HW3

2. MyMLP

Test Set Loss/Error Rate for *Adam* with Epoch 5					
Learning Rate	$1e-5$	$1e-4$	$1e-3$	$1e-2$	$1e-1$
Loss	498.6	476	468.9	517.4	734.8
Error Rate	0.09	0.056	0.036	0.19	0.88

Test Set Loss/Error Rate for *SGD* with Epoch 5					
Learning Rate	$1e-5$	$1e-4$	$1e-3$	$1e-2$	$1e-1$
Loss	711.68	559.9	485.6	469.8	476.8
Error Rate	0.6676	0.27	0.08	0.034	0.062

Test Set Loss/Error Rate for *Adagrad* with Epoch 5					
Learning Rate	$1e-5$	$1e-4$	$1e-3$	$1e-2$	$1e-1$
Loss	709	576.1	489.7	471.3	738.7
Error Rate	0.61	0.27	0.08	0.39	0.89

Test Set Loss/Error Rate for *RMSprop* with Epoch 5					
Learning Rate	$1e-5$	$1e-4$	$1e-3$	$1e-2$	$1e-1$
Loss	476.5	467.68	677.81	734.8	739.59
Error Rate	0.05	0.03	0.70	0.88	0.90

3. MyCNN

Test Set Loss/Error Rate for *SGD* with Epoch 10	
Learning Rate	$1e-2$
Loss	463.3
Error Rate	0.018
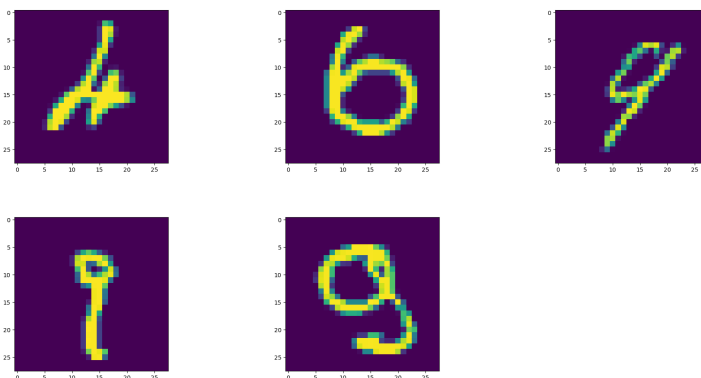
- Image 1 label is 4, predict is 2
- Image 2 label is 6, predict is 0
- Image 3 label is 9, predict is 4
- Image 4 label is 9, predict is 7
- Image 5 label is 9, predict is 8

In my opinion, the first reason the dataset is not big enough, if there are similar image in the training dataset, the accuracy in testing will be higher. Second, those image looks like in the middle of the true label and predict label, for example, image 3 looks like 9 nor 4 neither, it is like the shape between 9 and 4.