

HW4

Part 1:

The final test accuracy achieved is **60%**.

Description:

I used the architecture described in the homework, but I added convolution layers between the basic blocks to match the dimension of the residual from previous block with the dimension of the output from current block.

In the Mini-Batch Stochastic Gradient Descent, I used batch size = 100.

Also, I performed data augmentation by random flipping the images horizontally with $p = 0.2$ and random rotate the data by -8 to $+8$ degree. I also used learning rate schedule by decreasing the learning rate by a factor of 10 every 8 epochs.

Part 2:

After fine tuning the ResNet18 model with the CIFAR100 training data, I achieved a test accuracy of **79%**.