

IE 534 Homework 4 Report

Hanwen Hu

October 10, 2018

This homework asks us to train two networks, i.e. Residual Network (ResNet) specified in the homework sheet, as well as tuning a pretrained ResNet-18 on CIFAR 100.

1 Hyperparameter Settings and Time Cost

For the first ResNet, the training process runs 50 epochs in total, costing 101 minutes. ADAM is applied during training, and the original learning rate is given as $1e-3$. In addition, dropout rate is set as 0.2.

Input images are shuffled, then divided into mini batches, with batch size as 128. Two types of data augmentation are implemented: Random horizontal flip and random center crop with 32×32 and padding 4.

For the second ResNet, the training process runs 23 epochs in total, costing 5 hours and 55 minutes. ADAM is applied during training, and the original learning rate is given as $1e-5$.

Input images are shuffled, then divided into mini batches, with batch size as 100. Four types of data augmentation are implemented: Resize as 224×224 images, random horizontal flip, random color jitter and normalize with mean (0.485, 0.456, 0.406) and standard error (0.229, 0.224, 0.225).

2 Accuracy Graph

For the first ResNet, the accuracy graph is shown as Figure 1. The test accuracy achieves 60.05% at the 31st epoch, but is then oscillating around 60%. It seems that my code needs to be improved before it can provide a higher accuracy.

And for the pretrained ResNet, the accuracy graph is shown as Figure 2. The test accuracy achieves 70.26% at the 13th epoch, and achieves 73.47% at the 23rd epoch. Since the walltime is set as 6 hours, the algorithm is then terminated. If given more training time, this number may be able to be over 75%.

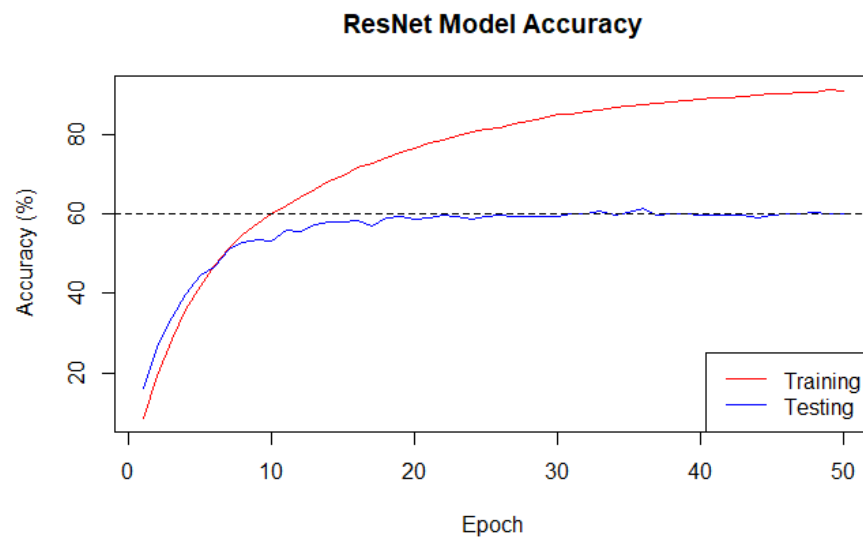


Figure 1: Accuracy graph of ResNet specified in homework sheet

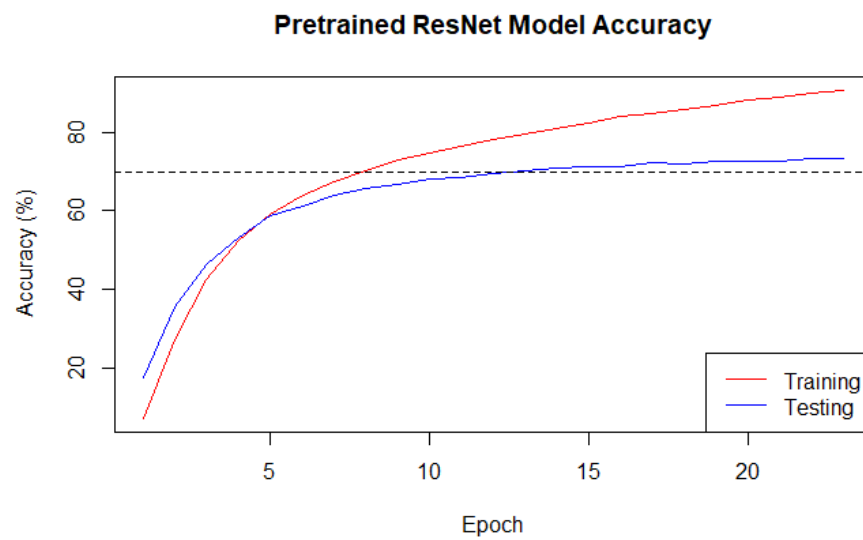


Figure 2: Accuracy graph of tuning pretrained ResNet-18