DTV HW1- Classification

Introduction

In this assignment, it's objective is to train a neural network to do classification. The dataset will be Cifar-10, which consist of 60,000 32x32 color images (RGB) in 10 classes, with 6,000 images per class. There are 50,000 training images and 10,000 test images.

Framework used in this code is pytorch, and trying to get the least number of parameter in the neural network. Testing data is not used for training in this case.

Experiment Setup

The model for this neural network is a simple CNN classifier, which consists of:

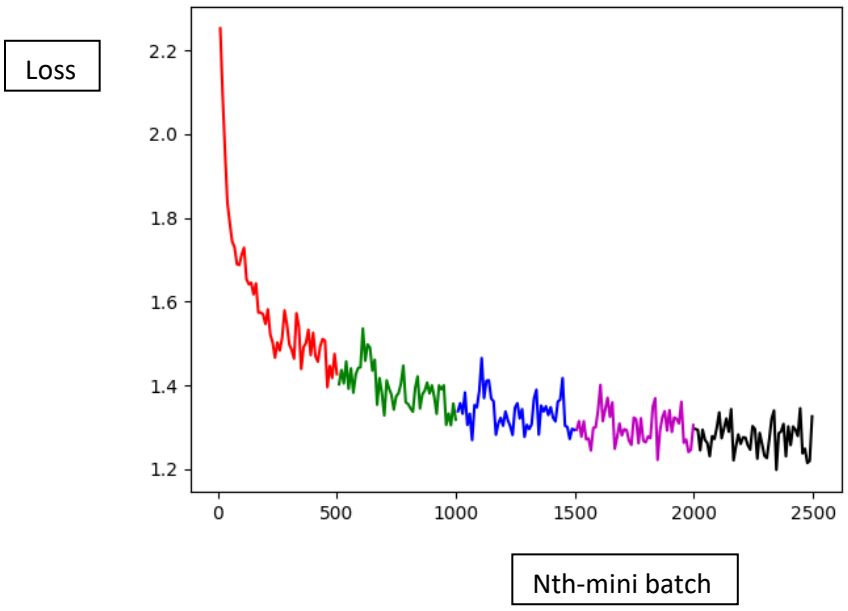
- Input Layer (3 x 32 x 32)
- Convolution Layer (5 filter and 5 unit filter size)
- ReLU Activation Function
- Max Pooling Layer (4 x 4)
- Fully Connected Layer (5x7x7 → 45)
- ReLU Activation Function
- Output Layer (45 → 10)

Number of Parameters: 11,910

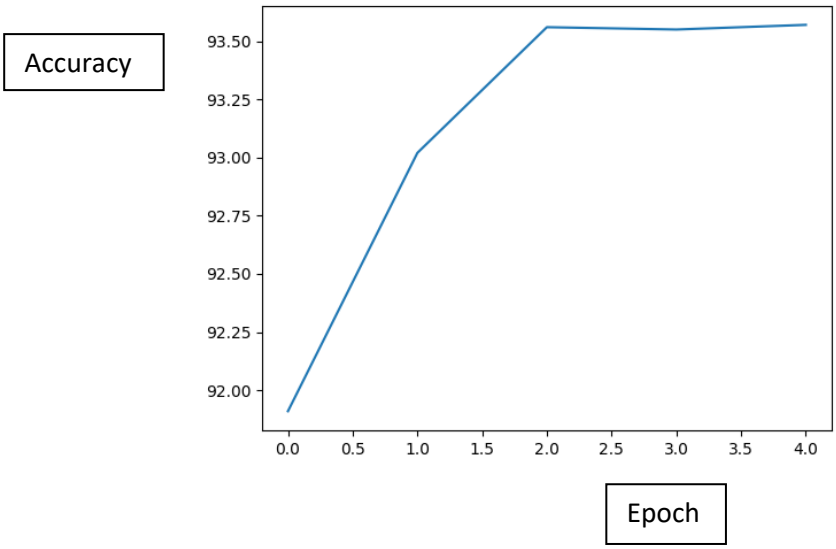
Result

Top 5 Accuracy Rate: 93.6%
Top 5 Error Rate: 6.4%

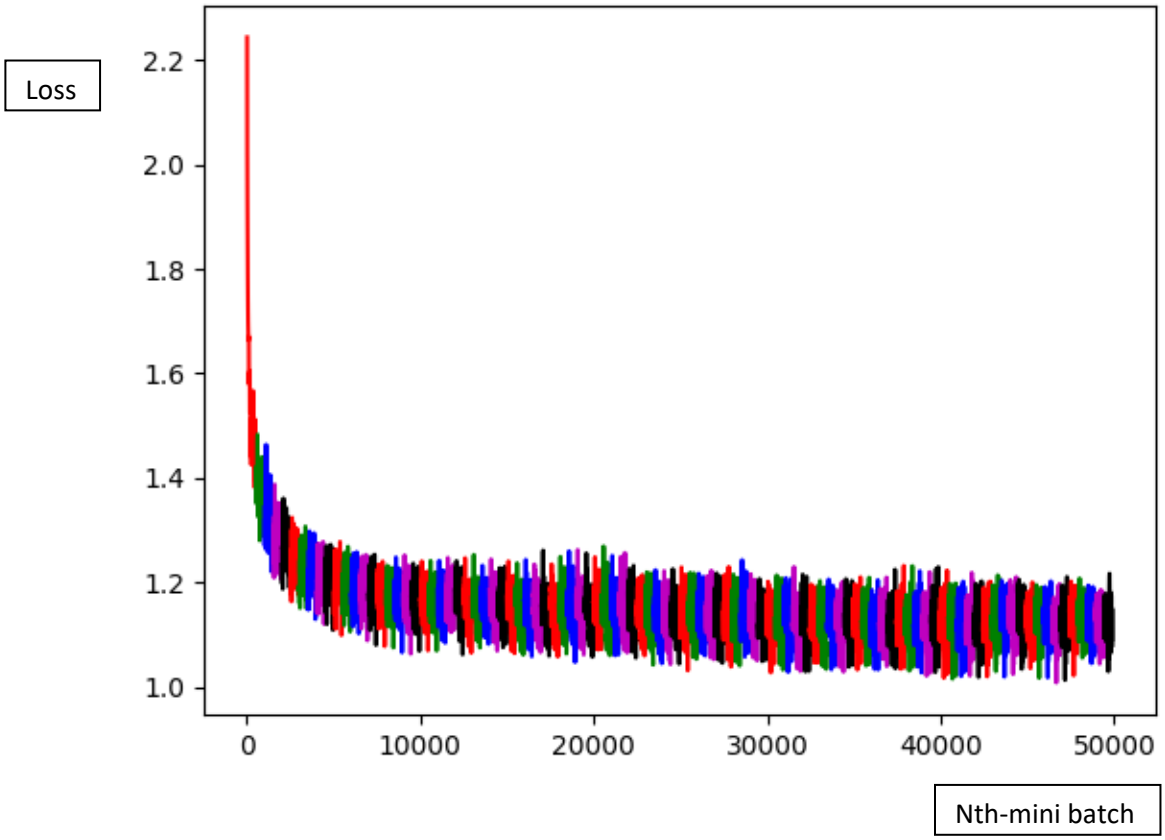
Graph for loss rate of the first 5 epoch of training: (Different color represent different epoch)



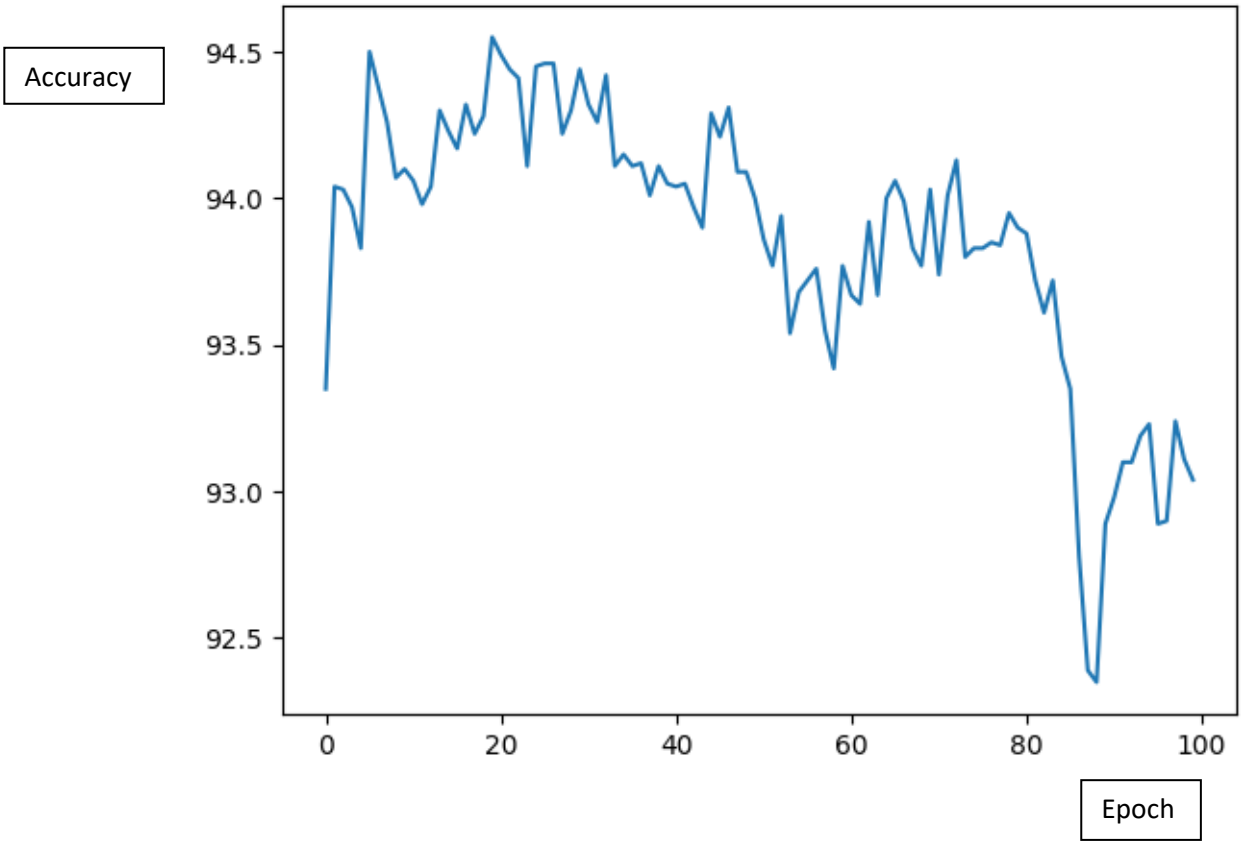
Graph of test classification top-5 accuracy of the first 5 epoch: (By each epoch)



Graph of loss rate for 100 epoch: (Different color represent different epoch)



Graph of test classification top-5 accuracy for 100 epoch:



Discussion

- Due to simple design of CNN, we need to run through the training data only once to gain good enough result. As can be seen in the graph for 100 epoch, no significant improvement can be made from iterating through the training data for more.
- Big filter in the Max Pooling Layer diminish many distinguishing feature of the image that has already been small enough (32x32 pixel). However, it is chosen to be quite big to reduce the number of parameter significantly as soon as possible.