

Name: Junzhe Wu NetID: junzhew3

1. Test accuracy: 0.9127
2. I use one convolutional layer and one fully connected layer. The kernel size is 4 and the number of channel is 3.
3. The learning rate is 0.01 in the 1st and 2nd epoch, then become 0.005 in the 3rd epoch. The number of epoch is 3, and the batch size is 100.
4. The K and w1 matrix are randomly initialized with a normalization constant 0.01. The b1 matrix are initialized to 0.
5. I use the minibatch stochastic gradient descent method and the batch size is 100. I shuffle the data each time.