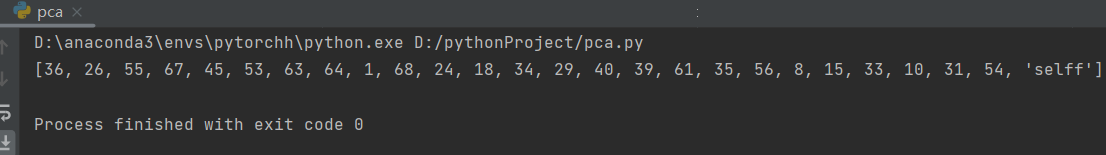


XUEJIALU A0263762Y

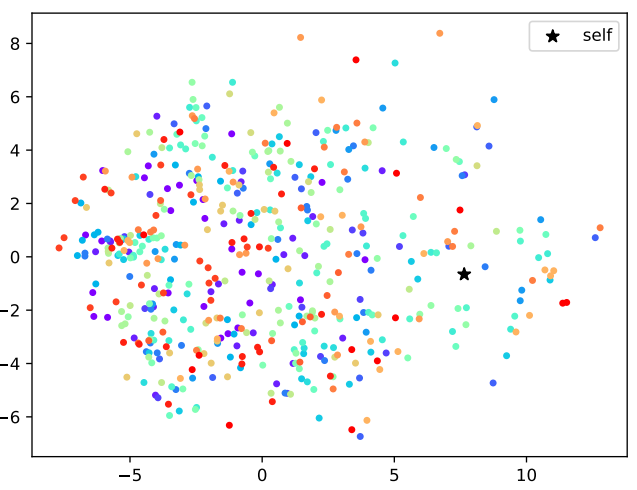
PCA

2D

Select subjects

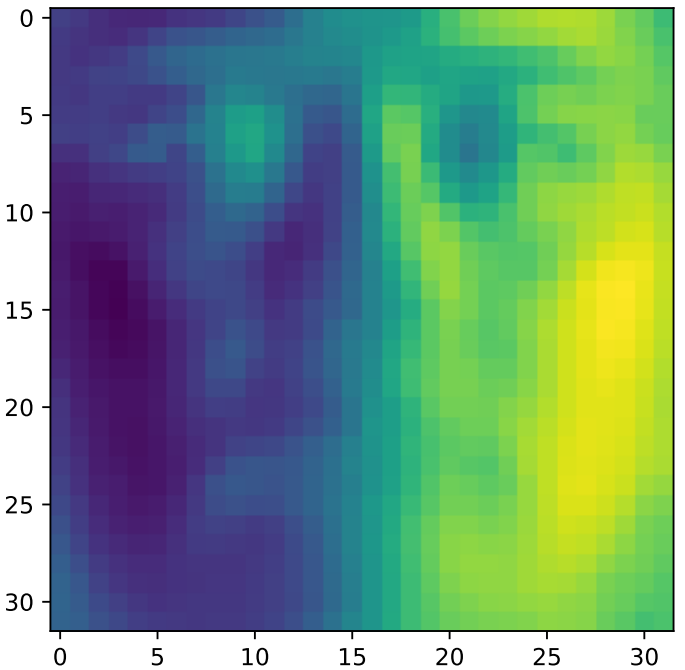
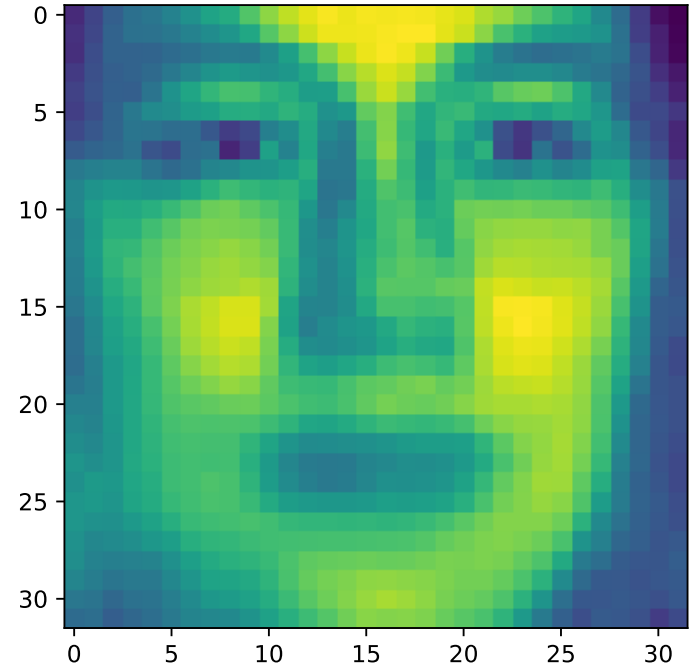


Data Distribution Visualization



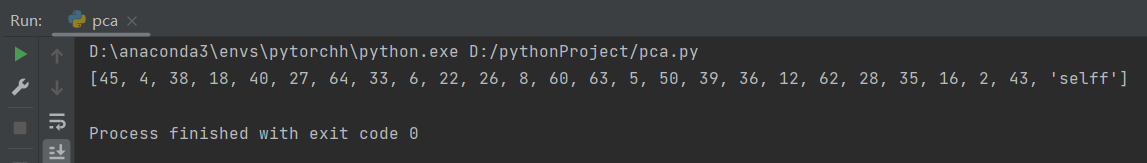
Different colors repersent projected points from different subjects, and my own selfie photos are marked with black star.

eigenfaces

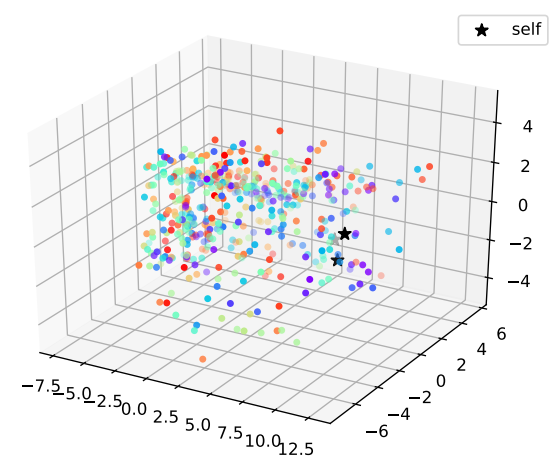


3D

Select subjects

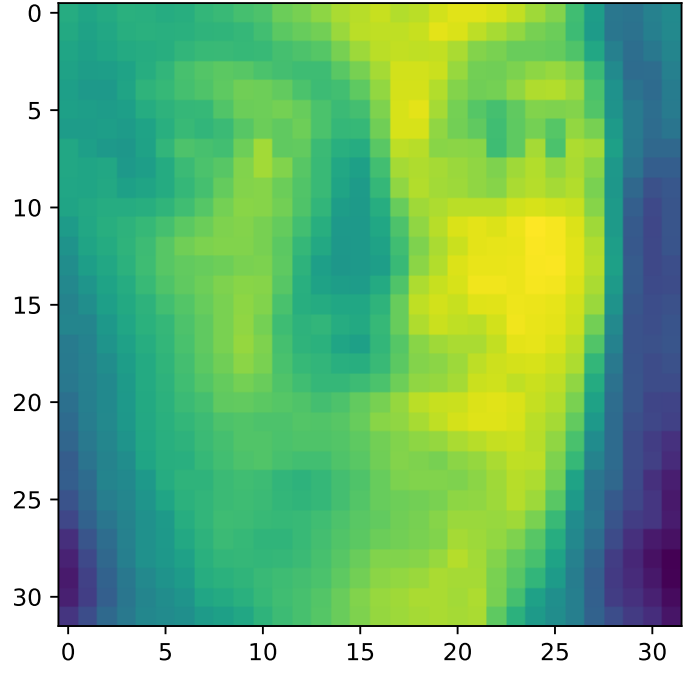
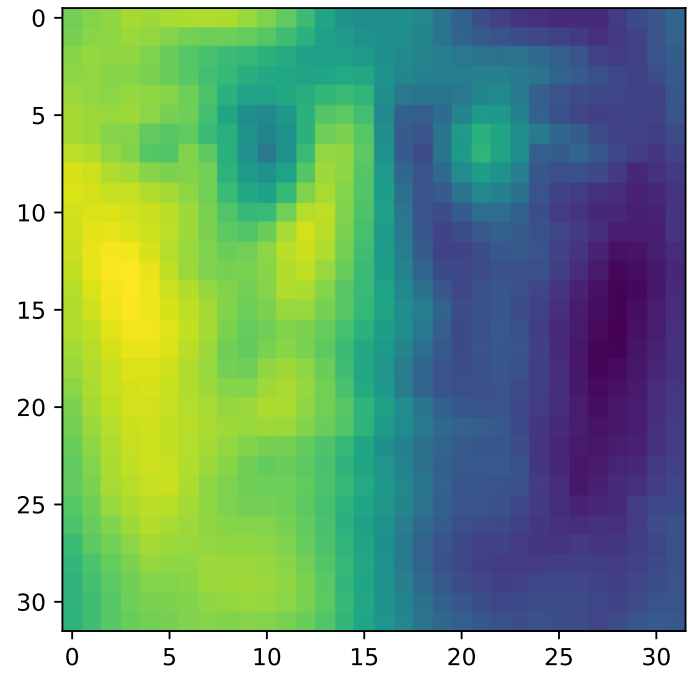
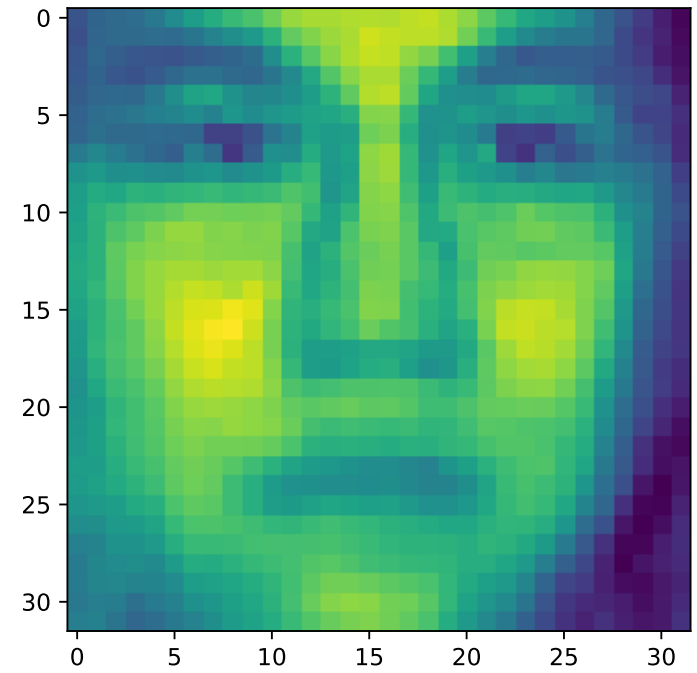


Data Distribution Visualization

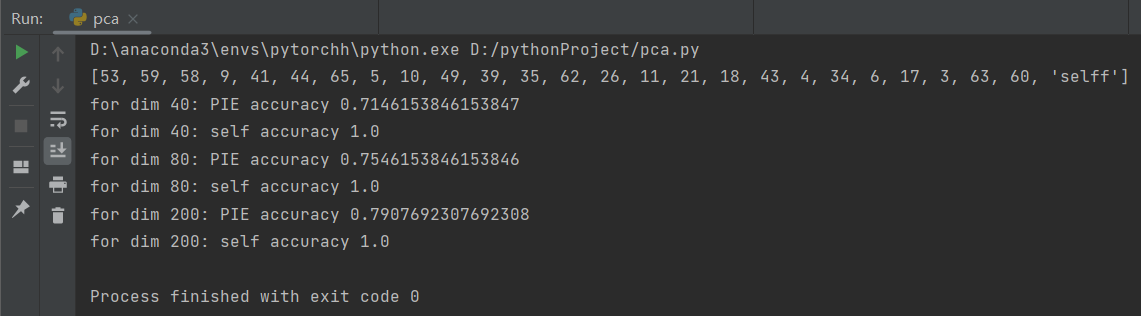


Different colors repersent projected points from different subjects, and my own selfie photos are marked with black star.

eigenfaces



Classification Results



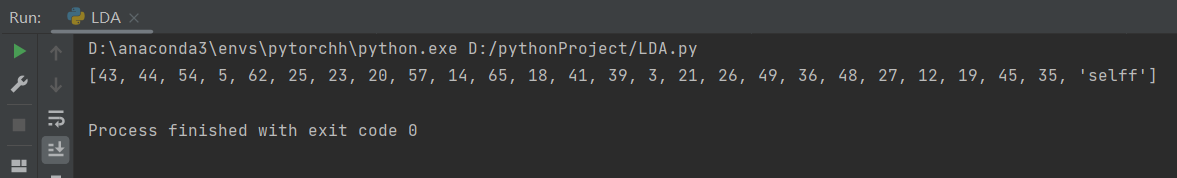
the accuracy of my own selfie have the same accuracy of 100% on PCA 40 80 200, the accuracy of the PIE data is 71.5%, 75.5% and 79.1% for PCA 40, 80 and 200.

For the PIE data, we can see that PCA at higher dimensions, the accuracy increases gradually.

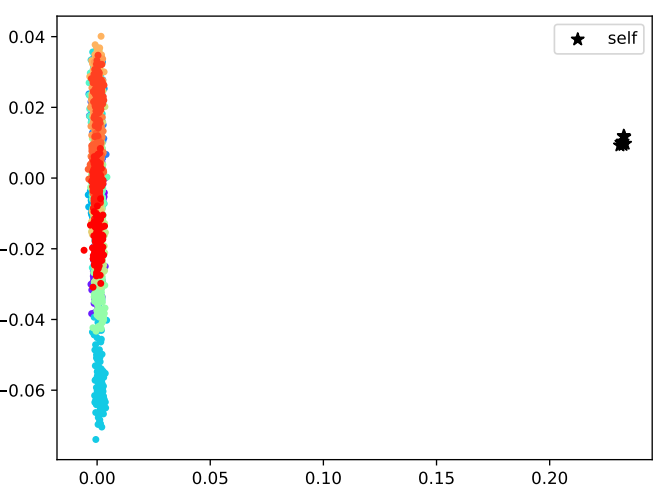
LDA

2D

Select subjects



Data Distribution Visualization

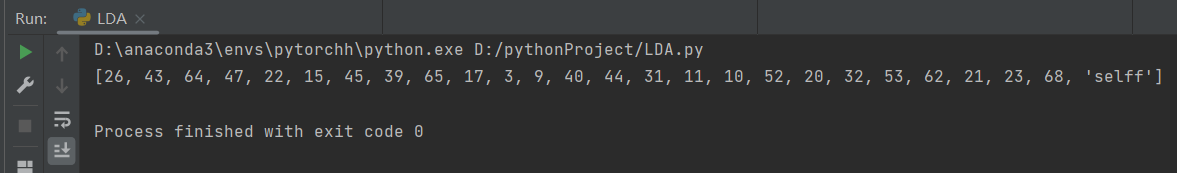


Different colors repersent projected points from different subjects, and my own selfie photos are marked with black star.

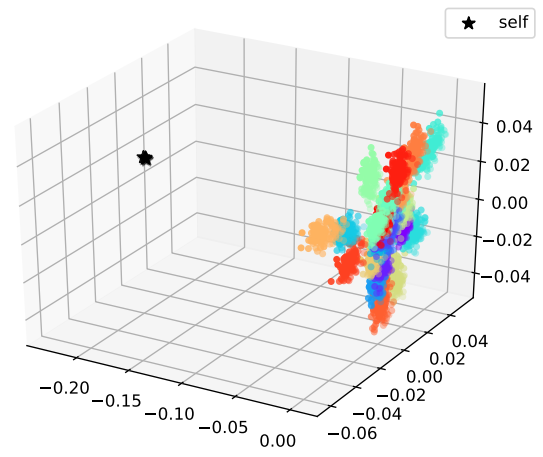
LDA realizes the distinction between classes by maximizing the "inter-class distance" and minimizing the "intra-class distance". Because the shooting environment of my selfie is quite different from that of the data set, it is easier to distinguish, and it is farther away from other data after projection.

3D

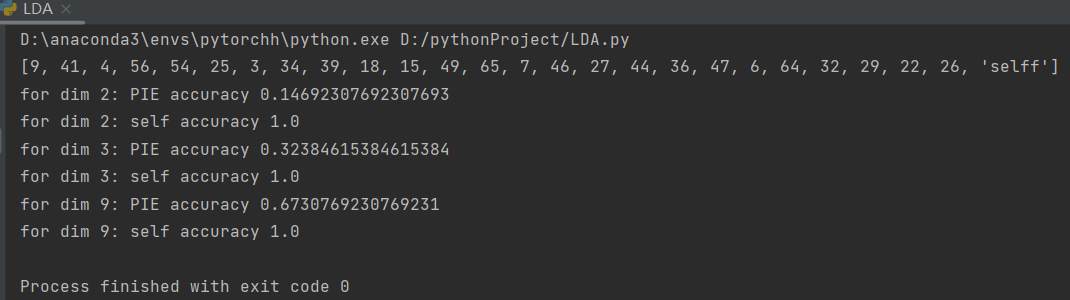
Select subjects



Data Distribution Visualization



Classification Results

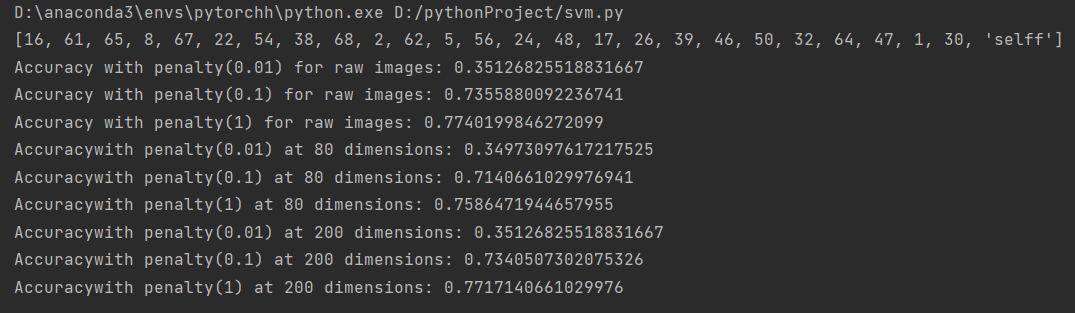


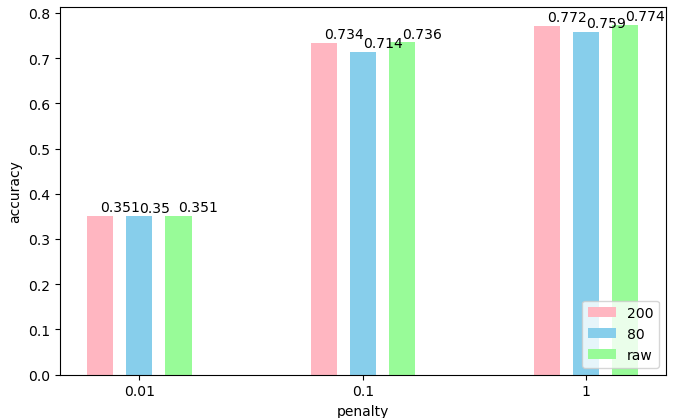
the accuracy of my own selfie have the same accuracy of 100% on LDA2, 3, 9, the accuracy of the PIE data is 14.7%, 32.4% and 67.3% for LDA 2, 3, 9.

For the PIE data, we can see that PCA at higher dimensions, the accuracy increases gradually.

SVM

Select subjects





Influence of dimension and penalty parameter

The penalty parameter C represents the extent of SVM can avoid misclassification of each training data. For larger C values, the optimization will select a hyperplane with smaller boundaries if the training points can be better classified correctly. Conversely, a smaller value of C causes the optimizer to look for a larger bounded hyperplane, at the cost of potentially misclassifying more points.

The original image and the reduction of dimension to 80 and 200 have no significant effect on the result. It is proved that the performance of SVM may be independent of the dimension of feature space.

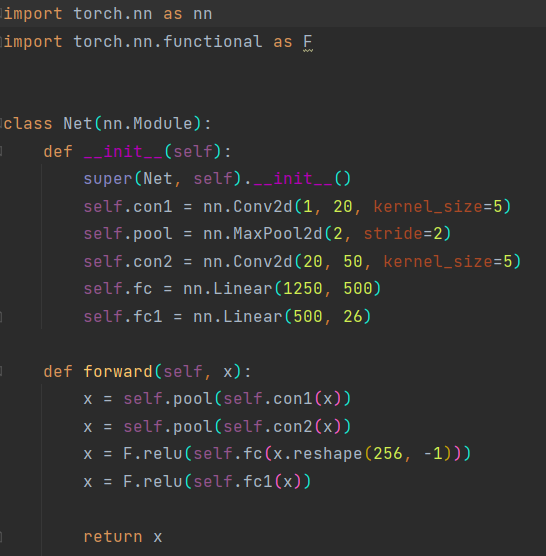
CNN

Select subjects

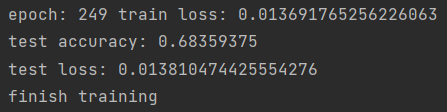


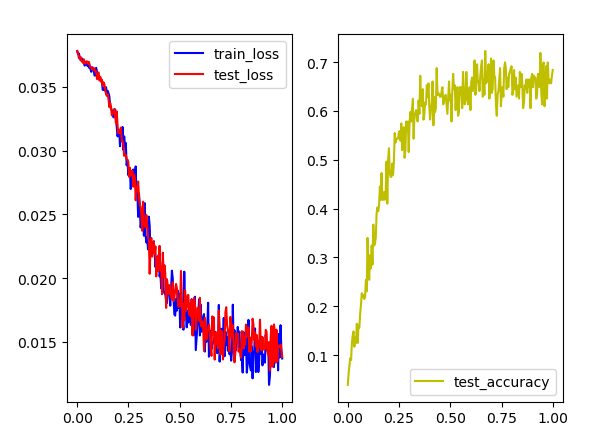
Network Architecture 1

two convolutional layers and one fully connected layer, with the architecture specified as follows: number of nodes: 20-50-500-26.



Result

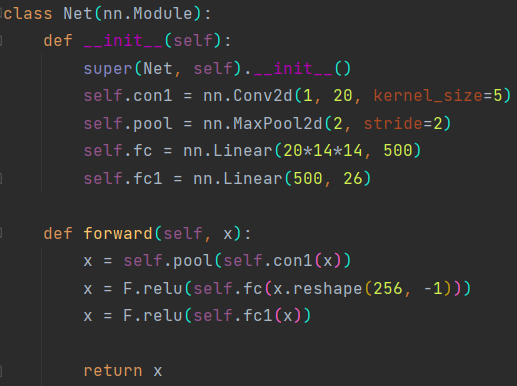




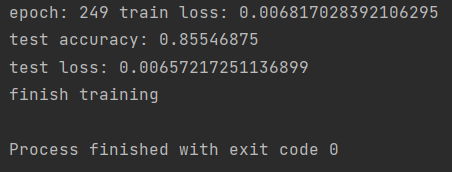
Finally, the test loss drop to 0.013, and the test accuracy rise to 68.4%

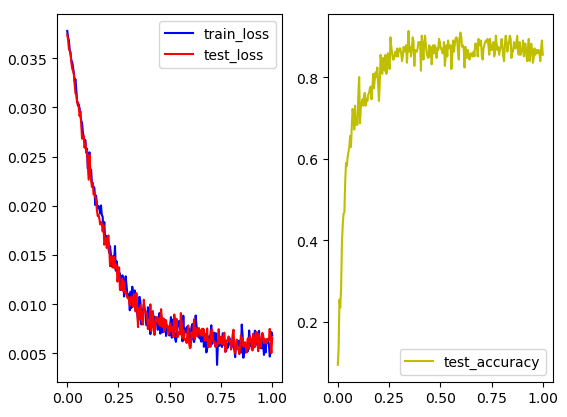
Network Architecture 2

I've removed the second convolution layer on the base of the first architecture.



result

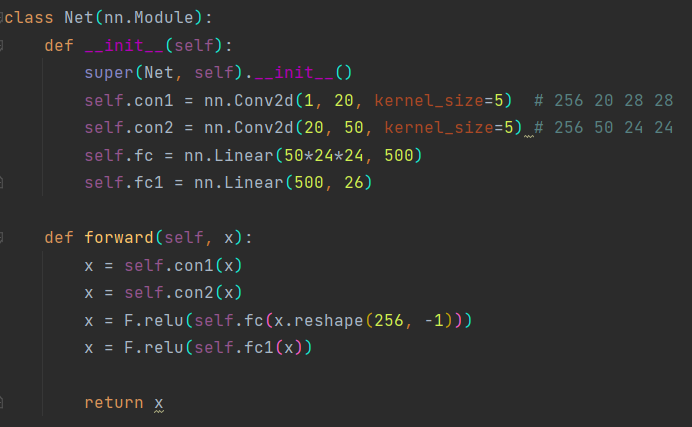




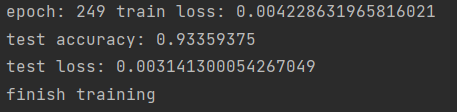
the test loss drop to 0.0065, and the test accuracy rise to 85.5%

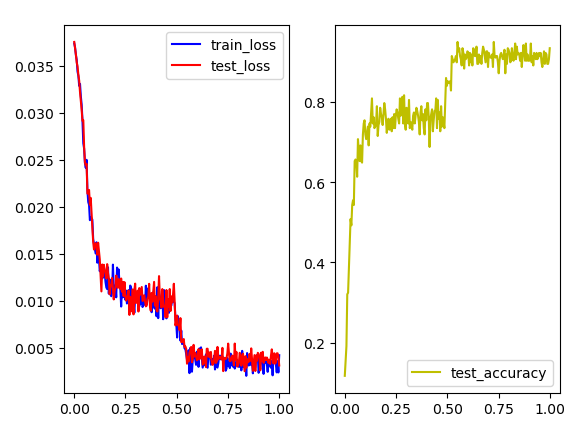
Network Architecture 3

I've removed all the maxpool layers on the base of the first architecture.



Result





The test loss drop to 0.0031, and the test accuracy rise to 93.3%. compare to the result of the first one, it can be seen that the existence of maxpool layer will reduce the performance of convolutional neural network.