

CS 535 Deep Learning Assignment 3

Zhou Fang

March 1, 2018

1

By adding a batch normalization layer after the first fully-connected layer, the validation accuracy improved. The best validation accuracy is 79.870%. The train loss, test loss, train accuracy and test accuracy are shown at Table 1, figure 1 and figure 2.

Table 1: Performance of CNN

epoch	train loss	test loss	train accuracy	test accuracy
1	1.000 83	1.072 45	64.114%	62.110%
2	0.631 02	0.759 11	77.968%	73.750%
3	0.469 84	0.693 01	83.856%	76.080%
4	0.353 20	0.688 44	88.020%	76.440%
5	0.254 66	0.672 88	91.876%	77.640%
6	0.158 99	0.694 85	95.334%	77.910%
7	0.110 97	0.712 65	96.808%	78.260%
8	0.076 13	0.742 00	98.094%	78.330%
9	0.049 92	0.796 38	98.768%	78.470%
10	0.035 68	0.805 64	99.224%	79.580%

2

By adding another fully connected layer with 512 nodes at the second-to-last layer, there is not much improve to the accuracy. The best validation accuracy is 78.130%. The train loss, test loss, train accuracy and test accuracy are shown at Table 2, figure 3 and figure 4.

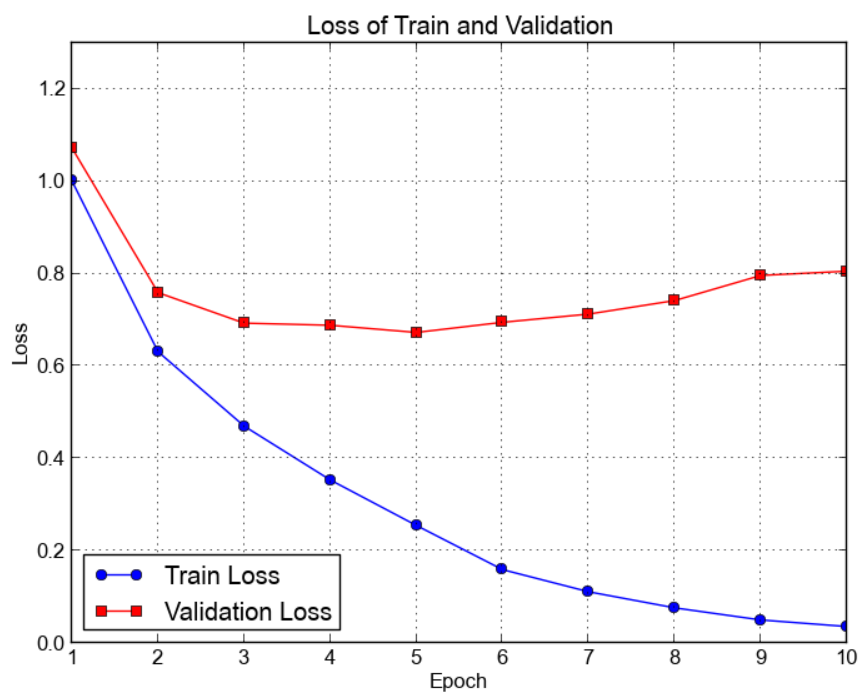


Figure 1: Loss of Train and Validation Adding a batch Normalization Layer

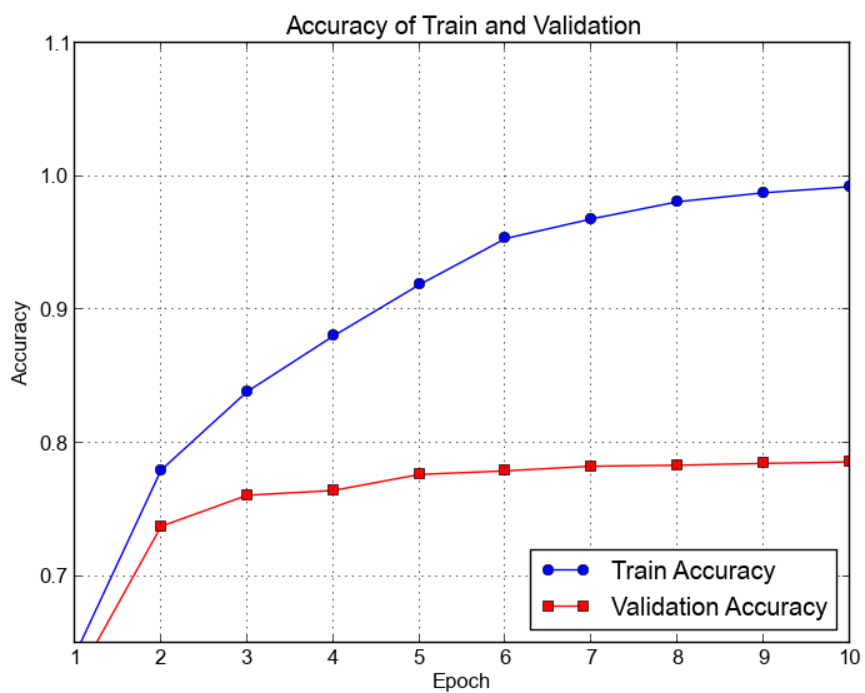


Figure 2: Accuracy of Train and Validation Adding a batch Normalization Layer

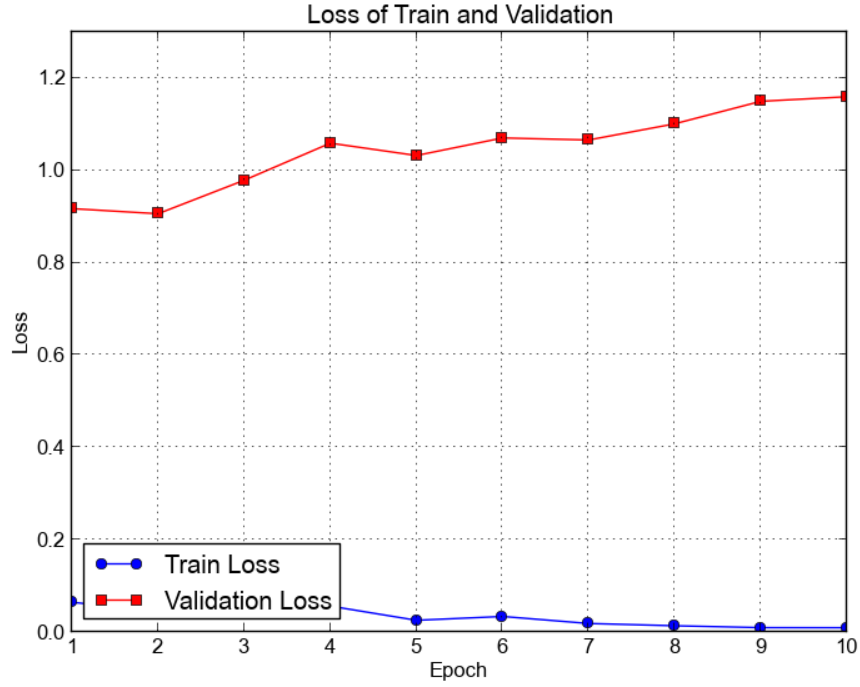


Figure 3: Loss of Train and Validation Adding Another Full Connected Layer

Table 2: Performance of CNN

epoch	train loss	test loss	train accuracy	test accuracy
1	0.064 53	0.917 29	97.910%	77.550%
2	0.044 20	0.906 23	98.678%	77.940%
3	0.042 61	0.978 80	98.632%	77.820%
4	0.056 13	1.059 15	98.028%	77.750%
5	0.025 36	1.032 38	99.220%	78.310%
6	0.033 67	1.070 41	98.918%	78.150%
7	0.018 46	1.066 04	99.460%	78.210%
8	0.013 50	1.101 12	99.614%	78.270%
9	0.009 46	1.149 86	99.730%	78.780%
10	0.009 31	1.159 67	99.744%	78.640%

3

We try to use Adam to tune the learning rate. The best validation accuracy is 79.000%. The train loss, test loss, train accuracy and test accuracy are shown at Table 3, figure 5 and figure 6.

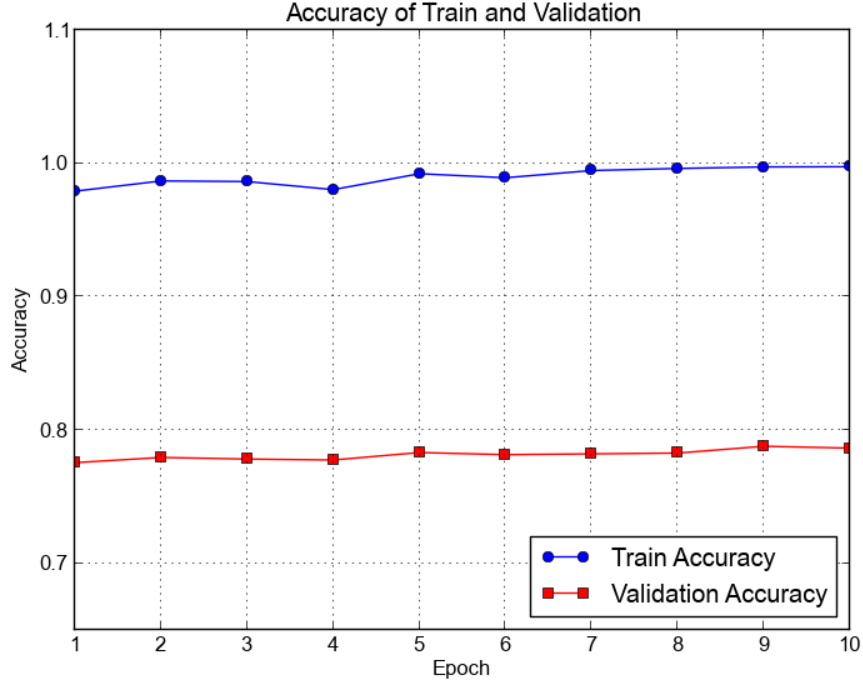


Figure 4: Accuracy of Train and Validation Adding Another Full Connected Layer

Table 3: Performance of CNN

epoch	train loss	test loss	train accuracy	test accuracy
1	0.761 38	0.845 99	73.318%	70.530%
2	0.540 61	0.703 42	81.264%	75.740%
3	0.401 38	0.660 40	86.904%	77.150%
4	0.340 15	0.713 22	88.330%	76.420%
5	0.231 29	0.728 79	92.362%	77.060%
6	0.117 31	0.698 97	96.628%	78.740%
7	0.077 96	0.777 49	97.798%	78.060%
8	0.084 51	0.867 91	97.262%	76.710%
9	0.048 17	0.855 02	98.636%	78.450%
10	0.036 68	0.867 39	99.010%	79.000%

4

Firstly, we tune our network by removing one full-connected layer. The best validation accuracy is 74.080%. The performance is worse than other situations. The train loss, test loss, train accuracy and test accuracy are shown at Table 4, figure 7 and figure 8.

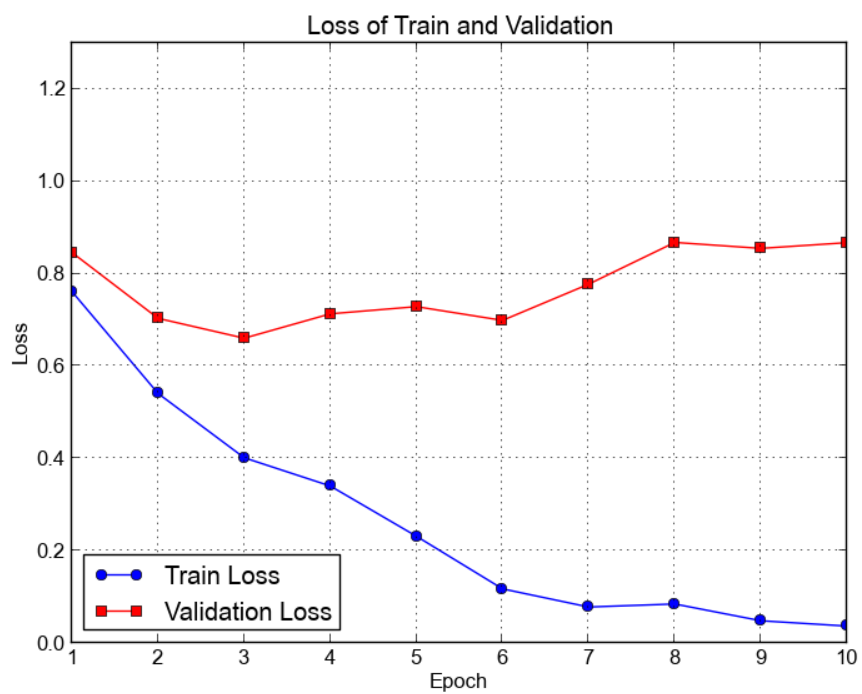


Figure 5: Loss of Train and Validation Using Adam Adaptive Schedule

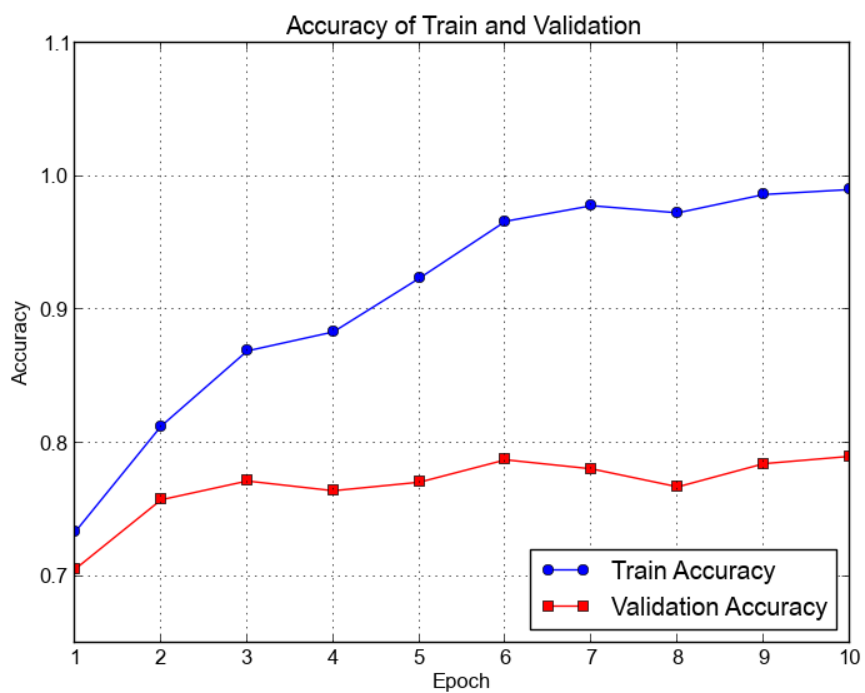


Figure 6: Accuracy of Train and Validation Using Adam Adaptive Schedule

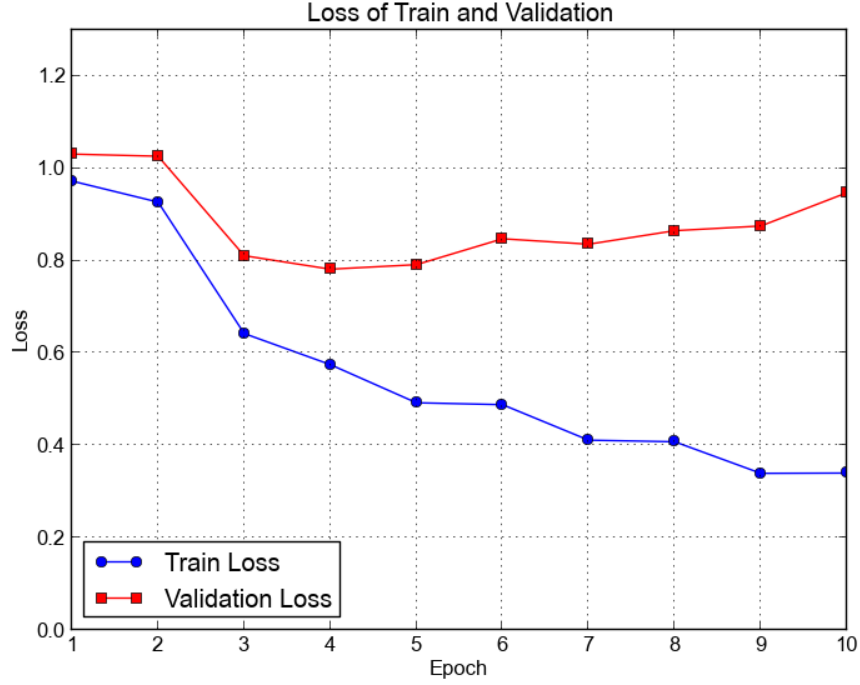


Figure 7: Loss of Train and Validation Removing One Full-connected Layer

Table 4: Performance of CNN

epoch	train loss	test loss	train accuracy	test accuracy
1	0.972 73	1.031 13	65.968%	63.700%
2	0.927 08	1.026 49	67.870%	64.990%
3	0.642 35	0.811 00	77.526%	71.980%
4	0.575 00	0.781 76	80.372%	73.160%
5	0.492 32	0.791 55	82.846%	74.080%
6	0.488 22	0.847 68	82.668%	72.510%
7	0.411 28	0.835 88	85.688%	73.190%
8	0.407 81	0.865 20	85.524%	72.190%
9	0.339 42	0.875 33	88.140%	73.230%
10	0.340 00	0.947 51	87.874%	72.250%

Secondly, we add two more convolution networks (128×128 filters and 256×256 filters) and decrease learning rate to 0.005. The best validation accuracy is 81.620%. This is the best performance I get by modifying the model. The train loss, test loss, train accuracy and test accuracy are shown at Table 5, figure 9 and figure 10.

Table 5: Performance of CNN

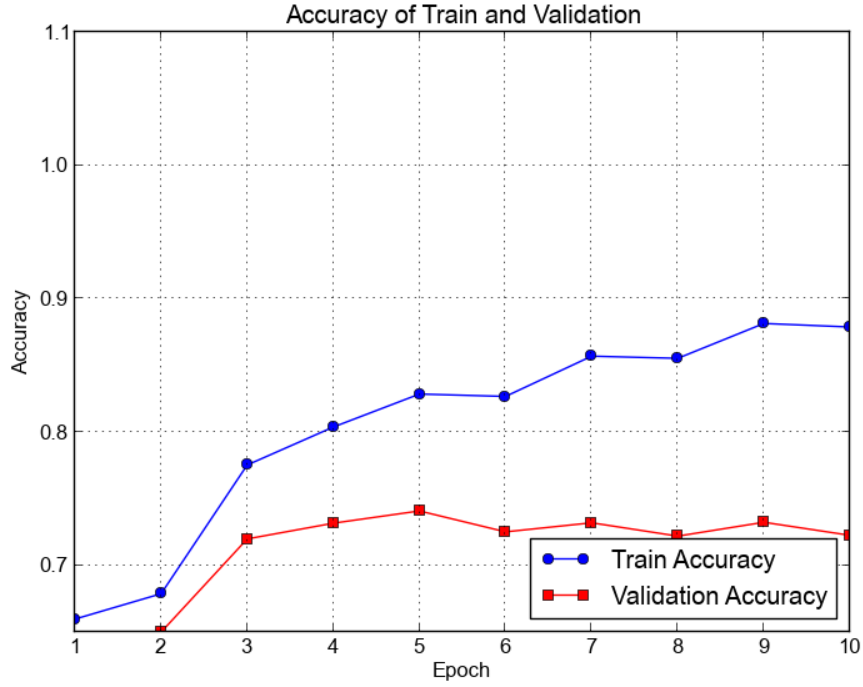


Figure 8: Accuracy of Train and Validation Removing One Full-connected Layer

epoch	train loss	test loss	train accuracy	test accuracy
1	1.619 09	1.638 49	46.800%	46.410%
2	0.866 42	0.927 77	69.500%	67.330%
3	0.753 25	0.838 97	73.508%	70.650%
4	0.632 48	0.768 28	77.874%	73.560%
5	0.507 96	0.687 51	82.418%	76.480%
6	0.502 83	0.722 61	82.610%	75.600%
7	0.345 76	0.630 78	88.022%	79.090%
8	0.290 08	0.624 89	90.142%	79.830%
9	0.210 73	0.612 90	92.906%	80.670%
10	0.166 42	0.648 21	94.572%	80.310%
11	0.123 89	0.626 89	96.150%	80.650%
12	0.108 81	0.676 80	96.520%	80.350%
13	0.093 55	0.717 92	96.872%	80.140%
14	0.084 26	0.762 88	97.210%	80.000%
15	0.095 36	0.817 63	96.652%	79.540%
16	0.065 47	0.841 22	97.700%	80.090%
17	0.052 65	0.852 79	98.308%	80.120%
18	0.019 34	0.780 01	99.500%	81.620%
19	0.026 99	0.832 29	99.198%	81.050%
20	0.032 54	0.860 33	98.942%	80.520%

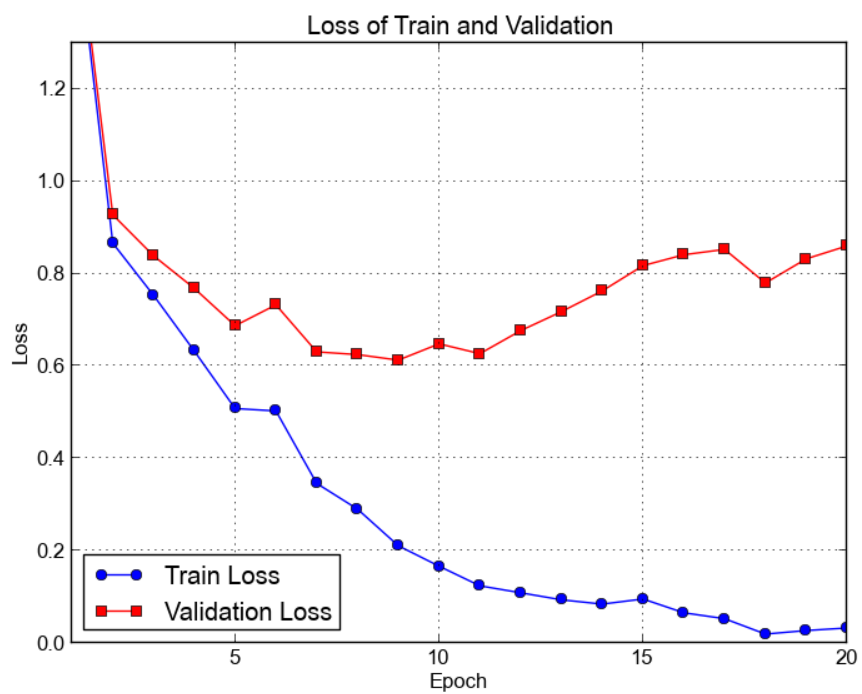


Figure 9: Loss of Train and Validation Removing One Full-connected Layer

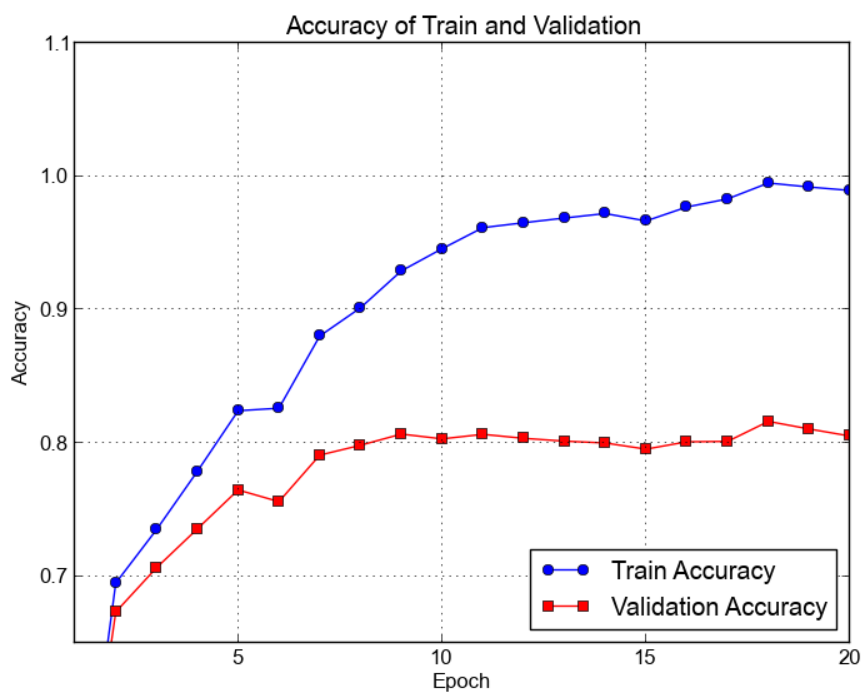


Figure 10: Accuracy of Train and Validation Removing One Full-connected Layer