CS5260 Homework6 Report

Hu Wangyang A0232499U

https://github.com/lonelyhwy/CS5260-Assignment6

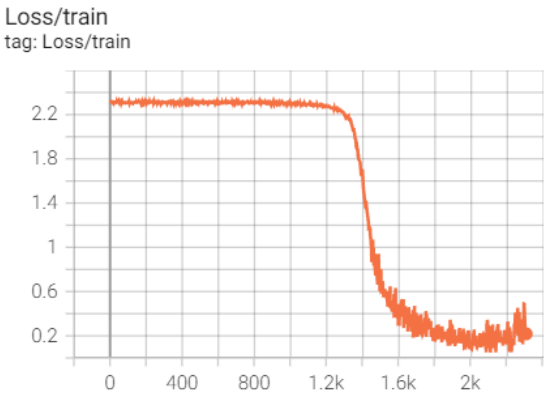
1. Optimizer:

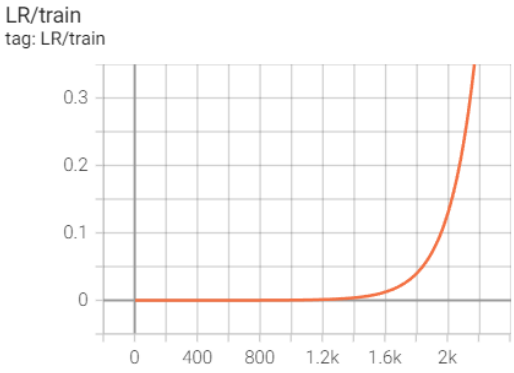
   SGD (model.parameters(), lr=0.1, momentum=0.9, weight_decay=5e-4)

2. Scheduler:

   OneCycle and Multistep

3. LR range test based on SGD optimizer.



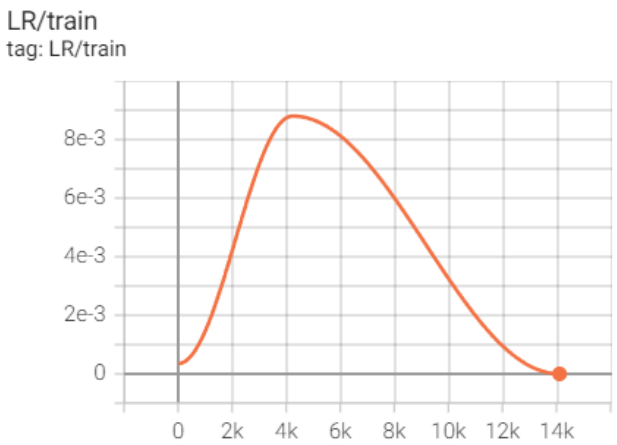


According to the left figure, I choose LR between step 1.3k and 1.5k because in this range, Loss dropped very fast. The corresponding LR is [$3.5e^{-3}$, $8.8e^{-3}$].

4. Training

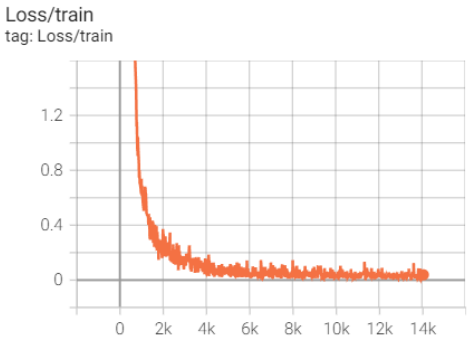
a. SGD optimizer(lr = 3.5e-3) + OneCycle scheduler(max_lr = 8.8e-3)

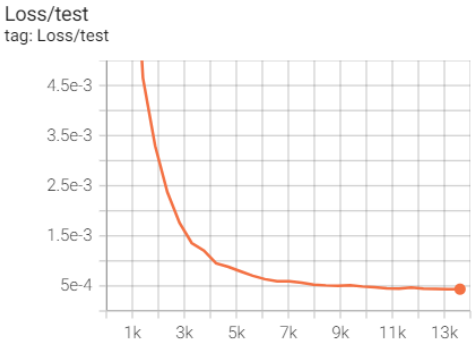
First, let's see the LR curve changes with steps. LR firstly increased to max_lr



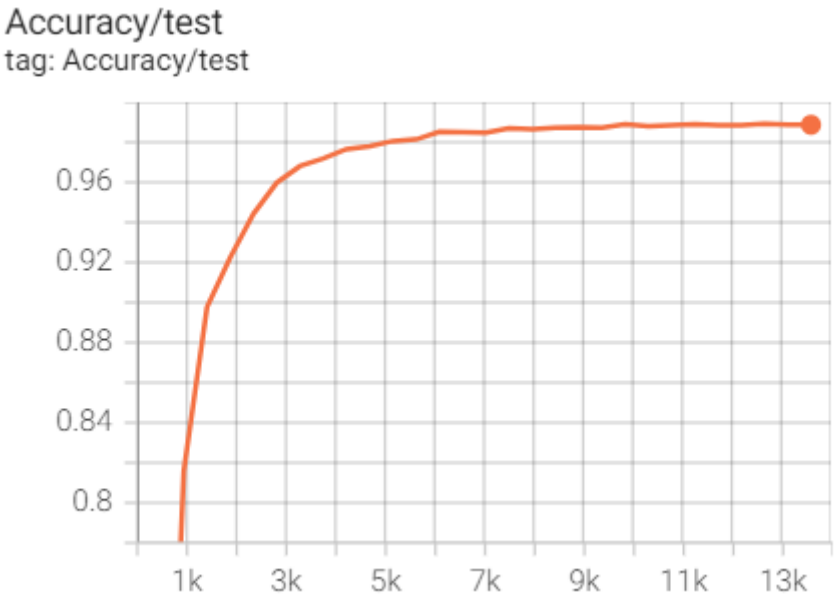
= 8.8e-3. Then LR reduces to nearly 0.

Next, let's see the train & test loss. Although training loss has experienced fluctuation, testing loss has kept dropping.





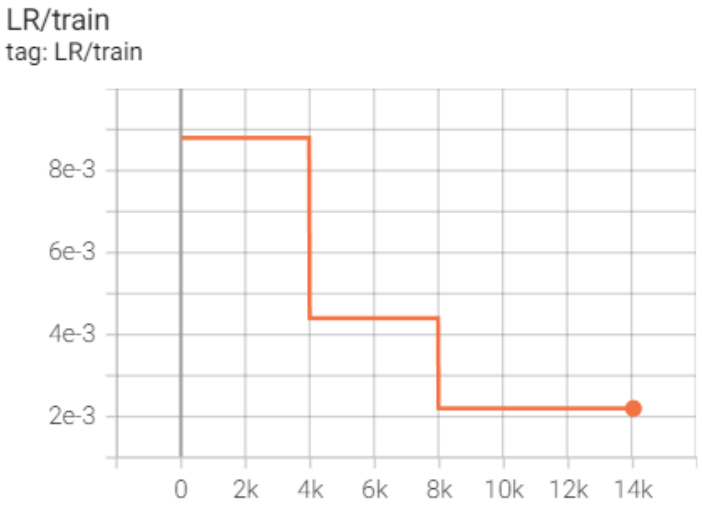
Finally, let's check the accuracy of testing. Turn on the Accuracy Hook by this line of code: hooks.AccuracyHook(accuracy_func=Accuracy()). We achieved 98.87% accuracy.



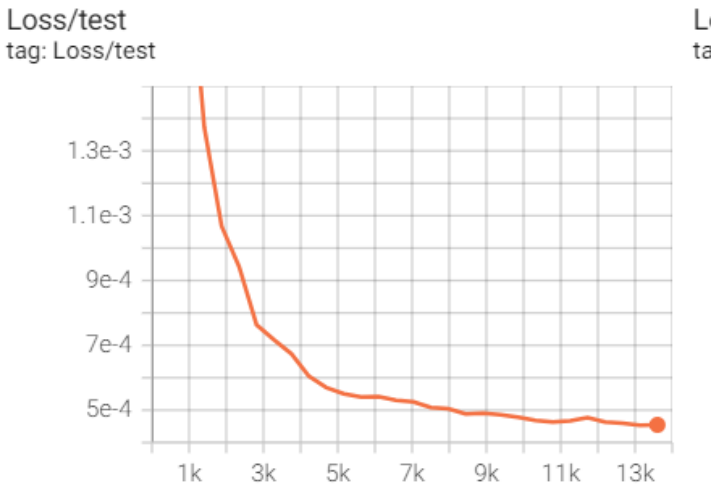
b. **SGD optimizer(lr = 8.8e-3), Multistep scheduler()**

Multistep scheduler lets the lr to decay by gamma once the number of step reaches one of the milestones. I initialized lr = 8.8e-3 according to LR range test. And let milestones be [4000, 8000] because the whole steps are around 13000.

First, let's see the LR curve. It seems to decay to half at step 4000 and step 8000.



Then we see the curve of testing loss, and we find that testing loss keeps dropping.



Finally we check the testing accuracy. It achieves 98.76% in the end.