Xin Yang
xy213

HW4
Result without dropout/batch normalization:

```
Train Epoch: 50 [39680/50000 (79%)] Loss: 1.249198
Train Epoch: 50 [40960/50000 (82%)] Loss: 1.158152
Train Epoch: 50 [42240/50000 (84%)] Loss: 1.234049
Train Epoch: 50 [43520/50000 (87%)] Loss: 1.141796
Train Epoch: 50 [44800/50000 (90%)] Loss: 1.349997
Train Epoch: 50 [46080/50000 (92%)] Loss: 1.387434
Train Epoch: 50 [47360/50000 (95%)] Loss: 1.241622
Train Epoch: 50 [48640/50000 (97%)] Loss: 1.399387
Train Epoch: 50 [31200/50000 (100%)]    Loss: 0.906030

Test set: Average loss: 1.0988, Accuracy: 6201/10000 (62%)

Traning and Testing total excution time is: 425.76764035224915 seconds

Process finished with exit code 0
```

Result with one additional dropout:

```
Train Epoch: 50 [35840/50000 (72%)] Loss: 1.291571
Train Epoch: 50 [37120/50000 (74%)] Loss: 1.654281
Train Epoch: 50 [38400/50000 (77%)] Loss: 1.324424
Train Epoch: 50 [39680/50000 (79%)] Loss: 1.487594
Train Epoch: 50 [40960/50000 (82%)] Loss: 1.530554
Train Epoch: 50 [42240/50000 (84%)] Loss: 1.528687
Train Epoch: 50 [43520/50000 (87%)] Loss: 1.418408
Train Epoch: 50 [44800/50000 (90%)] Loss: 1.309369
Train Epoch: 50 [46080/50000 (92%)] Loss: 1.412139
Train Epoch: 50 [47360/50000 (95%)] Loss: 1.343901
Train Epoch: 50 [48640/50000 (97%)] Loss: 1.486733
Train Epoch: 50 [31200/50000 (100%)]    Loss: 1.412356

Test set: Average loss: 1.3147, Accuracy: 5647/10000 (56%)

Traning and Testing total excution time is: 433.37776708602905 seconds
```

Result with one additional batch normalization:

```
Train Epoch: 50 [35840/50000 (72%)] Loss: 1.171893
Train Epoch: 50 [37120/50000 (74%)] Loss: 1.244684
Train Epoch: 50 [38400/50000 (77%)] Loss: 1.031173
Train Epoch: 50 [39680/50000 (79%)] Loss: 0.903002
Train Epoch: 50 [40960/50000 (82%)] Loss: 1.214403
Train Epoch: 50 [42240/50000 (84%)] Loss: 1.305582
Train Epoch: 50 [43520/50000 (87%)] Loss: 1.223900
Train Epoch: 50 [44800/50000 (90%)] Loss: 1.300613
Train Epoch: 50 [46080/50000 (92%)] Loss: 1.111347
Train Epoch: 50 [47360/50000 (95%)] Loss: 1.039284
Train Epoch: 50 [48640/50000 (97%)] Loss: 1.395892
Train Epoch: 50 [31200/50000 (100%)]    Loss: 1.258286

Test set: Average loss: 1.0632, Accuracy: 6420/10000 (64%)

Traning and Testing total excution time is: 422.8482120037079 seconds
```

Comparison:

The result of CNN with one dropout layer has the lowest test accuracy of 56%. One possible reason is the dropout layer randomly discards half of the weights, which relieves the overfitting issue, but decreases the accuracy.

The CNN with one batch normalization layer shows a highest accuracy of 64%. The batch normalization is performed before every iteration, which distributes the data in a more uniform way to accelerate the convergence.

The performance without dropout/batch normalization layer shows a 62% accuracy, which is between the one with dropout layer and the one with batch normalization layer.