HW3

The final test accuracy achieved is 80%.

Model Architecture

- Convolution layer 1: 64 channels, k=4, s=1, P=2.
- Convolution layer 2: 64 channels, k=4, s=1, P=2.
- Dropout: p=0.1.
- Convolution layer 3: 64 channels, k=4, s=1, P=2.
- Batch Normalization
- Convolution layer 4: 64 channels, k=4, s=1, P=2.
- Max Pooling: s=2, k=2.
- Dropout: p=0.1.
- Convolution layer 5: 64 channels, k=4, s=1, P=2.
- Batch Normalization
- Convolution layer 6: 64 channels, k=3, s=1, P=0.
- Dropout: p=0.1.
- Convolution layer 7: 64 channels, k=3, s=1, P=0.
- Batch Normalization
- Convolution layer 8: 64 channels, k=3, s=1, P=0.
- Batch Normalization
- Dropout: p=0.1.
- Fully connected layer 1: 500 units
- Fully connected layer 2: 500 units
- Linear -> Softmax function

In the Mini-Batch Stochastic Gradient Descent, I used batch size = 100. Also, I performed data augmentation by random flipping the images horizontally with p = 0.2 and random rotate the data by -8 to +8 degree.