PROJECT 2 REPORT

LIN YUWEI
20446054

1. Problem Description

This project aims to train a deep model to classify 5 kinds of flowers including daisy, dandelion, roses, sunflowers and tulips. There are three datasets, respectively training dataset with 2569 samples, validation dataset with 550 samples and test dataset with 551 samples without labels. Nowadays, many algorithms have been used for image recognition, from original Convolutional Neural Network to some advanced algorithms including VGG, ResNet, GoogleNet and so on. Thanks to ImageNet and updated computation techniques, the performance of latter algorithms can reach as highest as 90% in image classification field. This project attempts to build a pretrained ResNet model and use the training dataset to train the top layer of the model in order to get better performance. The validation data will be used for validating the accuracy of the model and the test dataset will be used to predict the label to calculate the generalization error. In this report, section 1 is the introduction of the project. Section 2 introduces the algorithm principle of ResNet. And at last, section 3 focuses on the experiment and the final result.

2. ResNet

ResNet is proposed by Kaiming He team from Microsoft Research. This algorithm is presented to make deeper neural networks easier to train. Though not directly solving the vanishing gradient problem in deeper layers, the paper introduces a deep residual learning framework: let the layers fit a residual mapping instead of fitting a desired underlying mapping of each layers. By assuming that it is easier to optimize the residual mapping than to optimize the original one, it combines the layers into residual blocks, and establishes an identity mapping from input to output of the blocks. Shown by figure 1, the desire underlying mapping is denoted as $H(x)$, and the residual block is made to fit another mapping of $F(x) \coloneqq H(x) - x$. If the identity mapping was optimal, the residue would be easily reduced to zero so that we don't need to optimize the original stacked layers. The deep residual nets are easy to optimize and easily enjoy accuracy gains when the layers go deeper. ResNet performs better than previous networks and won the 1st place in the ILSVRC 2015 classification competition.

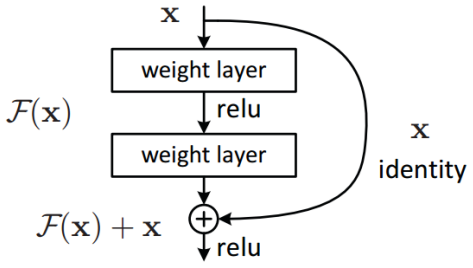

Figure 1

In order to utilize the ResNet framework in this experiment, fine-tune step is introduced. Since ResNet is a very deep neural networks containing a huge number of parameters, it will rise a overfitting problem if only training on the given small training dataset. Therefore, we would train the ResNet on a large dataset ImageNet (with over 1000-class images) and then fine-tune the existing networks by continuing training it on the given training dataset (only 2570 pictures) by very small steps (small learning rates used for not dramatically changing the pre-trained weights). Assumed that our dataset is not drastically different to the large one (ImageNet), the pre-trained model will already have learned features that are relevant to our 5-class flower classification problem.

3. Experiment Result

The experiment is conducted using Python 3.5.0, keras 2.1.1 and tensorflow 1.4.0 as backend.

First, the image should be resized to the standard dimension $224 \times 224 \times 3$ of ResNet model input. When some points outside the boundaries of the input because of transformation, the blank ones would be filled by constant 0.

Next, ResNet model is used for base model with the weights training by ImageNet. In order to map the previous 1000 classes into our 5 classes, it should not include the top layer in the model. Then we can fine-tune the model. Fitting the model by our small dataset, the optimizer should be in very small steps.

Third, after well trained the model, evaluation is made by validation dataset to get the accuracy score. In this experiment, the model reaches accuracy 73.2%.

At last, use the model to make a prediction on test dataset, if the test data is not drastically different to the training data, it will be classified at an accuracy rate lower but near the evaluation score.