

# Final Report on Flower Classification

Lujie Ma  
University of Adelaide  
School of Computer Science  
a1810558@adelaide.edu.au

Wenlong Dai  
University of Adelaide  
School of Computer Science  
a1777081@adelaide.edu.au

## Abstract

*Classification is become more and more popular in our daily life, and gradually become more useful. Image classifier can do lots things for us, e.g. determining people's gender, identifying colors for special populations. To achieve this, we can use deep learning model trained on the specific data set.*

## 1. Introduction

In this project, we implemented the recognition of five types of flowers which are daisy, dandelion, rose, sunflower and tulip. By using the base line model, the test accuracy is close to 71.3% and getting overfitting. Through our effort, the final accuracy can reach 92.6% by adopting the deeper network resnet18.

## 2. Before the start

Here shows some works before starting the project.

### 2.1. Review the dataset

By checking the dataset, we notice that the number of pictures under each category is close to 400. As shown in the Figure 1, we can see that although in the same category, there also has the huge difference which will influence the accuracy. Such as the second picture of sunflower, the woman actually occupy most of the entire picture.

### 2.2. Introduction about baseline model

This baseline uses two convolutional layers to extract features and four fully-connected layers to classify the extracted features. The problem is that the number of convolutional layers is too small, which leads to inadequate extraction of features, and the number of fully-connected layers is too large, which leads to overfitting of the whole network.



Figure 1. Original images

## 3. Apply Tricking

Because of many factors in the given dataset will influence the accuracy, so the baseline model do not works good. For this reason, how to make the neural network learn features without overfitting has become the most important goal of this competition. Four tricks were adopted to complete this competition as shown:

(1) The network structure of resnet18 is adopted and the redundant fully connected layers in the baseline are discarded in order to learn more discriminative features. [2]

(2) The pre-training method is adopted to prevent the resnet18 network from falling into overfitting, and the parameter weights trained on the imagenet dataset can better optimize the entire network parameters. [4]

(3) Most of the time Adam gets the optimal solution is Sharp Minimum, and SGD gets Flat Minimum, which is why we set optimizer to SGD. [3]

(4) The imagenet dataset is very large, so the provided flower dataset and the flower samples in the imagenet dataset have high similarity, and the number of samples in the provided dataset is also relatively large, so different learning rates are adopted in different Network layers. [1]

## 4. Experiment analysis

Four tricks were adopted in this competition, so ablation experiments have been done to verify the effectiveness of each trick. The effectiveness of various tricks obtained through ablation experiments are shown in Table 1. Trick1 adopts the network structure of resnet18, trick2 adopts the pre-training model, trick3 adopts the SGD optimizer, and trick4 adopts different learning rates for different network layers. As can be seen from Table 1, the four adopted tricks have achieved good results, and the entire model has learned more discriminative features.

Table 1. Accuracy – Tricks

| Trick1 | Trick2 | Trick3 | Trick4 | Accuracy |
|--------|--------|--------|--------|----------|
| N      | N      | N      | N      | 71.3%    |
| Y      | N      | N      | N      | 75.6%    |
| Y      | Y      | N      | N      | 86.1%    |
| Y      | Y      | Y      | N      | 88.7%    |
| Y      | Y      | Y      | Y      | 92.6%    |

More intuitive observations can be made in Figure 2 and Figure 3. Figure 2 represents the accuracy curve and the Figure 3 represents the loss curve after applying the different tricks respectively.

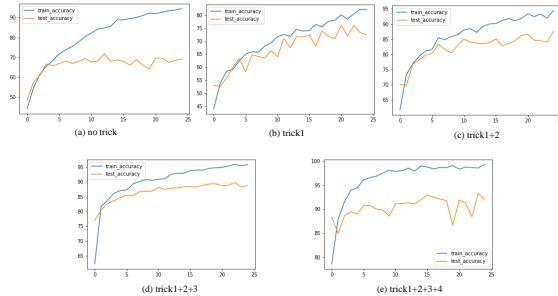


Figure 2. Accuracy under different tricks adopted

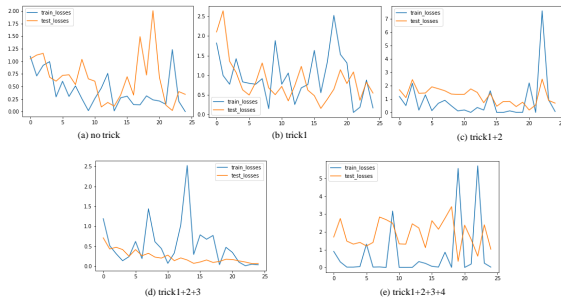


Figure 3. Losses under different tricks adopted

## 5. Observation

As shown in Figure 2, the training accuracy of the baseline is rising, while the test accuracy is falling, and the difference between the accuracy is getting higher and higher. This shows that the baseline easily falls into overfitting, and the entire network does not learn good features. As can be seen from Figure 3, the test loss in the baseline fluctuates greatly in the later stage, indicating that the generalization of the baseline is poor. The adopted series of methods can solve the above problems well, the network has learned more discriminative features, and the generalization of the network has also been improved.

## 6. Conclusion

Flower classification is a task of fine-grained image classification. Fine-grained images have large intra-class differences and small inter-class differences. How to learn more discriminative features is the core of this task. For this reason, the deeper network resnet18 was adopted, and the pre-training method with different learning rates on different layers and SGD optimizer were adopted to improve the efficiency of the entire network. The final classification accuracy reached 92.6%, and the flops was 3.59G.

## 7. Final

You can check our code in our Github repositories: <https://github.com/XiaoLinZzz/CV-project>.

## References

- [1] Dumitru Erhan, Yoshua Bengio, Aaron Courville, Pierre-Antoine Manzagol, Pascal Vincent, and Samy Bengio. Why does unsupervised pre-training help deep learning? *J. Mach. Learn. Res.*, 11:625–660, mar 2010. 1
- [2] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition, 2015. 1
- [3] Blake Woodworth, Kumar Kshitij Patel, Sebastian U. Stich, Zhen Dai, Brian Bullins, H. Brendan McMahan, Ohad Shamir, and Nathan Srebro. Is local sgd better than minibatch sgd?, 2020. 1
- [4] I. Zeki Yalniz, Hervé Jégou, Kan Chen, Manohar Paluri, and Dhruv Mahajan. Billion-scale semi-supervised learning for image classification, 2019. 1