Research on Flower Image Classification Algorithm Based on Convolutional Neural Network

Zhao Jiantao¹, Chu Shumin¹

¹School of Control and Computer Engineering, North China Electric Power University, Beijing 102206, China
1354457072@qq.com

Abstract
The fast progress of deep learning makes Convolutional Neural Network (CNN) emerges at the historic moment, and as an important achievement, it has been extensively used in all sorts of fields. Compared with traditional machine learning, CNN has more advantages, on the one hand, it has more hidden layers and complex network structure, and on the other hand, it has a stronger ability of feature learning and feature expression. With the fast progress of computer technology, the application research of fast and accurate recognition and classification of flowers by obtaining flower images through mobile devices has received extensive attention. The flowers images collected under natural conditions have large background interference, and it is difficult to recognize flowers because of their inter-class similarity and intra-class diversity. Therefore, in view of the lack of flower image data and low classification accuracy, the experiment sorted out the data sets of four kinds of flowers, and used the CNN to classify the images. Compared with the traditional approaches, the classification precision can be largely enhanced.

1. Introduction
Deep learning expands the scope of artificial intelligence and has made great achievements in recent years, among which the deep learning model of convolutional neural network can realize the extraction of relevant features through convolution and achieve great performance improvement in large-scale image classification tasks. The traditional image classification methods mainly extract the shape features, color features and texture features of the image. Using the manual feature extraction method, there will be difficulties in feature selection and insufficient feature extraction. Convolutional neural network has obvious advantages in the processing of input multidimensional signals, and it can be said that it is the first learning algorithm model of multilayer neural network successfully established and trained [1]. This paper presents a convolutional neural network model based on different kinds of flower image screening, classification and recognition method. First, four kinds of flower data sets are established, and then the convolutional neural network is built. The convolutional neural network is used to initially extract small features, then select and extract main features, and then summarize the features of each part. Finally, the image classification is realized, and a GUI interface is implemented for using the trained network. In addition, four groups of comparative tests were set from two aspects of learning rate and data enhancement, so as to realize rapid and accurate recognition of different flower images.
2. Convolutional Neural network
The birth of convolutional neural network is closely related to biological natural visual cognition, which is generally composed of input layer, convolutional layer, pooling layer, full connection layer and output layer, among which the convolutional layer and pooling layer are usually carried out in turn. Due to image after convolution layer calculation diagram output characteristics, the characteristics of the output figure with the input of each neuron are local connection, and through the weighted average of the corresponding connection weights, and then sum, increase after the offset feature map is obtained by nonlinear activation function, then the feature map using filter convolution [2], was also named the convolutional neural network. The convolutional neural network model used in the experiment is composed of one input layer, two convolutional layers, two pooling layers, two full connection layers and one Softmax layer for classification. ReLU activation function is used to add some nonlinear factors to the neural network, so that the neural network can better solve complex problems [3].

The first convolutional layer of the model used in the experiment consisted of 64 3x3 convolutional cores (3 channels), padding='SAME', representing the post padding convolution matching the original size and uses the activation function relu (), followed by a 3x3 maximum pooling layer with a step strides of 2. Behind this is still a convolutional layer of 16 3x3 convolutional cores (16 channels) and 3x3 maximum pooling layer with 2 length of step strides and after the pooling layer are two fully connected layers, both of 128 neurons.

2.1. Convolutional layer
The convolution layer is composed of several convolution units, and the parameters of each convolution unit, such as the size of the convolution kernel, are optimized by the back propagation algorithm. The convolution operation is mainly used to extract the features of the input image. The primary convolution layer can only extract some small image features, such as lines, shapes, edges, etc., while more multi-layer convolution layers can extract more complex image features from these small features. Just as we human beings recognize pictures, the brain cannot recognize a picture instantaneously, but first locally recognize some features in the picture, and then comprehensively operate the local features of the image from a higher level, so as to get the global information of the whole picture [4].

2.2. Pooling layer
Pooling is also called subsampling, and its main function is to reduce the size of the feature image while maintaining the features of the image. In general, there is a pooling layer between the two convolution layers. The pooling layer does not contain parameters, and it only makes a subsampling of the characteristic map transmitted by the convolution layer, that is, data compression, to prevent overfitting. The pooling layer can play a role because the features of the image will not be changed, that is, the original features of the image will not be missing through the subsampling. According to this feature, the image can be shrunk down for convolution calculation, which can greatly reduce the time of convolution operation [5]. There are usually two types of pooling: maximum pooling and average pooling. Maximum pooling is to select the maximum value within the defined neighborhood, while average pooling is to take the mean value within the neighborhood. Usually, maximum pooling is used more often, and the model in this paper also uses the maximum pooling method.

2.3. Activation function
When the activation function works, some neurons are stimulated to continue the activation information into the later layer. The neural network can solve the nonlinear problem because of the lack of linear model expression force and the activation function adds the nonlinear factors, supplementing the expression force of the linear model, so the characteristics of the activated neurons are saved and mapped to the next layer through the activation function. There are many kinds of activation functions, where ReLU functions are more used, and the formula of ReLU functions is f(x)
3

=\max (x, 0). The ReLU function is all equal to 0 at \( x < 0 \), and the \( f (x) \) derivative is 1 at \( x > 0 \), so the ReLU function can make the gradient unattenuated at \( x > 0 \), which can not only prevent the gradient from disappearing, but also converge faster, compensating for the sparse expression ability of the neural network[6]. The activation function used in the convolutional neural network model is the ReLU function.

2.4. **Full connection layer**

The full connection layer is usually at the end of the convolutional neural network. The nodes in the full connection layer are connected to all the nodes in the upper layer. The main function of the full connection layer is to integrate the features extracted from the front. After convolution, activation function and pooling, enough features are extracted to identify the image, then the image can be classified. Normally be enough to help identify the characteristics of the image by the neural network into a cuboid, at the end of the convolutional neural network will get into a vertical rectangle vector, the input to classify all connection to cooperate in the output layer, is about to get 2 d feature graph into a one dimensional vector. The layers preceding the full connection layer are mainly used to extract different features, while the full connection layer integrates local information after convolution, activation functions and pooling.

2.5. **softmax classifier**

Softmax classifier is the output layer of convolutional neural network. After receiving the real number vector of the full connection layer by Softmax, the output is also a real number vector. Each value of the real number vector output by Softmax represents the probability that this sample belongs to each class. Each value of the output real vector has a size between 0 and 1.

3. **Dataset**

The dataset contains four common types of flowers: tulips, dandelions, sunflowers, and roses. Each category had the same size, ranging from more than 600 to more than 800 images. Two hundred pieces from each category were selected as the test set, and the rest were used as the training set.

Due to the small amount of data in the data set, the model soon overfitted in the training, so data enhancement was used in the subsequent experiments. Data enhancement means that after flipping, rotating and changing the color saturation of the image, more images will be obtained for training. Data enhancement will only be used for the training set. After data enhancement, the number of each class was changed to about 2000 pieces, and some were deleted randomly, because all data enhancement would also overfit. Figure 1 shows the data set before the Sunflower class data enhancement, and Figure 2 shows the data set after the Sunflower data enhancement.

![Figure 1. Sunflower Dataset (Before Data enhancement)](image-url)
4. Experimental results and analysis

4.1. Experimental environment
The training environment is a Windows8 system, and the computer is configured as the Intel (R) Core (TM) i5-5200 UCPU. Training Convolutional neural networks employs the TensorFlow deep learning framework that, as one of the most used frameworks developed by Google, is not only easy to use and efficient, but also a variety of devices supported by TensorFlow, whether in computer clusters and smartphones. And you can both generate the training model immediately and reuse it.

4.2. Experimental setting
The initial learning rate is set to 0.0001, max step to 5000, print at 50 steps and accuracy, until the loss and accuracy values remain stable. loss and, accuracy represent the loss rate and accuracy of the training set, respectively. Batch Size, initially set to 20, is the number of samples selected for a workset.

4.3. Experimental Results
(1) The first set of experiments: with the learning rate of 0.0001, batchsize of 20 and step of 1700, loss and accuracy approached 0 and 100% respectively. As shown in Figure 3.

After the test set tested multiple sets, the results is not only not high, but also mostly wrong. Figure 4 is a GUI interface, and Figure 5 is the result of the test. The GUI interface shows the picture to test is dandelion, and the final identified result is sunflower.
The second set of experiments: when the learning rate is 0.001, batchsize 20, other parameters unchanged, when step is about 1050, loss and accuracy approached 0 and 100% respectively. As shown in Figure 6.

Multiple sets were tested by the test set, and most of the judgments were correct but with low accuracy. As shown in Figure 7, the dandelion in experiment 1 is still tested, and the final identification result is that the probability of the dandelion is only more than 20 percent.
(3) The third set of experiments: when the learning rate is 0.0001 and batchsize is 60, the data enhancement of the dataset is performed, and the other parameters remain unchanged. When step is around 2200, loss and accuracy approach 0 and 100% respectively. As shown in Figure 8.

![Figure 8. Accuracy and loss rate of the Group 3 experiment](image)

Multiple sets were tested through the test set, the results were correct and the accuracy was significantly improved. As shown in Figure 9, the original picture is tulips, and the final recognition accuracy is more than 80 percent.

![Figure 9. Test Results](image)

(4) The fourth set of experiments: when the learning rate is 0.001 and batchsize is 60, the data enhancement of the dataset and the other parameters remained unchanged. When step is around 1600, loss and accuracy approach 0 and 100% respectively. As shown in Figure 10.

![Figure 10. Accuracy and loss rate of the Group IV experiments](image)

Multiple sets were tested by the test set, the results were correct and the accuracy was higher than that of the third set. As shown in Figure 11, the same test picture case was used in the third group, with higher results than the third group.

![Figure 11. Test Results](image)
4.4. Experimental Analysis

From the comparison of one and two experiments and three experimental groups, we found that when the learning rate reached 0.001, the optimal solution reached 0.0001. Moreover, the results of the first three and two four groups showed that the data. Combining these four sets of experiments, the fourth set of experiments can both reach the optimal solution of the training set, namely, when the learning rate is 0.001, and perform the data enhancement of the dataset.

Learning rate is the parameter set before starting learning, is a very important hyperparameter, which constrains the learning progress of the network model, determines whether the network can succeed or how long it takes to find the global minimum, so as to obtain the global optimal solution, which is the optimal parameter. When the learning rate setting is too small, the network convergence is very slowly, increasing the time when the optimal value is found. While setting a very small learning rate is reachable, this is likely to only reach local extremal points, no really found optimal solution[7].

When the dataset has less data, all the characteristics of the dataset are fitted with too many parameters, but the commonality between the data is ignored, so the network model will quickly overfit. The network model largely fits the distribution of the training set data, and causes the low accuracy of the test set data and the lack of generalization ability. Using data enhancements can be used to prevent overfitting, but during training it is found that the accuracy and loss rate fluctuate substantially after data enhancement when the batchsize is invariant. Is because if the data enhancement is 1 picture into 3, batchsize is 3, when in the test, small data set of each batchsize of 3 pictures are different, this time three pictures of error probability is very low, but after the data enhancement, there is very likely that 3 pictures from the same original picture, this will cause right or wrong, big fluctuations[8]. So, when we do the data enhancement, we can also increase the same multiple of the batchsize.

5. The Conclusion

This paper addresses different kinds of flower image classification problems by training convolutional neural networks while comparing with the results after the learning rate differences and data enhancement. The experimental results show that the convolutional neural network can accurately extract the high-level semantic features of various flower images, and greatly improve the classification results, where when the learning rate is set too small, the real optimal solution cannot be found; too small will cause overfitting, and data enhancement can solve this problem[9]. After contrast, when the learning rate is 0.01 and enhanced test set accuracy for dataset data.

Acknowledgments

This work was supported by the Institute of Network and Information Systems of North China Electric Power University.

References

[1] Liu Shujun, Kou Xupeng, He Ying, Mo Xuefeng. Study on crop Pathological Image Classification Algorithm Based on Convolutional Neural Network [J]. Hubei Agricultural Science, 2021,60 (09): 131-134
[2] Lin Peijun, Geng Zengmin, Hong Ying, Li Xuefei. Application of Convolutional Neural Network in Textile and Clothing Imagery [J]. Journal of Beijing Institute of Clothing (Natural Sciences), 2021,41 (01): 92-99 + 108
[3] Hu Xiaochun, Zhu Chengyu, Chen Yan. Research and analysis of deep convolutional neural network model [J]. Information Technology and Information Technology,2021(04):107-110.
[4] Wang Bin, Gao Jiaping, Si Shhtao. Image Classification and Application Based on Convolutional Neural Network [J]. Electronics and Packaging, 2021,21 (05): 76-80
[5] Dong Liang, Duan Chint, Gong Bin. Image Classification Based on Convolutional Neural Networks and Migration Learning [J]. Information and Computer (Theory Edition), 2021,33
[6] Li Ying, Song Lijuan. Study on Remote Sensing Image Recognition and Classification Based on GoogLeNet Model [J]. Computer Knowledge and Technology, 2021, 17(12): 4-6

[7] Abade Andre, Ferreira Paulo Afonso, de Barros Vidal Flavio. Plant diseases recognition on images using convolutional neural networks: A systematic review[J]. Computers and Electronics in Agriculture, 2021, 185.

[8] XU Weijun. Application of Deep Learning in the Field of Computer Vision [J]. Electronic Technology, 2021, 50(05): 20-21.

[9] LI Yan. Research on garbage image recognition and classification based on ResNet algorithm [J]. Yangtze River Information and Communication, 201, 34(05): 25-27.