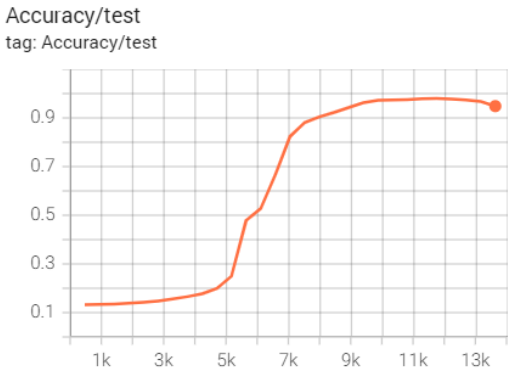
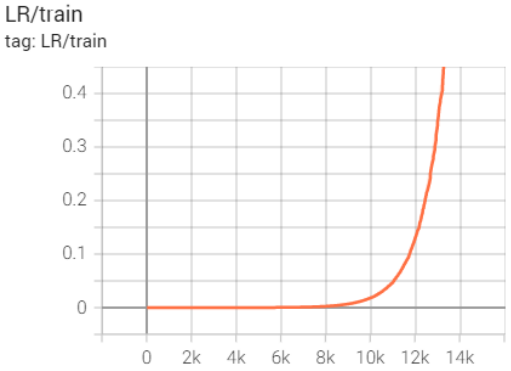
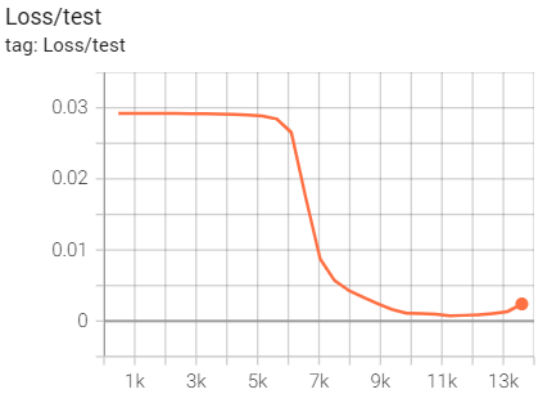
WANG JIYU A0248372Y
CS5260 Assignment 6

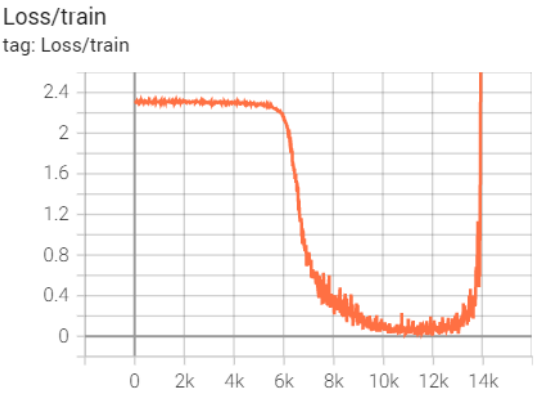
In this assignment, I have conducted the experiment under the following settings:
1. Using SGD optimizer and use LambdaLR to adjust the learning rate (baseline);
2. Using SGD optimizer and use MultiStepLR to adjust the learning rate;
3. Using SGD optimizer and use ExponentialLR to adjust the learning rate;
5. Using Adam optimizer and use LambdaLR to adjust the learning rate.

Baseline:



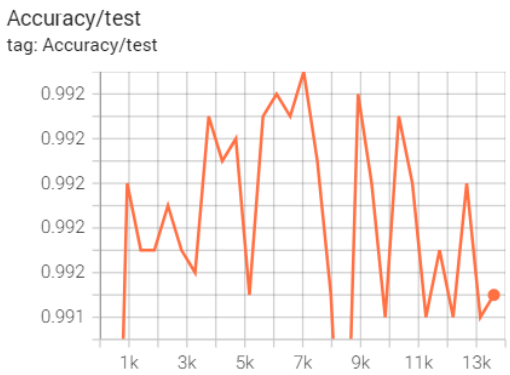


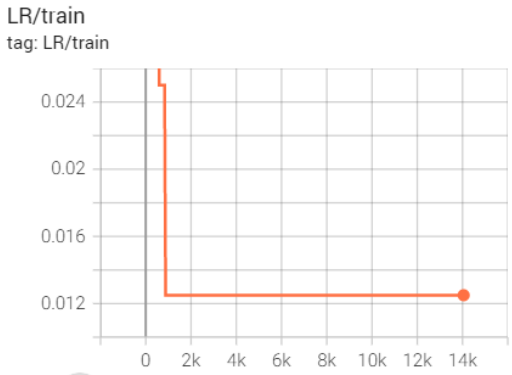


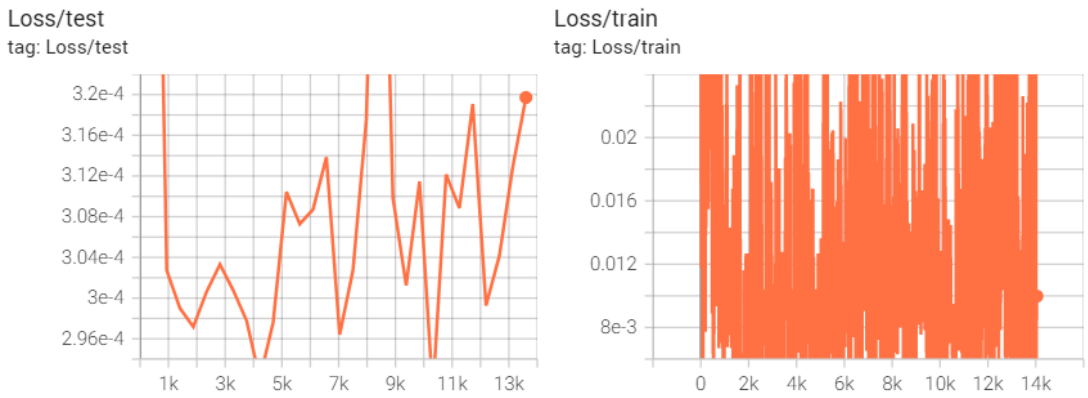


It can be observed that the test accuracy is steadily increasing, but at nearly 1.4k steps, the training loss plummets.

Second setting (MultiStepLR):

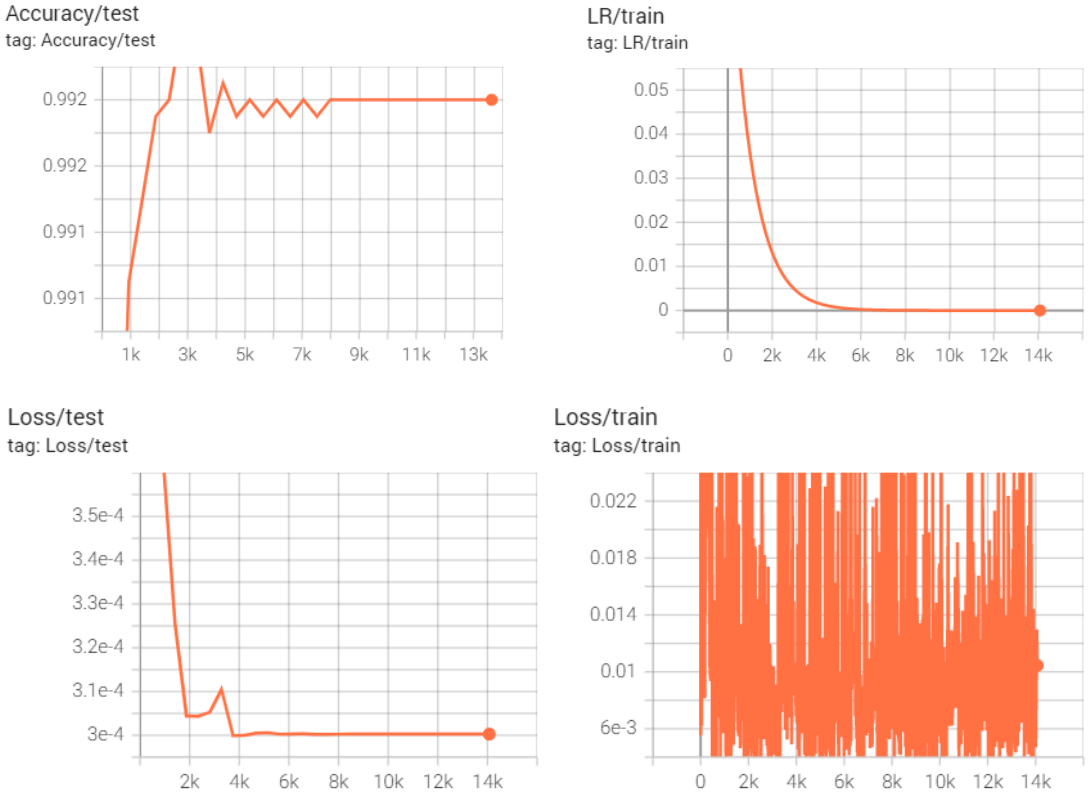






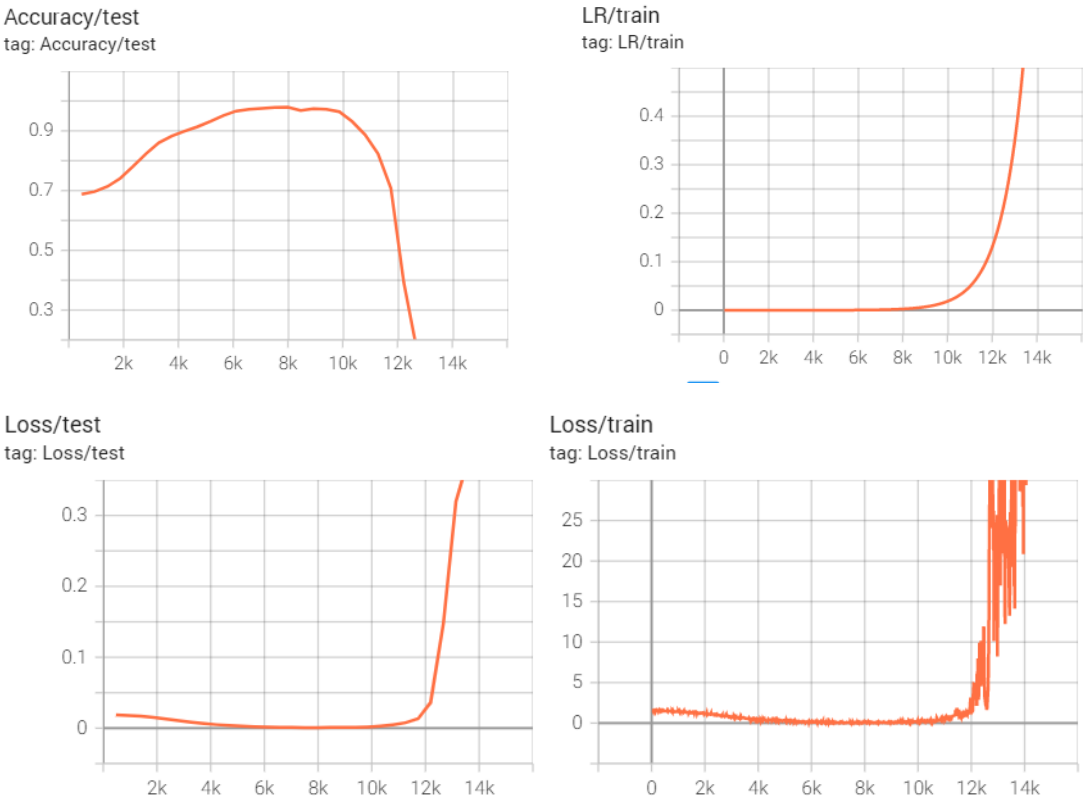
It can be observed that the test accuracy fluctuates up and down from 0.99, accompanied by a constant learning rate to around 0.012 after a sudden drop. The value of the loss fluctuates up and down from a value close to 0. Overall, the results of setting 2 are better than those of baseline.

Third setting (ExponentialLR):



It can be observed that the test accuracy rises to 0.992 and floats for a period of time before stabilizing at 0.992, accompanied by a smooth decrease in the learning rate to 0. The value of test loss stabilizes and decreases to 3e-4, but the value of train loss is still floating up and down at a value close to 0.

Forth setting (Adam):



After changing the Adam to optimizer, it is found that the increased learning rate has a significant negative impact on the accuracy and loss of the model. After about 11k steps, the increased learning rate causes the test accuracy to plummet to 0 and the loss to increase. It can be seen that SGD is more suitable than adam to be chosen as optimizer.

In summary, among these changes, the combination with the best results should be to use SGD as an optimizer and ExponentialLR to change the learning rate.

The Github link: https://github.com/WeenyJY/CS5260_assginment6