

# Research on Flower Image Classification Based on Transfer Learning

Zhiling Wang\*

School of Information Technology Engineering, Tianjin University of Technology and Education, Tianjin 300222, China

\* Corresponding author: Zhiling Wang (Email: wzl822@126.com)

**Abstract:** Due to the high similarity between flowers, it is difficult to identify them if they do not have the corresponding biological knowledge when classifying varieties manually. Given the above problems, to improve the accuracy and efficiency of flower classification, this paper proposes a migration parameter pre-training and fine-tuning VGG16 model based on the ImageNet data set to solve this problem. In this paper, the grid coverage enhancement method enhances the flower classification data set to expand the training sample data. The model uses the migration learning pre-training and fine-tuning method to improve network stability and accelerate network convergence. The results of comparative experiments show that the performance of the improved model has been significantly improved, and the result is better on the flower image data set, which has specific practical value.

**Keywords:** Deep learning, Transfer learning, Neural networks, Image classification.

## 1. Introduction

Flower classification and recognition are one of the research hotspots in the field of plant recognition. It is an essential issue in computer vision. It requires computers to recognize different kinds of flowers automatically. This is of great significance for botany research, agriculture, and ecological monitoring. With the development of deep learning technology, flower image classification has made significant progress. The researcher trained the model on a large data set by using a deep learning model such as CNN convolution neural network and obtained a high accuracy rate on the test set. However, due to the huge diversity of flower images and the complexity of environmental conditions, there are still many challenges, such as occlusion, light changes, and posture changes. Therefore, flower image classification is still an active and challenging research field.

In the aspect of flower image research, literature [1] proposed integrating SIFT features and HOG features of flower images, and using support vector machine SVM to classify and recognize flowers, with a classification accuracy of 76.3%; Literature [2] designed an 8-layer convolutional neural network and tested it on the flower data set. Due to the shallow depth of the model, the classification effect is not ideal. Literature [3] concatenated the features of AlexNet and VGG16 models, used the mRMR algorithm to select effective features, and finally used SVM classifier for classification. This method is slightly cumbersome and the classification accuracy is not high. Literature [4] fused the multi-level deep convolution features based on the VGG16 model, It has been tested in the flower data set, but the accuracy of classification is not ideal. Literature [5] uses the method of multi-level feature fusion and extraction of regions of interest to classify flowers, and the model is relatively complex.

This paper proposes a flower image classification method based on the VGG16 network, which uses a model-based transfer learning method and a pre-training and fine-tuning model. The experiment shows that the improved model of this method has significantly improved the effect of flower recognition..

## 2. Data Preprocessing

### 2.1. Data set

The flower data set used in this paper takes five common flowers as the research object: daisy, dandelion, rose, sunflower, and tulip. There are 3670 image data samples of five flowers, including 633 daisies, 651 roses, 699 sunflowers, 799 tulips, and 898 dandelions. The flower data set, see Figure 1.



Figure 1. Flower data set

### 2.2. Grid coverage augmentation method

BHide-and-Seek [6] is a data enhancement method, mainly used for image classification. Its basic idea is to randomly cover some areas on the image to enhance the model's generalization ability. However, because the Hide-and-Seek is a random coverage, there is a problem with completely covering the objects in the image when processing the data, resulting in the loss of important feature information.

Because there are few picture data in the flower data set, to prevent over-fitting and improve the performance of the network model, this paper uses the controllable and discontinuous grid coverage enhancement method to expand the data set based on the weak data enhancement such as random rotation, random rotation, random translation, center clipping, color enhancement, etc.

The implementation of the grid coverage enhancement method is to select a grid pattern and then cover the grid pattern on the unconnected area of the image pixel set according to the needs of the task and finally get a new image sample. The size, spacing, and location of the overlay grid can be adjusted by parameters to control the amount of information not covered in the new image data. The grid coverage method corresponds to a total of five parameters, namely, the abscissa axis and the ordinate axis that control the orientation of the grid, the width and height of the grid spacing, and the side length that control the size of the grid itself. The grid coverage parameters are shown in Figure 2.

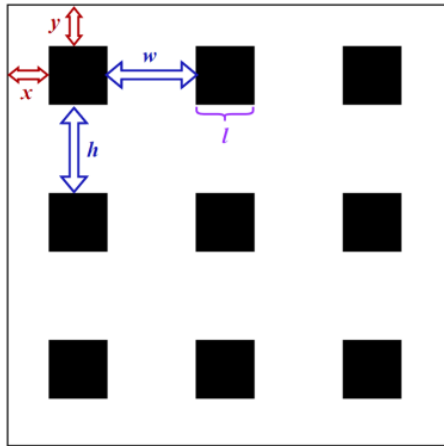


Figure 2. Grid covering method

The grid coverage method can make the model encounter image samples with different information without losing the key information of the image. In this study, the grid coverage abscissa and ordinate of image data are set to 21, the grid width and height spacing are set to 45, and the grid size is set to 30. Figure 3 is an example image after the grid coverage enhancement method. The left is the initial sample image, and the right is the image after the grid coverage.



Figure 3. Sample diagram of grid covering method

Through the above data enhancement operation, the flower classification data set was expanded from 3670 to 22020. In this paper, the expanded data set is divided into a training set and a test set according to the ratio of 8:2. The training set is used to train the model and the test set is used to evaluate the performance of the model.

### 3. Model Design

#### 3.1. VGGNet

VGGNet is a deep network model jointly developed and proposed by the Google team and Oxford University in 2014. It is the first to use the method of repeatedly overlapping convolution kernel to deepen the network depth in a disguised

form to improve the classification recognition rate [7]. The VGGNet network structure is composed of the convolution layer and the entire connection layer as a whole. Although the amount of parameters required for training is huge, it still has research value in the field of image classification today because of its strong expansibility, good generalization ability, and simple structure.

Nowadays, VGG16 is often used for research in classification tasks. There are 13 volume layers and 3 full connection layers. The number of convolution cores in the five modules of the VGG16 model is 64, 128, 256, 512 and 512 respectively, and the size of all convolution cores is  $3 \times 3$ . The step length is 1. In terms of the maximum pooling layer, there are 5 layers with a size of  $2 \times 2$ . Pool nucleus with a step of 2. In the end, there are three full connection layers, the first two full connection layers are 4096 output nodes, and the last full connection layer has a total of 1000 output nodes. The VGG16 network structure, see Figure 4.

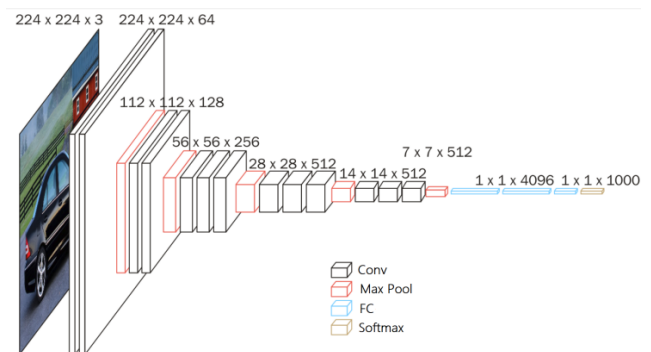


Figure 4. VGG16 network model

#### 3.2. Transfer Learning

In this paper, VGG16 and AlexNet [8] models are trained using the pre-training and fine-tuning method in transfer learning [9], and the AlexNet model is used as the reference for comparative experiments.

Select VGG16 and AlexNet trained by ImageNet large data set as the pre-training model. When the AlexNet model is loaded, the parameters of the input layer, pooling layer, and volume layer of the pre-training model are reserved and loaded.

Two schemes are adopted when loading the VGG16 model. Scenario A is loaded in the same way as the AlexNet model. In Scheme B, first, freeze all network layers except the full connection layer of the VGG16 model, then "unfreeze" the last three layers of the convolution layer and load the pre-training parameters, and finally select the ReLU activation function as the activation function of the first two layers of the full connection layer, and fine-tune the last three layers of the convolution layer and the full connection layer to randomly initialize and retrain them.

It should be noted that before the comparative experiment, the output quantity of AlexNet and VGG16 models is fine-tuned to 5, corresponding to the five flower types in the flower data set in this paper. The learning process of flower image classification and transfer learning, see Figure 5.

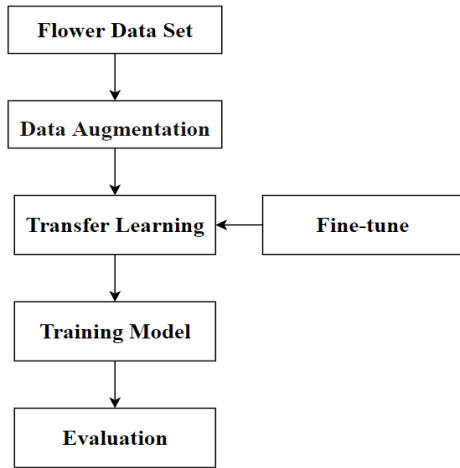


Figure 5. Flower image classification transfer learning

## 4. Experimental Design

### 4.1. Training Strategy

Adam (Adaptive Moment Estimation) is a variant of the gradient descent algorithm SGD, which can adaptively adjust the learning rate and improve the optimization efficiency [10]. Adam's algorithm introduces momentum, which can help the optimization algorithm converge faster, and can better find the optimal global solution near the optimal local solution. At the same time, the Adam algorithm also introduces the second stage gradient moment estimation, which can adjust the learning rate adaptively, so that the model can automatically adjust the learning rate in different training stages.

RAAdam (Rectified Adam) is an improved version of the Adam optimizer [11]. RAAdam algorithm introduces a new learning rate adjustment mechanism so that the learning rate can be adjusted according to the gradient of each parameter. In this way, parameters with larger gradients can be handled better and can converge faster.

To obtain a better effect of flower image classification, Adam and RAAdam are used to train the model in this comparative experiment.

### 4.2. Experimental Environment

This experiment was conducted on the 64-bit Windows 10 Professional operating system, using the in-depth learning framework of Python version 1.7. CUDA version is cuda11.0, cuDNN version is 8.1; The development platform is PyCharm Community 2021.12.10, and the programming language used is Python 3.7. The hardware configuration of the device is CPU: Inter (R) Core (TM) i5-11600KF, GPU: NVIDIA GeForce RTX 3070TI, and the operating memory is 16G.

### 4.3. Experimental Parameter Setting

This paper sets the number of training rounds epoch as 50, and the sample size of each batch is 32. A warmup learning rate optimization strategy [12] is adopted. The initial learning rate is 0.00001. Finally, the model is trained, and the results are observed and recorded.

## 5. Experimental Analysis

In this paper, the AlexNet model, the AlexNet model after transfer learning, the VGG16 model after transfer learning, and the VGG16 model after transfer learning are trained on the network based on the flower image data set using the Adam and RAAdam optimizers respectively. The experimental

results show that the accuracy of the flower image classification model is better than that of the Adam optimizer when using the RAAdam optimizer. At the same time, comparing the classification effect of the model flower image before and after the transfer learning, we can also find that the transfer learning method can effectively improve the accuracy of the model and accelerate the convergence speed of the network. In the case of using the Adam optimizer, the accuracy of AlexNet and VGG16 models trained by the transfer learning method has reached more than 80%, and the transfer learning method (Scheme B) proposed in this paper has achieved the highest accuracy in the experiment, reaching 89.1%. This shows that the migration learning method of thawing part of the convolution layer not only enhances the effect of network feature extraction, and improves the performance of the VGG16 model, but also achieves a good recognition effect on small-scale image data sets. The experimental results, see Table 1.

Table 1. Experimental Result

Method	Epoch	Optimizer	Acc
AlexNet	50	Adam	67.1%
VGG16	50	Adam	65.0%
Per-AlexNet	50	Adam	79.9%
Per-VGG16(A)	50	Adam	87.1%
Per-VGG16(B)	50	Adam	87.7%
AlexNet	50	RAAdam	68.5%
VGG16	50	RAAdam	65.8%
Per-AlexNet	50	RAAdam	81.6%
Per-VGG16(A)	50	RAAdam	88.3%
Per-VGG16(B)	50	RAAdam	89.1%

## 6. Conclusion

This paper proposes a flower image recognition and classification method based on the transfer learning method designs a pre-training and fine-tuning method suitable for flower image classification, constructs a VGG16 flower image classification model, and carries out comparative experiments to verify. The experimental results show that this method can not only avoid the shortcomings of the training model but also save a lot of training time and achieve high classification accuracy. This paper proposes a new pre-training and fine-tuning idea based on the VGG16 model for flower image classification. Whether this method can also achieve good results in other data sets needs further research. In this paper, only one model of transfer learning is proposed to identify five kinds of flowers. In the subsequent research, further optimization of the model is considered to improve the classification accuracy, and the method is tested on other types of data sets to improve the generalization ability of the model.

## Acknowledgment

This Project Supported by the Natural Science Foundation of Tianjin (No.18JCYBJC84900).

## References

- [1] Nilsback, ME, Zisserman, et al. Automated flower classification over a large number of classes[J]. -, 2008.
- [2] Liu Y, Tang F, Zhou D, et al. Flower classification via convolutional neural network[C]//2016 IEEE International

- Conference on Functional-Structural Plant Growth Modeling, Simulation, Visualization and Applications (FSPMA). IEEE, 2016: 110-116.
- [3] Cıbuk M, Budak U, Guo Y, et al. Efficient deep features selections and classification for flower species recognition[J]. Measurement, 2019, 137: 7-13.
- [4] Yin H, Fu X, Zeng J X, et al. Flower image classification with selective convolutional descriptor aggregation [J][J]. Journal of Image and Graphics, 2019, 24(05): 0762-0772.
- [5] YANG Wanggong, HUAI Yongjian. Flower fine-grained image classification based on multilayered feature fusion and region of interest. Journal of Harbin Engineering University, 2021, 42(4): 588-594. DOI: 10.11990/jheu.201912064.
- [6] Singh K K, Yu H, Sarmasi A, et al. Hide-and-seek: A data augmentation technique for weakly-supervised localization and beyond[J]. arXiv preprint arXiv:1811.02545, 2018.
- [7] Simonyan K , Zisserman A . Very Deep Convolutional Networks for Large-Scale Image Recognition[J]. Computer Science, 2014.
- [8] Yu W, Yang K, Bai Y, et al. Visualizing and comparing AlexNet and VGG using deconvolutional layers [C]//Proceedings of the 33 rd International Conference on Machine Learning. 2016.
- [9] Zhuang F, Qi Z, Duan K, et al. A comprehensive survey on transfer learning[J]. Proceedings of the IEEE, 2020, 109(1): 43-76.
- [10] Kingma D P, Ba J. Adam: A method for stochastic optimization[J]. arXiv preprint arXiv:1412.6980, 2014.
- [11] Liu L, Jiang H, He P, et al. On the variance of the adaptive learning rate and beyond[J]. arXiv preprint arXiv:1908.03265, 2019.
- [12] He K, Zhang X, Ren S, et al. Deep residual learning for image recognition[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2016: 770-778.