

Research Article

Deep Learning Model of Image Classification Using Machine Learning

Qing Lv, Suzhen Zhang , and Yuechun Wang

Shijiazhuang Posts and Telecommunications Technical College, Shijiazhuang 050021, China

Correspondence should be addressed to Suzhen Zhang; zhangsz_2022@163.com

Received 26 May 2022; Revised 19 June 2022; Accepted 29 June 2022; Published 19 July 2022

Academic Editor: Qiangyi Li

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

Not only were traditional artificial neural networks and machine learning difficult to meet the processing needs of massive images in feature extraction and model training but also they had low efficiency and low classification accuracy when they were applied to image classification. Therefore, this paper proposed a deep learning model of image classification, which aimed to provide foundation and support for image classification and recognition of large datasets. Firstly, based on the analysis of the basic theory of neural network, this paper expounded the different types of convolution neural network and the basic process of its application in image classification. Secondly, based on the existing convolution neural network model, the noise reduction and parameter adjustment were carried out in the feature extraction process, and an image classification depth learning model was proposed based on the improved convolution neural network structure. Finally, the structure of the deep learning model was optimized to improve the classification efficiency and accuracy of the model. In order to verify the effectiveness of the deep learning model proposed in this paper in image classification, the relationship between the accuracy of several common network models in image classification and the number of iterations was compared through experiments. The results showed that the model proposed in this paper was better than other models in classification accuracy. At the same time, the classification accuracy of the deep learning model before and after optimization was compared and analyzed by using the training set and test set. The results showed that the accuracy of image classification had been greatly improved after the model proposed in this paper had been optimized to a certain extent.

1. Introduction

Not only is using manual method to select image features time-consuming and laborious, but also the feature quality depends on professional knowledge and practical experience to a certain extent, which limits the application of manual feature selection method [1]. With the advent of the era of big data and artificial intelligence, the traditional artificial way to obtain image information has been unable to meet the needs of applications in related fields. In recent years, the method of obtaining images by computer has made some progress, but, for the massive image information processing, the traditional computer vision is still inefficient. For example, traditional machine vision methods have low accuracy in detecting low-level features and high-level features of images [2]. Classification and processing of massive image information is the prerequisite for in-depth mining of image

information. It is also of great significance for the expansion of computer vision in the field of image processing application. At present, the research of image classification by computer has been gradually applied to agriculture, industry, aerospace, and other fields. The research of image classification mainly includes efficient and accurate classification of large-scale image information and accurate classification of image semantic similar feature information.

In order to recognize and classify the image, we need to use the image acquisition tool to obtain the original image information, then clean the image data and extract and filter the features, and finally use the machine learning algorithm to realize the image classification. With the deepening of the application of machine vision in many fields, people can use computer vision technology to quickly realize the processing operations of image information, such as noise reduction and feature extraction [3]. The research of image

classification based on machine learning has made some progress. From the adoption of artificial recognition to the continuous application of computer vision technology in image classification, scholars have made some achievements in the research of image classification and recognition.

Machine learning is an important branch of artificial intelligence. Although machine learning has experienced half a century of development, there are still some unsolved problems, for example, complex image understanding and recognition, natural language translation, and recommendation system [4]. Deep learning is an important branch developed on the basis of machine learning. It makes full use of the hierarchical characteristics of artificial neural network and biological neural system to process information and obtains high-level features by learning low-level features and adopting feature combination method, so as to realize image classification or regression. Different from traditional machine learning, deep learning uses multilayer neural network to automatically learn the image and extract the deep-seated features of the image. Different depth learning models can be formed according to different feature learning and its combination. However, the accuracy of image classification is not high and the operation efficiency of the existing deep learning model is low. Therefore, based on the existing basic theory of convolution neural network, this paper establishes a deep learning model of image classification by improving the structure of convolution neural network, which provides a basis for image classification under complex environmental conditions.

2. Related Works

Image classification provides an important basis for image depth processing and the application of computer vision technology in related fields. Traditional image classification mainly goes through different stages, such as image preprocessing, feature extraction, classifier construction, and learning training [5]. Traditional image classification methods mainly use the extracted basic image features to realize image classification, which can provide a basis for further obtaining the semantic information of images by computer. Traditional image classification generally uses image color, texture, and other information to calculate image features and uses support vector machine and logistic regression to realize image classification [6]. The results of image classification not only depend on the extracted features to a great extent but also are affected by the knowledge and experience of relevant fields.

Not only are the manually acquired features difficult to apply to image classification but also a lot of time is spent in analyzing feature data. At the same time, the traditional machine learning cannot be applied to the processing of large datasets, and it is difficult to realize the optimization of feature design, feature selection, and model training, which makes the classification effect of the model poor. Therefore, image classification methods using traditional machine learning are affected in many application fields [7]. Research shows that because texture, shape, and color features can be used for image classification and recognition, low-level basic

features can be used as the basis of image classification. Traditional image classification methods generally use single feature extraction or feature combination and take the extracted features as the input value of support vector machine. In recent years, some progress has been made in image classification using artificial neural network classifier. In order to improve the accuracy of image classification, we can focus on the standardized design of low-level features such as texture, shape, and color.

Deep learning realizes the training of large-scale datasets through multilevel network model and adopts the method of layer-by-layer feature extraction to obtain the high-level features of the image. Not only is the deep learning network model used to extract the basic features of the image but also it can obtain the deep features of the image through multiple hidden layers. Compared with traditional machine learning methods, the features obtained by deep learning method are not only accurate but also conducive to image classification. In the process of image recognition and classification, the way of feature learning and combination is mainly determined by the deep learning model [8]. At present, the commonly used deep learning models are sparse model, restricted Boltzmann machine model, and convolution neural network model. Although these models have some differences in feature extraction, they have similarities in image classification and recognition. They all go through the steps of image information input, data preprocessing, feature extraction, model training, and classification output.

In image classification, some scholars have carried out a lot of research work on image feature representation and classifier selection. For example, the deep learning model based on feature representation can be effectively applied to the recognition and classification of various images. Some scholars use deep convolution neural networks (DCN) to deeply extract image features and apply them to large-scale dataset ImageNet [9]. Experiments show that the model can effectively classify large data image sets. In addition, the deep learning model can effectively learn and describe image features. For example, the deep learning model can better describe the hierarchical features through unsupervised learning, and the features extracted by the model not only have strong expression ability but also improve the efficiency of image classification. A large number of research results show that, with the deepening of the research of image classification methods in related fields, deep learning model has gradually replaced the traditional artificial feature extraction and machine learning methods and will be widely concerned by scholars in the field of image recognition and classification [10].

3. Fundamentals of Image Classification Algorithm

3.1. Basic Theory of Neural Network. The traditional neural network, referred to as artificial neural network (ANN), is a hot spot in the field of early artificial intelligence [5]. Artificial neural network mainly uses the neurons of network model to abstract the characteristics of external things, so as to be used by computer to complete information processing.

Artificial neural network generally establishes the corresponding network structure according to the different construction methods of neurons. Neural network is an operation model composed of several different nodes or neurons connected with each other. Each node in the model is a processing function, and the connection between different nodes uses weight to represent the memory ability of artificial neural network. The output of neural network depends on the connection form, weight value, and excitation function of different nodes. At the same time, the neural network model is mainly constructed according to some algorithm or function to express some specific logical operation.

A basic neural network model usually includes information input layer, hidden layer, and calculation result output layer. Different layers can contain several neurons [11]. Neurons represent a transformation or operation, which is completed by the activation function of neurons. Two adjacent layers of neurons are connected to each other, as shown in Figure 1.

As can be seen from Figure 1, the neural network model includes 11 neurons: 3 input layers, 5 hidden layers, and 3 output layers. The structure belongs to a two-layer neural network model, where W_1 and W_2 are the weight matrices of the hidden layer and the output layer, respectively.

Deep learning method is a part of machine learning. It is widely used in natural language recognition and image detection and classification. Moreover, deep learning comes from the theory of artificial neural network. By referring to the human brain for hierarchical processing of information, different levels of neural networks are established. Deep learning effectively extracts multilevel feature information by simulating human brain, so as to obtain the key feature information of image, text, and other data. Deep learning mainly describes the specific object characteristics through hierarchical processing according to a large amount of edge feature information. It is a process from low-level feature extraction to high-level feature combination. As an important method of machine learning, deep learning is an effective method to process big data and obtain abstract features by using neural network model.

Deep learning is a multilayer deep neural network model. According to the connection law of human brain neurons, the sample features are processed in different layers of the model, and the deep features of the sample data are obtained in turn. Similar to deep neural network, artificial neural network belongs to hierarchical structure. The model is composed of multilayer perceptron, including input layer, hidden layer, and output layer. Different from the deep neural network, the artificial neural network model only contains two to three layers of forward neural network, and the number of neurons in each layer is small, so the processing ability of large datasets is limited. Because the deep neural network model contains many layers and each layer contains a large number of neurons, the neural network model not only can realize the abstract expression of data but also has strong learning function. Compared with the traditional machine learning methods, the deep learning model does not need to rely on manual design and feature

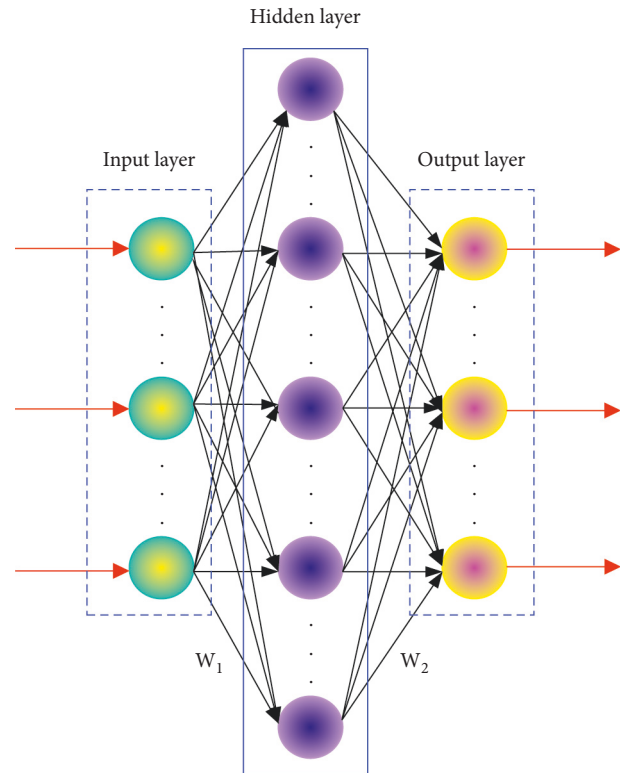


FIGURE 1: Connections between neurons in different layers.

extraction. It not only solves the shortage of manual knowledge but also avoids the preference problem in feature extraction. At the same time, the deep learning model can obtain representative high-level features through the organic combination of low-level features [12]. As shown in Figure 2, the working processes of deep learning and traditional machine learning are compared.

In the neural network model, the activation function is used to perform nonlinear operation on the input data of neurons in order to extract effective feature information from the original input data. Activation function is a nonlinear function. With the increase of the number of layers of neural network model, the most effective feature information can be obtained after many iterations and data training.

In order to realize the classification of features, Softmax function is often used as the activation function in the neural network model and used in the output layer of the model [13]. Softmax function is generally used as a classifier. The calculation formula is as follows:

$$S_k = \frac{\exp(R_k)}{\sum_{k=1}^n \exp(R_k)}. \quad (1)$$

In the above formula, n is the number of neurons in the current layer and R_k is the nonlinear transformation value of the k -th neuron in this layer.

Softmax is a classifier that can output different feature categories. The value of each neuron contained in Softmax can be considered as the probability of the corresponding category. The calculation process of Softmax function is shown in Figure 3.

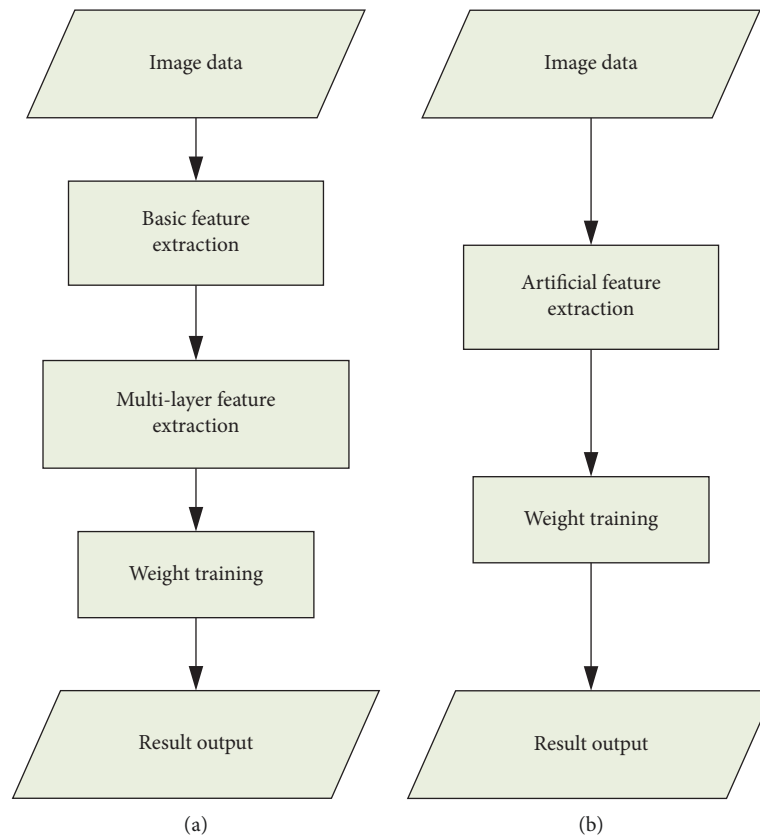


FIGURE 2: Comparison between deep learning model and traditional machine learning algorithm. (a) Deep learning algorithm. (b) Traditional machine learning algorithm.

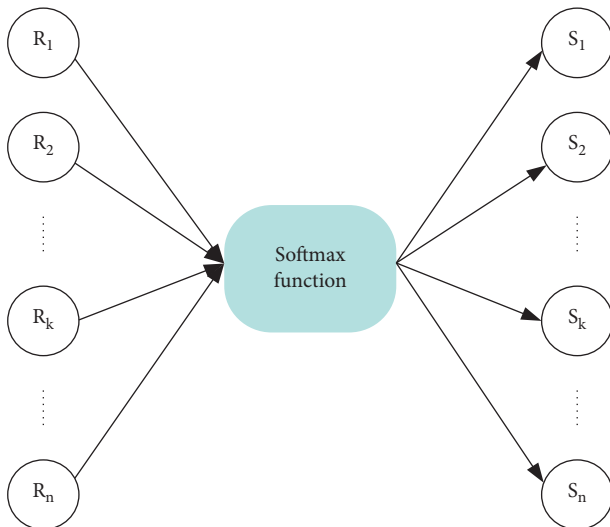


FIGURE 3: Schematic diagram of Softmax function calculation.

3.2. Basic Theory of Convolution Neural Network. Convolution neural network (CNN) is a typical network structure in deep learning model [14]. Different from traditional machine learning, convolution neural network can be better used for image and time series data processing, especially for image classification and language recognition.

The basic structure of convolution neural network is shown in Figure 4.

Convolution neural network generally includes three different types of processing layers: convolution layer, pooling layer, and full connection layer. Among them, the feature extraction task is completed by the convolution layer, and the pooling layer is used for feature mapping. The full connection layer is similar to the general neural network structure. All nodes in this layer are not connected to each other but completely connected to the nodes of the previous layer. In addition, like other neural networks, convolution neural networks also have data input layer and result output layer.

The calculation task of convolution neural network is mainly completed through the convolution layer, and the convolution kernel in the convolution layer is the core of the convolution neural network model. The convolution layer uses convolution check to convolute the input image and extract the characteristic information of the image. The images processed by convolution operation will gradually become smaller, and the pixels at the edge of the image have little effect on the output results.

As shown in Figure 5, assuming that the original input image is a 3×3 matrix, the original image is convoluted through a convolution check with a size of 2×2 , and the corresponding feature map is output.

Generally, there is a strong correlation between adjacent pixels in the image. Convolution kernel mainly extracts

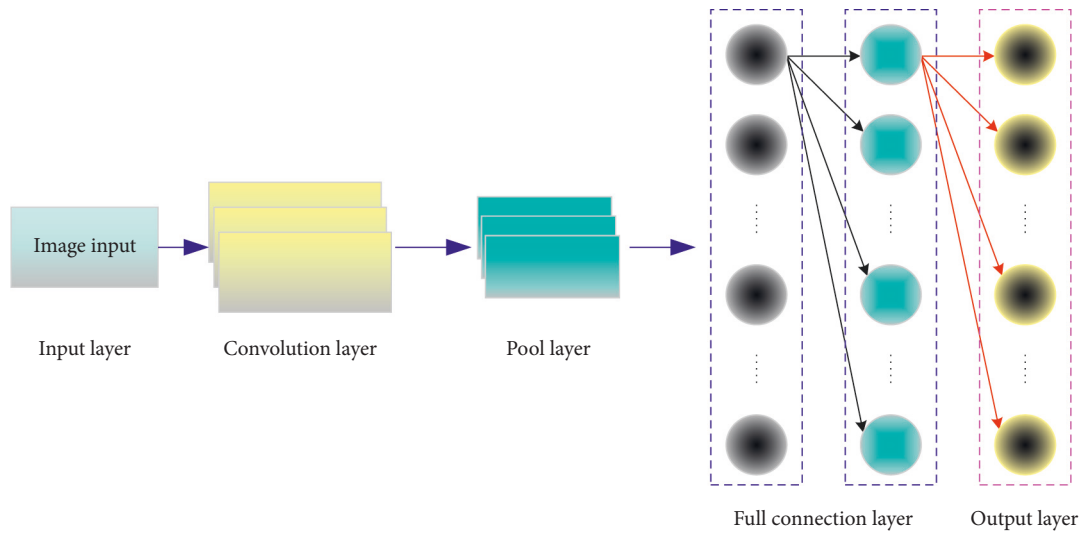


FIGURE 4: Structure diagram of convolution neural network.

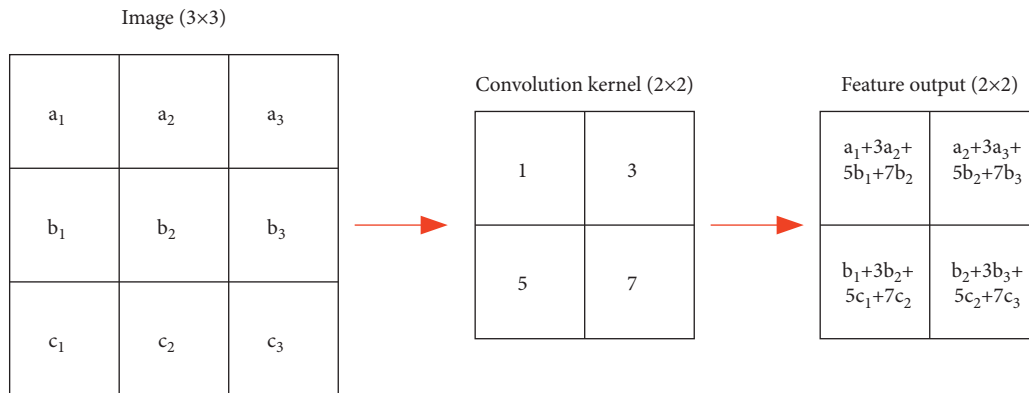


FIGURE 5: Schematic diagram of convolution operation process.

features from the local area of the image and sends the extracted local features to the high level for integration processing. Because the bottom feature of the image is independent of its position, it can not only use the same convolution check to extract the relevant features but also reduce the number of parameters of the neural network through the shared weight characteristic of the convolution kernel, so as to improve the training efficiency of the network model.

For complex images, in order to reduce the amount of parameter training of the model, the pooling layer in convolution neural network can be used to reduce the size of feature map. During pooling, the depth and size of the image can remain unchanged. The operation of pooling layer generally includes max pooling and average pooling [15], as shown in Figure 6.

3.3. Convolution Neural Network Model. In the existing image detection and recognition models, such as ResNet, Mask-RCNN, and Faster R-CNN models, they are usually based on common network models [16]. LeNet is the most basic convolution neural network model [9]. After transforming the LeNet model, the LeNet-5 model is established

to classify ordinary images. Because the LeNet-5 model is not deep enough, it cannot extract enough image features during model training. Therefore, it cannot be applied to the classification of complex images.

Therefore, some people put forward the AlexNet model based on the LeNet structure, applied the convolution neural network to the processing of complex images, and provided a theoretical basis for the application of deep learning model in the field of computer vision [17]. The network structure of AlexNet is shown in Figure 7.

AlexNet is a network structure with 8 layers. The model includes 5 convolution layers and 3 full connection layers. The model uses the ReLU function as the activation function to avoid the gradient dispersion phenomenon caused by the large number of layers of the network model. In order to reduce network training time, AlexNet uses multiple GPUs for training. In order to suppress neurons with small response ability, AlexNet uses LRN (local response normalization) processing layer to establish a competition mechanism for neurons, so as to make the value with large response ability increase continuously, so as to increase the generalization ability of the model. In addition, in order to prevent neurons from forward propagation and back

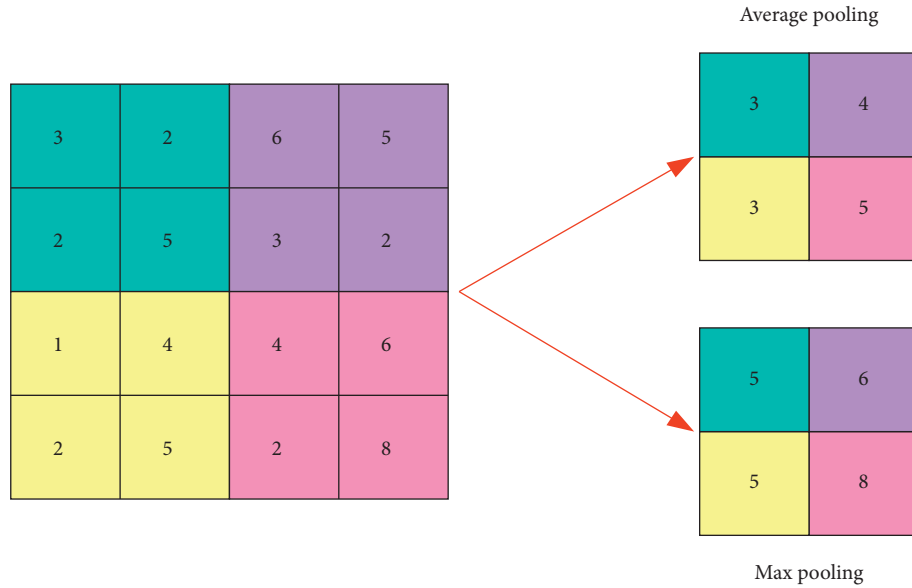


FIGURE 6: Schematic diagram of pool layer operation type.

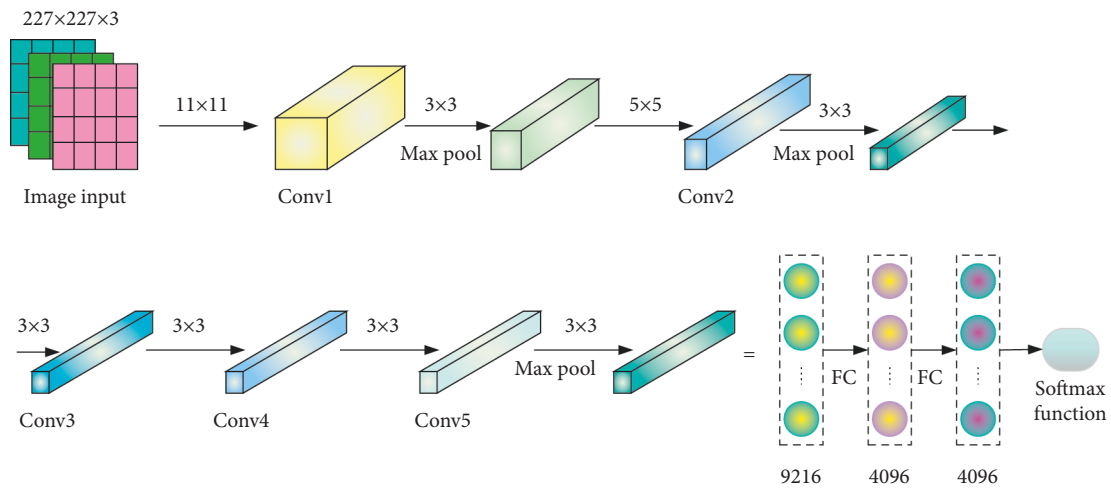


FIGURE 7: Schematic diagram of AlexNet network model structure.

propagation, the dropout mechanism is adopted in the full connection layer of AlexNet model, so that the output results of all hidden layer neurons are 0, so as to avoid the complex interaction between neurons.

At present, VggNet is a widely used deep convolution neural network model. Compared with other models, VggNet not only has better generalization ability but also can be effectively used for the recognition of different types of images [11]. For example, convolution neural networks such as FCN, UNet, and SegNet are based on VggNet model. In recent years, Vgg-16 and Vgg-19 networks have been commonly used for VggNet models [18]. The network structure of Vgg-16 is shown in Table 1.

The Vgg-16 network structure has 16 layers in total, excluding the pooling layer and Softmax layer. The convolution core size is 3×3 , the pooling layer size is 2×2 , and the pooling layer adopts the maximum pooling operation with step size of 2.

TABLE 1: The network structure of Vgg-16.

Layer type	Convolution kernel size	Characteristic diagram size	Convolution kernel number
Input layer	—	448×448	—
Convolution layer	3×3	448×448	128
Pool layer	3×3	448×448	—
Convolution layer	3×3	224×224	256
Pool layer	2×2	112×112	—
Convolution layer	3×3	112×112	768
Pool layer	2×2	56×56	—
Full connection layer	4096	—	—
Full connection layer	4096	—	—
Full connection layer	1000	—	—
Softmax classifier	—	—	—

Vgg-16 network uses convolution blocks instead of convolution layers, in which each convolution block contains 2~3 convolution layers, which is conducive to reducing the network model parameters. At the same time, Vgg-16 network adopts ReLU activation function to enhance the training ability of the model. Although Vgg model has more layers than AlexNet model, the convolution kernel of Vgg model is smaller than that of AlexNet model. Therefore, the number of training iterations of Vgg model is less than that of AlexNet model.

4. Image Classification Model Based on Improved CNN

4.1. Improved Image Classification Model Framework. Because image classification algorithms are usually used in systems with high real-time requirements, image classification algorithms need to consider real-time performance. For complex neural network models, image classification needs to consume a lot of time. Therefore, this paper simplifies VggNet model and takes it as the model basis of image classification.

Considering the distribution characteristics of datasets used for model training, a typical dataset can be selected as the weight of the model to initialize the training dataset. When the model is pretrained and reaches a certain accuracy, the number of nodes in Softmax layer is reduced by ten times, and then the dataset is used for weight training. Considering that the data processed by the model may be affected by various noises, a noise reduction automatic encoder is added to the model to eliminate the noise interference, and the existing dataset is extended through the data enhancement method to enhance the generalization ability of the model.

Considering that the image classification algorithm needs to meet certain real-time performance, the corresponding image classification model is established and optimized based on VggNet model. Among them, the algorithm combining convolution neural network and noise reduction automatic encoder can be used. Because there may be overfitting problem in image classification, it can be optimized by data enhancement. Compared with other algorithms, this classification algorithm has certain generalization performance in the case of small amount of data. In addition, the algorithm also adds a noise reduction automatic encoder, which can effectively reduce the impact of data noise on the performance of the model, so as to ensure that the model has good generalization ability. Since the improved algorithm is based on VggNet network, the training time of the model may increase [19]. The improved image classification algorithm is shown in Figure 8.

In order to solve the problem of automatic noise reduction of complex image structure, the normalized network of encoder is used for classification in this paper. Based on the existing convolution neural network model, the noise reduction automatic encoder and sparse automatic encoder are organically combined, and the input original image information is normalized on the sparse automatic encoder. Then, the improved convolution neural network model is

used to extract the image feature information, and the Softmax classifier is used to classify the features [14].

When the improved convolution neural network model is used to classify images, it is very necessary to preprocess the image such as noise reduction and grayscale, select a certain number of training sets and test sets from the dataset, and then take the training set as the input object of the model after unsupervised learning processing. Secondly, the hidden layer of the noise reduction automatic encoder is used to encode and decode the input object, and the processing results are output to the sparse automatic encoder of the next layer for normalization. The data is trained layer by layer through the hidden layer of sparse automatic encoder, and finally the training results of sparse automatic encoder are output to Softmax classifier. In order to improve the classification accuracy, gradient descent method can be used to strengthen the training of classifier model parameters in order to improve the performance of image classification depth learning model. Finally, the network model is verified by using the image test set, and the effectiveness of the image classification method is tested according to the classification results output by the model.

The improved convolution neural network model can overcome the problem that the traditional neural network is only limited to some features in image classification. Through the normalization of sparse automatic encoder, the overfitting phenomenon of model in data processing can be avoided, and more abstract representative features can be obtained by using the hidden layer of sparse automatic encoder to train the data layer by layer. The improved model adopts Softmax classifier, which can make the classification result closer to the real value. The improved deep learning network model is mainly divided into two stages: training and testing. The training stage is mainly used to build an effective image classification model, and the testing stage is mainly to evaluate and analyze the model according to the experimental classification results. Figure 9 shows the workflow of the improved deep learning network model.

4.2. Image Classification Model Optimization. It is known from the existing research that the convolution neural network model can be optimized from the aspects of data enhancement and adjusting training methods, and the optimization of convolution neural network model is related to the type of model used. For example, for the deep learning model with more layers, the training parameters can be optimized.

According to the structure of convolution neural network, convolution layer is used to extract features, and the size of convolution kernel determines the extraction quality of image features. From the perspective of image composition, adjacent pixels can generally form the edge lines of the image. Several edge lines form the image texture, and the image texture is combined into several local patterns. The local pattern is the basic element of the image. Through the convolution layer of the network model, different types of features can be extracted and the local pattern of the image can be formed. When the convolution kernel is smaller,

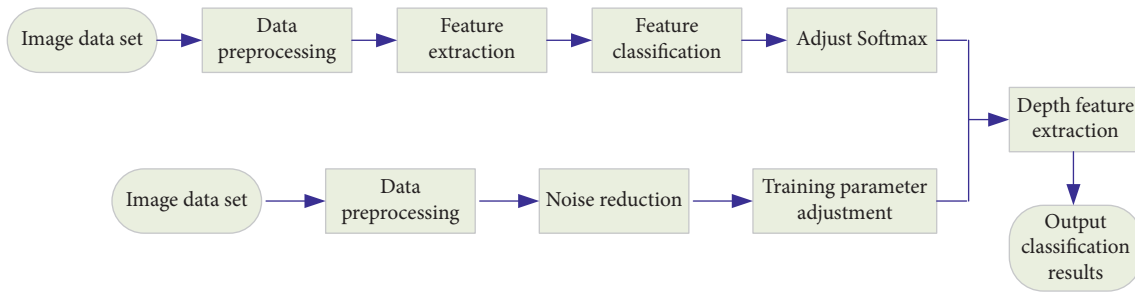


FIGURE 8: Working diagram of improved image classification algorithm.

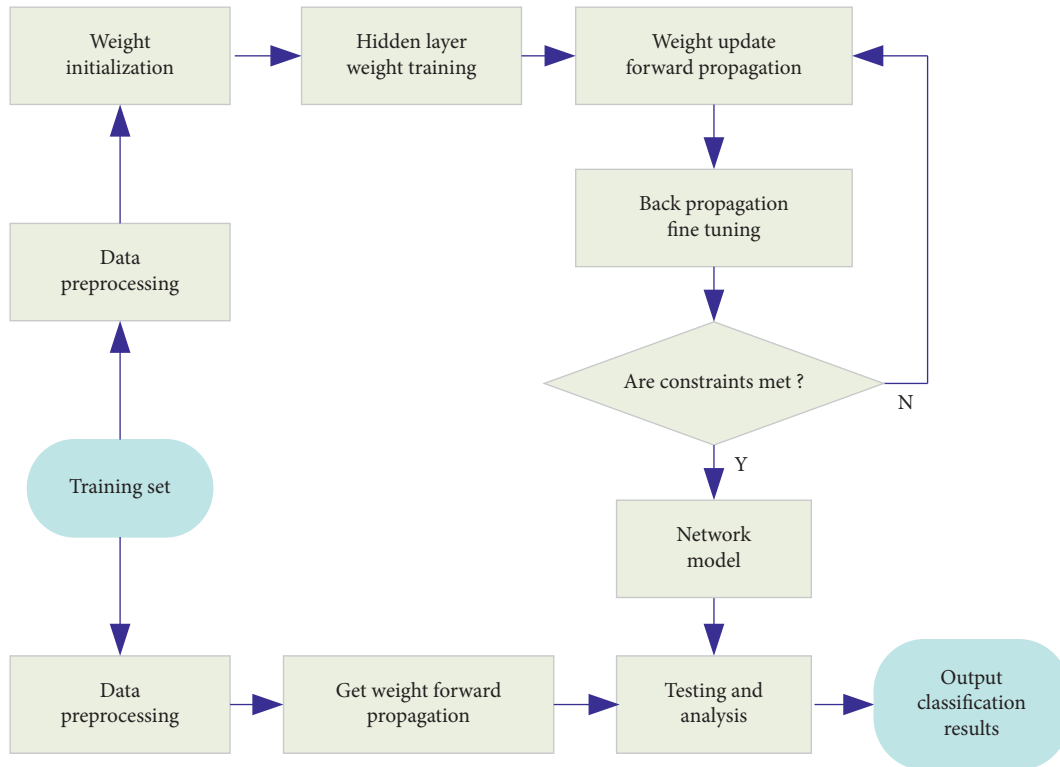


FIGURE 9: Working flow diagram of improved deep learning network model.

although the convolution layer extracts more features, there may be overfitting problems in the model. On the contrary, the larger the convolution kernel is, the fewer features extracted by the convolution layer are, and the worse the image classification effect may be. Therefore, the reasonable optimization of the size of convolution kernel can improve the accuracy of image classification.

Because the convolution neural network model mainly extracts the image features layer by layer through different convolution layers, the number of convolution layers will affect the feature extraction quality of the model to a certain extent. Similar to the number of convolution kernels contained in the convolution layer, the more convolution layers, the finer the features obtained by the model classifier, which may lead to overfitting phenomenon, while the less convolution layers, the coarser the features obtained by the model classifier, which may lead to the decline of image classification accuracy. Therefore, the optimization of

convolution layers can improve the classification accuracy of the model.

In order to improve the accuracy of image classification and recognition, the depth learning model proposed in this paper needs to be optimized. Firstly, a smaller convolution kernel is selected in the first convolution layer in order to extract more image feature details. Secondly, the maximum pool sampling operation is adopted in the model to avoid the overfitting problem. As shown in Figure 10, the optimized image classification model consists of three convolution layers, in which the convolution kernel of each convolution layer decreases in turn. At the same time, after each convolution layer, the features are processed by ReLU activation function, and the generated features are used as the input of the maximum pooling layer. The model adopts three full connection layers, takes the processing result of the last full connection layer as the input of Softmax classifier, and then generates the classification result of the image.

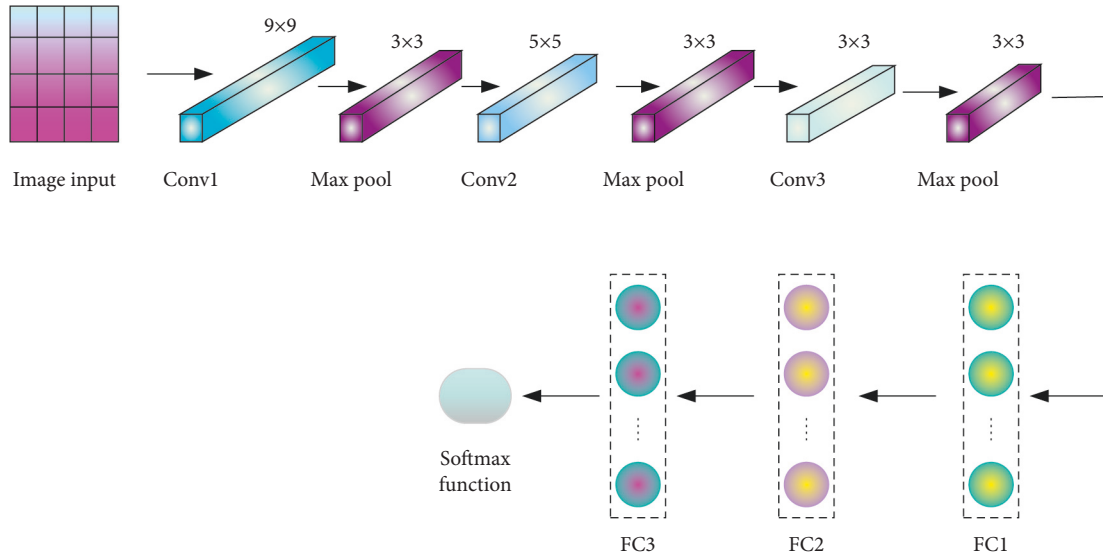


FIGURE 10: Structure diagram of optimized image classification model.

The convolution layer parameters of the model include the size and number of convolution kernels. The first convolution layer is close to the image input layer and is mainly used to extract the basic features of the image, so the parameters of the first convolution layer have a great influence on the features. In order to facilitate the further processing of features in the subsequent convolution layer, a smaller convolution kernel needs to be used to extract the attribute information such as shadow, boundary, and light of the image.

The convolution layer maps the obtained features through the activation function. Therefore, the optimized convolution neural network model adopts ReLU activation function [20], which can be expressed by mathematical function as follows:

$$y(x) = \text{Max}(0, x). \quad (2)$$

When the traditional convolution neural network model uses ReLU activation function to train features, it may lose useful feature information in the process of image classification. In order to prevent the loss of useful features during image classification, it can be improved on the basis of the existing ReLU activation function [20]. The optimized activation function can be expressed as

$$y(x_i) = \begin{cases} \frac{x_i}{c_i}, & x_i < 0, \\ x_i, & x_i \geq 0. \end{cases} \quad (3)$$

From the improved calculation formula of activation function, when the input feature is less than zero, it can not only retain the negative value information in the feature map but also increase the reinforcement learning of effective features.

The optimized convolution neural model uses Softmax function to classify the images, and Softmax function uses supervised learning algorithm to regress the features [21]. In

the classification process, category y of the image target can have M different values. If the image training set is $\{(x_1, y_1), \dots, (x_i, y_i)\}$, where x_i represents the image training sample, y_i is the image classification category, and $y_i \in \{1, 2, \dots, M\}$, the cost function of Softmax regression algorithm can be expressed as

$$r(\alpha) = \frac{1}{N} \left[\sum_{i=1}^N \sum_{k=1}^M 1\{y_i = k\} \log \left(\frac{\exp(\alpha_k^T x_i)}{\sum_{j=1}^M \exp(\alpha_j^T x_i)} \right) \right]. \quad (4)$$

Assuming that M markers are accumulated in the cost function, the probability calculation of training sample x as category k can be obtained, which is expressed as

$$\lambda(y_i = k | x_i; \alpha) = \frac{\exp(\alpha_k^T x_i)}{\sum_{j=1}^M \exp(\alpha_j^T x_i)}. \quad (5)$$

5. Experiment and Analysis

5.1. Selection of Datasets and Experimental Methods. In order to verify the effectiveness of the image classification depth learning model proposed in this paper, the Flower dataset provided by Oxford University was used in the experiment [22]. These images describe the proportion, shape, and light changes of different types of flowers, and the images of some flower categories change greatly. The dataset contains 17 categories of flower datasets; each category contains 80 pictures, for a total of 1360 pictures. In the experiment, the dataset is randomly divided into three parts as the training set, verification set, and test set of the model. A partial image of the Flower dataset is shown in Figure 11.

The experiment is based on Matlab, Python, and deep learning framework. Through the preprocessing, feature extraction, training, and classification of Flower dataset, the effectiveness and accuracy of the model in image classification are verified.

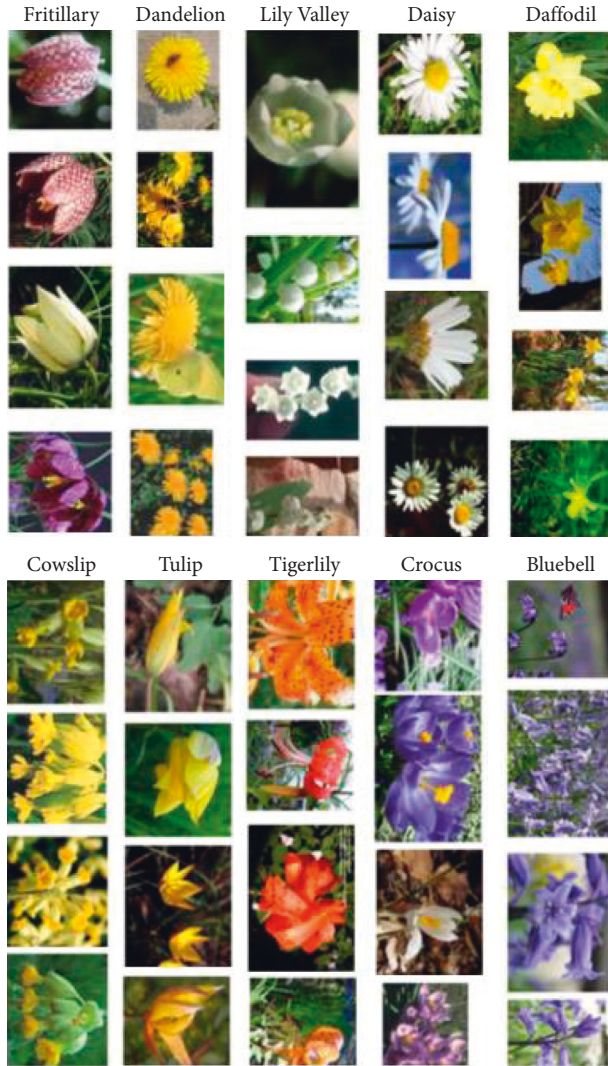


FIGURE 11: partial image of the Flower dataset [22].

In the experiment, the classification accuracy is used to evaluate the robustness of the deep learning network model. According to different evaluation angles, classification accuracy generally includes overall accuracy and category classification accuracy [23]. The overall accuracy AC is expressed by the ratio of the number of correctly classified samples to the total number of samples, while the classification accuracy CCA_i is expressed by the ratio of the number of correctly classified samples to the number of such tests. The calculation formula is as follows:

$$AC = \frac{t_r}{t_n}, \quad (6)$$

$$CCA_i = \frac{t_{ri}}{t_{ni}}$$

In the above formula, t_{ni} represents the number of correctly classified samples, t_n denotes the number of test samples, t_{ri} indicates the number of correctly classified samples of type i , and t_{ni} expresses the number of test samples of type i .

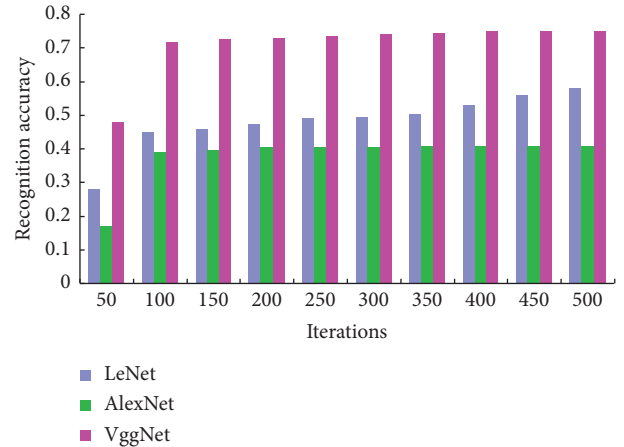


FIGURE 12: The relationship between the classification accuracy of the model and the number of iterations.

5.2. Results and Analysis. In order to facilitate comparative analysis, three common network models, LeNet, AlexNet, and VggNet, are selected for comparison in the experiment. According to the experimental comparison results, the relationship between the classification accuracy and the number of iterations of the model when using different network models to classify flower images is shown in Figure 12. When iterating 50 times, the accuracy of LeNet model is 28%, that of AlexNet model is 17%, and that of VggNet model is 48%. When the number of iterations is 100, the accuracy of LeNet model is 45%, that of AlexNet model is 39%, and that of VggNet model is 72%. In addition, when the VggNet model converges, the highest accuracy rate is as high as 75%, the highest accuracy rate of LeNet model is 58%, and the highest accuracy rate of AlexNet model is 41%.

In order to test the effect of image classification after the model is optimized to a certain extent, the accuracy of flower image classification before and after the model optimization is compared in the experiment, as shown in Figure 13. The comparison results show that, for the training dataset, the optimized model converges faster in the early stage of training and slower in the middle stage of training, while the two models in the later stage of training are basically the same. For the test dataset, the optimized model is higher than the nonoptimized model in terms of convergence speed and image classification accuracy. Therefore, the model optimization method proposed in this paper can effectively improve the accuracy of image classification.

In addition, in order to test the relationship between the loss value function of the optimized model and the number of iterations, the training set and test set are used to compare the models before and after optimization, as shown in Figure 14. The loss value function of the nonoptimized model shows an upward trend with the increase of the number of iterations, indicating that the nonoptimized model has the phenomenon of overfitting, while the loss value function of the optimized model shows a downward trend with the increase of the number of iterations. It can be seen that the cost of parameter training can be reduced through model optimization.

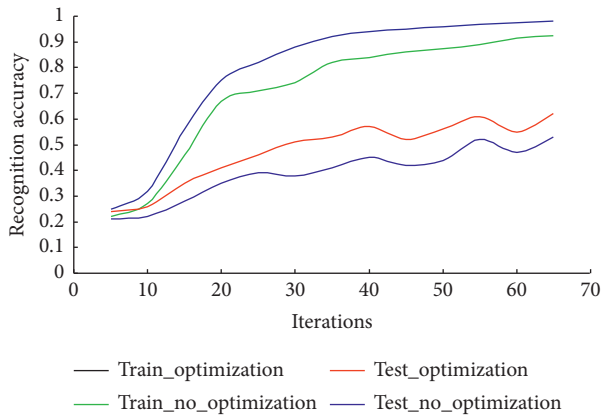


FIGURE 13: Comparison of accuracy before and after model optimization.

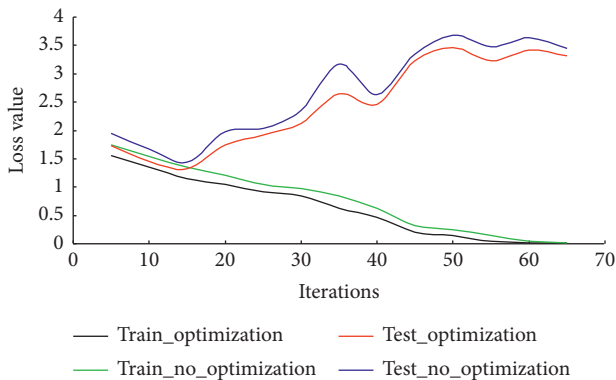


FIGURE 14: Iterative comparison of model loss value before and after optimization.

From the above experimental comparison results of the relationship between the accuracy of common network models in image classification and the number of iterations, it is known that the model proposed in this paper is superior to other models in classification accuracy. By comparing the classification accuracy of the deep learning model on the training set and the test set before and after optimization, it is known that the accuracy of image classification can be significantly improved after a certain degree of optimization.

6. Conclusion

Aiming at the problems of large time overhead and low classification accuracy in traditional image classification methods, a deep learning model of image classification based on machine learning was proposed in this paper. By analyzing the basic theory of neural network, this paper expounded the types of convolution neural network and its application in image classification. Using the existing convolution neural network model, through noise reduction and parameter adjustment in the feature extraction process, an image classification depth learning model based on improved convolution neural network was proposed. In order to improve the classification efficiency and accuracy of

the model, this paper optimized the proposed deep learning model. Finally, the accuracy of several common network models in image classification was compared through experiments. The results showed that the proposed model was better than other models in classification accuracy. At the same time, the classification accuracy before and after the optimization of the deep learning model was compared and analyzed. The results showed that the optimized model had a great improvement in the accuracy of image classification. How to classify dynamic targets in complex environment is a problem that needs further research in the future.

Data Availability

The labeled dataset used to support the findings of this study is available from the corresponding author upon request.

Conflicts of Interest

The authors declare that there are no conflicts of interest.

Acknowledgments

This work was supported by the Shijiazhuang Posts and Telecommunications Technical College.

References

- [1] S. H. Kim and H. L. Choi, "Convolutional neural network-based multi-target detection and recognition method for unmanned airborne surveillance systems," *INTERNATIONAL JOURNAL OF AERONAUTICAL AND SPACE SCIENCES*, vol. 20, no. 4, pp. 1038–1046, 2019.
- [2] P. W. Song, H. Y. Si, H. Zhou, R. Yuan, E. Q. Chen, and Z. D. Zhang, "Feature extraction and target recognition of moving image sequences," *IEEE Access*, vol. 8, pp. 147148–147161, 2020.
- [3] W. Y. Zhang, X. H. Fu, and W. Li, "The intelligent vehicle target recognition algorithm based on target infrared features combined with lidar," *Computer Communications*, vol. 155, pp. 158–165, 2020.
- [4] M. Li, H. P. Bi, Z. C. Liu et al., "Research on target recognition system for camouflage target based on dual modulation," *Spectroscopy and Spectral Analysis*, vol. 37, no. 4, pp. 1174–1178, 2017.
- [5] S. J. 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, 2020.
- [6] O. Kechagias-Stamatis and N. Aouf, "Evaluating 3D local descriptors for future LIDAR missiles with automatic target recognition capabilities," *The Imaging Science Journal*, vol. 65, no. 7, pp. 428–437, 2017.
- [7] M. Ding, Z. J. Sun, L. Wei, Y. F. Cao, and Y. H. Yao, "Infrared target detection and recognition method in airborne photoelectric system," *Journal of Aerospace Information Systems*, vol. 16, no. 3, pp. 94–106, 2019.
- [8] W. L. Xue and T. Jiang, "An adaptive algorithm for target recognition using Gaussian mixture models," *Measurement*, vol. 124, pp. 233–240, 2018.
- [9] F. Liu, T. S. Shen, S. J. Guo, and J. Zhang, "Multi-spectral ship target recognition based on feature level fusion," *Spectroscopy and Spectral Analysis*, vol. 37, no. 6, pp. 1934–1940, 2017.

- [10] S. Razakarivony and F. Jurie, "Vehicle detection in aerial imagery: a small target detection benchmark," *Journal of Visual Communication and Image Representation*, vol. 34, pp. 187–203, 2016.
- [11] O. K. Stamatidis 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.
- [12] L. Y. Ma, X. W. Liu, Y. Zhang, and S. L. Jia, "Visual target detection for energy consumption optimization of unmanned surface vehicle," *Energy Reports*, vol. 8, pp. 363–369, 2022.
- [13] Z. Geng, H. Deng, and B. Himed, "Ground moving target detection using beam-Doppler image feature recognition," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 54, no. 5, pp. 2329–2341, 2018.
- [14] Z. M. Guo, Y. Jiang, and S. H. Bi, "Detection probability for moving ground target of normal distribution using an imaging satellite," *Chinese Journal of Electronics*, vol. 27, no. 6, pp. 1309–1315, 2018.
- [15] Y. K. Bai, "Target detection method of underwater moving image based on optical flow characteristics," *Journal of Coastal Research*, vol. 93, no. sp1, p. 668, 2019.
- [16] W. Z. Wu, J. W. Zou, J. Chen, S. Y. Xu, and Z. P. Chen, "False-target recognition against interrupted-sampling repeater jamming based on integration decomposition," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 57, no. 5, pp. 2979–2991, 2021.
- [17] L. L. Yu, Q. X. Yang, and L. M. Dong, "Aircraft target detection using multimodal satellite-based data," *Signal Processing*, vol. 155, pp. 358–367, 2019.
- [18] I. Mahmud and Y. Z. Cho, "Detection avoidance and priority-aware target tracking for UAV group reconnaissance operations," *Journal of Intelligent and Robotic Systems*, vol. 92, no. 2, pp. 381–392, 2018.
- [19] T. Yulin, S. H. Jin, G. Bian, and Y. H. Zhang, "Shipwreck target recognition in side-scan sonar images by improved YOLOv3 model based on transfer learning," *IEEE Access*, vol. 8, pp. 173450–173460, 2020.
- [20] Z. M. Guo, Y. Jiang, and S. H. Bi, "Detection probability for moving ground target of normal distribution using infrared satellite," *Optik*, vol. 181, pp. 63–70, 2019.
- [21] S. Matteoli, M. Diani, and G. Corsini, "Automatic target recognition within anomalous regions of interest in hyperspectral images," *Ieee Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 11, no. 4, pp. 1056–1069, 2018.
- [22] M. E. Nilsback and A. Zisserman, "Automated flower classification over a large number of classes," in *Proceedings of the Sixth Indian Conference on Computer Vision, Graphics & Image Processing*, pp. 722–729, Bhubaneswar, India, December 2008.
- [23] C. Y. Zhang, B. L. Guo, N. N. Liao et al., "A public dataset for space common target detection," *KSII TRANSACTIONS ON INTERNET AND INFORMATION SYSTEMS*, vol. 16, no. 2, pp. 365–380, 2022.