A Flower Auto-Recognition System Based on Deep Learning

Lin Shi¹, Zhigang Li² and Dingli Song³

¹ Computer Science Department, North China University of Science and Technology, China
² Computer Center, TangShan College, China
³ College of Science, North China University of Science and Technology, China
Email: pattiesl@126.com; lzghello@126.com; SongDL@126.com

Abstract. Building a flower auto recognition system with high accuracy is very significant. At the same time, it can bring convenience to our daily life. But Inter-class similarities between different species and the intra-class variation among the same species is a giant challenge needing addressed. Thus, this paper proposes a flower auto-recognition system based on deep learning, we get pictures by mobile smartphone and send the image to the CNN network, which is retrained by transfer learning based on Inception-V3. At last, an experiment was taken to verify the recognition accuracy is higher than other methods.

1. Introduction

Flower recognition is a very challenging problem, because there are thousands of flower species in the world. It’s a very demanding and time-consuming task, which has been mainly carried out by botanists. While in our daily life, people usually use mobile phones, cameras to shoot flowers, sometimes they are confused because of the flower species. Therefore, the design of an auto recognition system for flowers is necessary in Botany and also will bring so much fun to people’s lives.

The previous work about flower classification usually rely on some features, such as color, texture or shape. Color is the main feature to distinguish the foreground from the complex background. Takeshi Saitoh proposes an automatic algorithm for flower recognition, the experiment uses 600 pictures and ultimately handled a 90% recognition rate [1]. Maria-Elena Nilsback and Andrew Zisserman from the University of Oxford was studying the method of flower recognition from 2006 to 2010. They proposed a visual vocabulary for flower classification in 2006[2] and developed an automatic segmentation algorithm for flowers in 2007 [3]. In 2010, they improved the segmentation algorithm once more[4]. In 2004, Tiay used some color features and the K-nearest neighbor algorithm to classify [5]. Although numerous algorithms[6-8] were used for flowers recognition, it’s quite difficult to analyze because of the high inter-class similarities and intra-class variation. We can see from above that during this time, people put more energy to preprocess of the flower image and the feature selection to classify. The feature extraction needs people to participate, which will undoubtedly have a greater impact on the recognition rate.

This article proposes a faster way to retrain the CNN network based on Inception-V3 based on a smaller dataset to get a higher accuracy than some other feature extracting methods.

2. Related Work

2.1 DataSet

At present, there are tens of thousands of kinds of flowers category in the world. If we want to classify so many kinds of flowers, we must choose a standard flower dataset. Maria-Elena and Andrew
Zisserman, they came from the University of Oxford and they founded two standard datasets: Oxford-17 flower dataset and Oxford-102 flower dataset. Both datasets are widely used during the process of recognizing different kinds of flowers.

Oxford-17 [2] dataset includes 17 categories of flowers, each category includes 80 images with different posture and lighting. The flowers of Oxford-17 have inter-class similarities between different species and the intra-class variation among the same species.

The Oxford-102 [4] flower dataset contains 102 species of flowers and 8189 images in total. There are about 50-250 images for every category of flowers. It includes more species than Oxford-17 dataset, so the classification will be more complex.

After comprehensive consideration, we decided to choose 15 species that are not included in Oxford-17 from Oxford-102 to combine a new dataset, we call it FLOWERS32 dataset (32 classes of flower species). The FLOWERS32 has 2560 images totally and 80 images for each category. Then we use FLOWERS32 for the flower classification model training.

![Figure 1: The images of flowers on dataset FLOWERS32](image)

2.2 TensorFlow

TensorFlow is Google's second-generation artificial intelligence learning system based on DistBelief[16]. TensorFlow is an open source software library using data flow graphs for numerical computation. TensorFlow was originally developed by researchers and engineers from the Google Brain Group (affiliated to the Google Machine Intelligence Research Institute) for machine learning and deep neural network research, but its versatility allows it to be widely used in other computing fields as well. It can run on various improvements ranging from a smartphone to thousands of data center servers. TensorFlow is completely open source and can be used by anyone.

2.3 Inception-V3

Inception-V3[7] is one of the pretrained model released on TensorFlow’s official website in 2015. It is trained on ImageNet 2012 data set that contains 1.2 million images over 1000 classes. The highest quality version of Inception-v3 reaches 21.2%, top-1 and 5.6% top-5 error for single crop evaluation on the ILSVR 2012 classification, setting a new state of the art.

2.4 Transfer Learning

If a system has the ability of transfer learning, which means the ability of a system to recognize and apply knowledge and skills learned in previous tasks to novel tasks. The parameters of the trained model can be easily adjusted and trained, and can be transferred to other different data sets. At the same time, it does not need a lot of computational support, so it can be trained in a short time to achieve satisfactory results. This is transfer learning. Essentially, although the image data sets are different, the underlying features are mostly universal.

3. Flower-Recognition System

In the following part, we described the system’s main architecture, the construction process of the flower recognition system, which consists of four steps: data labelling, training process, verification process and testing process.

3.1 Main Architecture
The system’s architecture overview is as shown in Figure 2. Users get flowers pictures by their mobile phone (when we take a photo for the flower, we’d better make the flower as the only foreground).

![architecture overview of the system](image)

Figure 2 architecture overview of the system

The picture obtained by the user is converted to a Tfrecord format. Tfrecord is a standard format officially recommended by TensorFlow. It can store image data and labels together into binary files, and realize fast copy, move, read and store operations in TensorFlow. When training network, we can load the data in Tfrecord format into the queue beforehand by establishing the queue system. The queue will automatically realize the random or orderly entry and exit of the data, and the queue system and the model training are carried out independently, which speeds up the reading and training of our model. The flower image in Tfrecord format is sent to the pre-trained CNN system on FLOWERS32 dataset for identification and output the category label of the image. Then it sends the label to the cloud storage, where the mobile phone receives the flower name automatically. The whole process from capturing a flower picture to receiving the flower name on the smartphone takes about 0.8S tested with MI8, the mobile phone’s configuration is as follow: android system, 6GB of RAM and Snapdragon eight-core processor.

### 3.2 Data labelling
Labeling the pictures of the dataset before training is necessary, because learning process of CNN is supervised learning.

### 3.3 Feature Extracting
Because the data amount of the images, including width, height and channels is huge. And in the recognition process, there are a large number of iterations and epochs to train a neural network, so we choose to use Deep Learning in this situation.

Training a traditional CNN needs a huge training dataset and testing dataset, and before classifying, it needs feature extracting which usually needs manual intervention. In recent days, transfer learning is more and more popular in the process of machine learning. There is no need to re-collect and calibrate huge new data sets at great cost, and it is also possible that data can not be obtained at all. Transfer learning makes it possible to quickly migrate and apply.

Inception-v3 is the 3rd version of Inception. Inception network is an important milestone in the development of CNN classifier. Before Inception, most popular CNNs just stacked more and more convolution layers, making the network deeper and deeper, hoping to get better performance. The greatest feature of Google LeNet is the use of Inception module. Its purpose is to design a network with excellent local topology, that is, to perform multiple convolution or pooling operations on input images in parallel, and to stitch all output results into a very deep feature map. Its continuous evolution has led to the emergence of a variety of Inception network versions.
Inception-v3 is a pre-trained network on ImageNet. In the process of the transfer learning, keeping the bottleneck layer and its trained parameters unchanged, only changing the final full connection layer can greatly reduce the training data and shorten the training time of the model. We also change the output numbers to 32 outputs, which corresponds to the 32 categories of flowers in the FLOWERS32 dataset.

3.4 Software Requirements
To achieve the proposed system, we used two programming languages, Java and Python. JAVA is used to implement mobile programming on mobile phone, mainly for image data acquisition. Python is used to construct CNN network structure for classification.

4. Experiment
The software platform of this experiment is TensorFlow. The transfer learning model used is inception-v3. The hardware configuration of the experiment is as follows: Lenovo, processor 1.8GHz, quad-core, Intel i7, memory 8GB 2400MHz DDR4. The dataset for the experiment is FLOWERS32 which we achieved based on Oxford-17 and Oxford-102. In this part, we detailed statement of the experimental process and then, experiment results are analyzed. At last, we compare the results with other methods.

4.1 Experimental steps
For the FLOWERS32 dataset, we take 1920 images as training set and 640 images as testing set. Generally, we do migration training, which is to remove the top layer, followed by a variety of custom new layers. The output layer first converts the output of 8 * 8 * 2048 into a tensor of 1 * 2048 using the Global Average Pooling 2D function. After that, a 1024-node full-connection layer is added, and finally an output layer of 32 nodes is used to activate the function with softmax.

4.2 Results Analysis
The recognition rate and based on FLOWERS32 is shown in Fig.3. The cross entropy is shown in Fig.4. The training dataset is represented by the red curve and the testing dataset is represented by the blue curve. We can learn from the figures that for the training data, the accuracy can reach about 100%, for the test data, the accuracy can reach about 95%. The cross entropy of the training data and the test data is separately 0.01 and 0.07.

Figure.3 recognition rate on FLOWERS32 dataset
We also compared our system with some other methods. In order to verify the effectiveness of this method, we select some typical feature methods for comparison. For example, color histogram, HU moments, LBP and combinations of these features. All these features are used on the training set and cooperated with the classifier to classify. The comparison of recognition efficiency is shown in Figure.5. We can learn from the figure the recognition efficiency of our system is better than any other global feature descriptors.

**Figure.5** Typical features vs our system with CNN (FLOWERS32)

5. **Future Work**

In this paper, we propose a transfer learning method to retrain inception-v3, which improves the recognition rate and reduces the hardware requirement of the system terminal and the size requirement of the training set. We chose 32 categories of flowers based on Oxford-17 and Oxford-102 as the dataset. The classification accuracy can reach about 95%, which is higher than other methods. However, there are millions of flower categories in the world, so the system should be more universal so that it can recognize more kinds of flowers. In addition, the future work should focus on building an automatic recognition system, which can not only identify flowers, but also accurately identify the heel, stem and leaf of plants. This is of great significance to Chinese medicine. And it is very helpful to people's daily life. For example, once you have such a recognition system, if people go out to play and need first aid, you can identify the plants around you.

If we want to complete the system, we need a lot of training set test data. Therefore, the establishment of data sets is the focus. We can take pictures through the terminal or use a public database.

6. **Acknowledgments**

This paper is partly supported by Key projects of Hebei Provincial Education Department (ZD2015059).
7. References

[1] I. F. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci. 2002. Wireless Sensor Networks: A Survey. Comm. ACM 38, 4 (2002), 393–422.

[2] David A. Anisi. 2003. Optimal Motion Control of a Ground Vehicle. Master’s thesis. Royal Institute of Technology (KTH), Stockholm, Sweden.

[3] P. Bahl, R. Chancre, and J. Dungeon. 2004. SSCH: Slotted Seeded Channel Hopping for Capacity Improvement in IEEE 802.11 Ad-Hoc Wireless Networks. In Proceeding of the 10th International Conference on Mobile Computing and Networking (MobiCom’04). ACM, New York, NY, 112–117.

[4] Jacques Cohen (Ed.). 1996. Special Issue: Digital Libraries. Commun. ACM 39, 11 (Nov. 1996).

[5] Saitoh T, Aoki K, Kaneko T. 2004. Automatic recognition of blooming flowers. In IEEE International Conference on Pattern Recognition: 27-30.

[6] Maria-Elena Nilisback, Andrew Zisserman. 2010. Delving deeper into the whorl of flower segmentation. Image Vision Comput. 2010, 28(6): 1049-1062.

[7] Avishikta Lodh, Ranjan Parekh. "Flower recognition system based on color and GIST features", 2017 Devices for Integrated Circuit (DevIC), 2017

[8] H. Kulkarni, H. M. Rai, K. A. Jahagirdar and P. S. Upparamani, 2013. A Leaf Recognition Technique for Plant Classification Using RBPNN and Zernike Moments, International Journal of Advanced Research in Computer and Communication Engineering, Vol. 2, Issue 1, 984-988.

[9] T. Tiay, P. Benyaphaichit, P. Riyamongkol. 2014. Flower Recognition System Based on Image Processing. IEEE Third ICT International Student Project Conference, 99-102.

[10] Christian Szegedy, Vincent Vanhoucke, et al. 2015. Rethinking the Inception Architecture for Computer Vision. arXiv:1512-1516.

[11] Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich. 2015. Going Deeper with Convolutions, IEEE Conference on Computer Vision and Pattern Recognition (CVPR’15).

[12] Robert M. Haralick, K. Shanmugam, Its’hak Dinstein. 1973. Textural Features for Image Classification, IEEE Transactions on Systems, Man and Cybernetics, Vol.SMC-3, No. 6, November 1973, 610-621.

[13] Siraj, Fadzilah, Muhammad Ashraq Salahuddin, and Shahrul Azmi Mohd Yusof. 2010. Digital Image Classification for Malaysian Blooming Flower, Computational Intelligence, Modelling and Simulation (CIMSiM’10).

[14] Nilisback, M.E., Zisserman. 2008. Automated flower classification over a large number of classes, Indian Conference on Computer Vision, Graphics and Image Processing, 722-729.

[15] Kumar, N., Belhumeur, P.N., Biswas, A., Jacobs, D.W., Kress, W.J., Lopez, I.C., Soares, J.V.B. 2012. LeafSnap: A computer vision system for automatic plant species identification, European Conference on Computer Vision. 502-516.

[16] Zheng Zou, Niannian Wang, Peng Zhao, Xuefeng Zhao. "Real-Time Detection of Ancient Architecture Features Based on Smartphones", Volume 2: Mechanics and Behavior of Active Materials; Structural Health Monitoring; Bioinspired Smart Materials and Systems; Energy Harvesting; Emerging Technologies, 2018