CV Assignment 6: Fine-Grained Visual Classification

Jinna Cui

Due Date: June 20, 2017

1 Introduction

In this assignment, you need to do at least eight experiments including image classification experiments and fine-grained visual classification (FGVC) experiment on three datasets. To achieve image classification with deep neural networks, you'd better install **Caffe** (`http://caffe.berkeleyvision.org/`) on your own computer firstly. All of our image classification and FGVC experiments can be done with Caffe.

2 Evaluation

The evaluation of our experiments is overall accuracy.

$$overall\ accuracy = \frac{number\ of\ correctly\ predicted\ items}{total\ number\ of\ items\ to\ predict}$$

3 Experimental details

3.1 Caffe

GPU is not essential to our experiments and we can run our experiments with "CPU only" model. When you install Caffe on your own computer, "CPU only" version is better for you.

3.2 Datasets

In this assignment, we need three datasets: MNIST, CIFAR-10, CUB-200-2011. You can download these datasets from following websites:

MNIST: `http://yann.lecun.com/exdb/mnist/`

CIFAR-10: `https://www.cs.toronto.edu/~kriz/cifar.html`

CUB-200-2011: `http://www.vision.caltech.edu/visipedia/CUB-200-2011.html`

Each dataset contains a training set and a test set. The training set and test set for each dataset are already divided by their publishers so that you can use them directly.

3.3 Experiments

3.3.1 MLP and LeNet-5 on MNIST dataset

1. Choose appropriate numbers of MLP's hiden layers and layer nodes by yourself. In this part, you can get the appropriate numbers of MLP's hiden layers and layer nodes by doing several experiments.

2. Compare the classification results of MLP and LeNet-5 on MNIST dataset.

3.3.2 LeNet-5 on CIFAR-10 dataset

1. Images' size in CIFAR-10 dataset is 64×64 and the input images' size of LeNet-5 is 28×28. As a result, you are required to resize the images of CIFAR-10 dataset before inputing them into LeNet-5.

2. Compare the classification result with the result of LeNet-5 on MNIST dataset and analyze LeNet-5's different classification effectivenesses on CIFAR-10 dataset and MNIST dataset.

3.3.3 CIFAR-10 net on CIFAR-10 dataset

1. CIFAR-10 net is a typical convolutional neural network and the network architecture is shown in Figure 1.

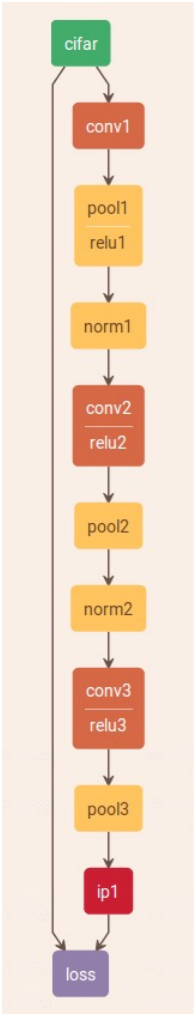

Figure 1: The network architecture of CIFAR-10 net

2. The activation function of CIFAR-10 net is ReLU. In this experiment, you are required to do three different experiments with ReLU, sigmoid and tanh as different activation functions respectively. Only the activation function need to be changed and the other parts of CIFAR-10 net remain the same.

3. The pooling method of CIFAR-10 net is max pooling. In this experiment, you are required to do two different experiments with max pooling and average pooling respectively. Only the pooling methods need to be changed and the other parts of CIFAR-10 net remain the same.

3.3.4 AlexNet on CUB-200-2011

In this experiment, you need to train and test AlexNet on CUB-200-2011 dataset and analyze the experimental result.

3.4 Conclusion

In conclusion, you need at least to do following experiments:

Networks	Datasets	Details
MLP	MNIST	At least one experiment
LeNet-5	MNIST	One experiment
LeNet-5	CIFAR-10	One experiment
CIFAR-10 net	CIFAR-10	ReLU + Max pooling
CIFAR-10 net	CIFAR-10	Sigmoid + Max pooling
CIFAR-10 net	CIFAR-10	tanh + Max pooling
CIFAR-10 net	CIFAR-10	ReLU + Average pooling
AlexNet	CUB-200-2011	One experiment

4 Submission

You need to submit a paper of your experimental analysises especially the comparisons between different deep neural networks, different activation functions and different pooling methods.

Submit your assignment to ouceecv@163.com before due date with the subject: Yourname_Assignment6. The filename of your paper should be: Yourname_Assignment6.pdf.