kumar, M. Burbano-Fernandez, G. Ramirez-Gonzalez, E. Abdulhay, and V. H. C. De Albuquerque, "An automated remote cloud-based heart rate variability monitoring system," *IEEE Access*, vol. 6, pp. 77055–77064, 2018.

[20] S. Rodríguez, T. Gualotuña, and C. Grilo, "A system for the monitoring and predicting of data in precision agriculture in a rose greenhouse based on wireless sensor networks," *Procedia Computer Science*, vol. 121, pp. 306–313, 2017.

[21] M. Riahi Manesh and N. Kaabouch, "Analysis of vulnerabilities, attacks, countermeasures and overall risk of the automatic dependent surveillance-broadcast (ADS-B) system," *International Journal of Critical Infrastructure Protection*, vol. 19, pp. 16–31, 2017.

[22] M. Shahid, L. Henrik, M. S. H. Kazi, R. Jonathan, and R. Ayman, "Smart direct-LTE communication: an energy saving perspective," *Ad Hoc Networks*, vol. 12, pp. 296–311, 2014.

[23] M. Z. Khan, S. Harous, S. U. Hassan, M. U. Ghani Khan, R. Iqbal, and S. Mumtaz, "Deep unified model for face recognition based on convolution neural network and edge computing," *IEEE Access*, vol. 7, pp. 72622–72633, 2019.