This is the work for the locally connected with no weights shared in the first three layers

Task 1

Our locally connected network is designed to classify handwritten digits. Instead of using the standard convolutional layers with shared weights, our network uses locally connected layers where each spatial location has its own unique filter. This allows the network to learn very specific features from different parts of the image.

The network consists of three locally connected layers. With a 16×16 input image, the first layer reduces the size to 14×14, the second to 12×12, and the third to 10×10. After each layer, we apply batch normalization to stabilize the training and use the ReLU activation function to introduce non-linearity. In one of the layers, we also use the tanh function to provide a bounded output.

Once the features have been extracted by these three layers, the output is flattened into a one-dimensional vector and passed through a dropout layer to reduce overfitting. Finally, a fully connected layer maps the 3200 features to 10 classes, corresponding to the digits 0 through 9.

This design effectively captures local patterns in the data and, together with batch normalization and dropout, leads to stable and generalizable learning.

Task 2

(1) Parameter initialization strategies

(a) Effective Learning

```
Epoch 1/10 - Loss: 0.4660, Acc: 85.24%
Epoch 2/10 - Loss: 0.2195, Acc: 93.50%
Epoch 3/10 - Loss: 0.1652, Acc: 94.95%
Epoch 4/10 - Loss: 0.1466, Acc: 95.47%
Epoch 5/10 - Loss: 0.1190, Acc: 96.12%
Epoch 6/10 - Loss: 0.1124, Acc: 96.15%
Epoch 7/10 - Loss: 0.1094, Acc: 96.45%
Epoch 8/10 - Loss: 0.0953, Acc: 96.91%
Epoch 9/10 - Loss: 0.0932, Acc: 96.68%
Epoch 10/10 - Loss: 0.0868, Acc: 97.22%

Test Loss: 0.2893, Test Accuracy: 93.52%

Process finished with exit code 0
```

(b) Fast Learning

```
Epoch 1/10 - Loss: 0.4756, Acc: 85.31%
Epoch 2/10 - Loss: 0.2024, Acc: 93.95%
Epoch 3/10 - Loss: 0.1611, Acc: 95.16%
Epoch 4/10 - Loss: 0.1370, Acc: 95.71%
Epoch 5/10 - Loss: 0.1206, Acc: 95.97%
Epoch 6/10 - Loss: 0.1018, Acc: 96.83%
Epoch 7/10 - Loss: 0.0840, Acc: 97.17%
Epoch 8/10 - Loss: 0.0881, Acc: 96.80%
Epoch 9/10 - Loss: 0.0753, Acc: 97.26%
Epoch 10/10 - Loss: 0.0767, Acc: 97.49%

Test Loss: 0.3210, Test Accuracy: 93.02%
```

(c) Slow Learning

```
Epoch 1/10 - Loss: 0.4654, Acc: 86.04%
Epoch 2/10 - Loss: 0.2088, Acc: 93.76%
Epoch 3/10 - Loss: 0.1573, Acc: 95.35%
Epoch 4/10 - Loss: 0.1329, Acc: 95.76%
Epoch 5/10 - Loss: 0.1169, Acc: 96.45%
Epoch 6/10 - Loss: 0.1114, Acc: 96.43%
Epoch 7/10 - Loss: 0.1018, Acc: 96.49%
Epoch 8/10 - Loss: 0.0893, Acc: 97.05%
Epoch 9/10 - Loss: 0.0841, Acc: 97.26%
Epoch 10/10 - Loss: 0.0722, Acc: 97.59%

Test Loss: 0.2767, Test Accuracy: 93.57%

Process finished with exit code 0
```

Comment: All three settings converge to roughly the same accuracy, but the effective setup reaches that point more quickly and steadily. The fast setting can converge quickly as well but risks instability, while the slow setting is more stable but takes longer to reach the same performance. Overall, the effective learning rate balances speed and stability best.

(2) Learning rate

(a) Effective Learning

```
Epoch 1/10 - Loss: 0.4644, Acc: 85.61%
Epoch 2/10 - Loss: 0.2161, Acc: 93.75%
Epoch 3/10 - Loss: 0.1564, Acc: 95.21%
Epoch 4/10 - Loss: 0.1350, Acc: 95.62%
Epoch 5/10 - Loss: 0.1148, Acc: 96.24%
Epoch 6/10 - Loss: 0.1036, Acc: 96.53%
Epoch 7/10 - Loss: 0.1011, Acc: 96.71%
Epoch 8/10 - Loss: 0.0860, Acc: 97.05%
Epoch 9/10 - Loss: 0.0876, Acc: 96.86%
Epoch 10/10 - Loss: 0.0793, Acc: 97.16%

Test Loss: 0.2889, Test Accuracy: 93.87%
```

(b) Slow Learning

```
Epoch 1/10 - Loss: 1.2361, Acc: 65.33%
Epoch 2/10 - Loss: 0.5878, Acc: 84.80%
Epoch 3/10 - Loss: 0.4334, Acc: 88.31%
Epoch 4/10 - Loss: 0.3583, Acc: 90.19%
Epoch 5/10 - Loss: 0.3109, Acc: 91.29%
Epoch 6/10 - Loss: 0.2772, Acc: 92.32%
Epoch 7/10 - Loss: 0.2527, Acc: 93.17%
Epoch 8/10 - Loss: 0.2358, Acc: 93.18%
Epoch 9/10 - Loss: 0.2138, Acc: 94.09%
Epoch 10/10 - Loss: 0.2035, Acc: 94.28%

Test Loss: 0.3124, Test Accuracy: 91.23%
```

© Fast Learning

```
Epoch 1/10 - Loss: 1.4011, Acc: 69.26%
Epoch 2/10 - Loss: 0.8242, Acc: 73.86%
Epoch 3/10 - Loss: 0.7434, Acc: 76.78%
Epoch 4/10 - Loss: 0.7527, Acc: 77.04%
Epoch 5/10 - Loss: 0.7005, Acc: 77.89%
Epoch 6/10 - Loss: 0.6729, Acc: 79.37%
Epoch 7/10 - Loss: 0.7020, Acc: 78.07%
Epoch 8/10 - Loss: 0.7440, Acc: 77.74%
Epoch 9/10 - Loss: 0.7137, Acc: 77.78%
Epoch 10/10 - Loss: 0.7262, Acc: 77.67%

Test Loss: 0.5258, Test Accuracy: 85.05%
```

Comment: All three runs ultimately converge to respectable accuracies, but the effective setup does so more reliably and with higher final accuracy. The slow run eventually catches up but takes longer, while the fast run struggles with volatile updates and ends up with notably lower test accuracy. This confirms that a balanced learning rate tends to achieve both speed and stability in training.

(3) Momentum

```
Training model 1 of 3...
Epoch 1/10 - Loss: 0.5498, Acc: 82.84%
Epoch 2/10 - Loss: 0.2408, Acc: 92.96%
Epoch 3/10 - Loss: 0.1764, Acc: 94.68%
Epoch 4/10 - Loss: 0.1528, Acc: 95.49%
Epoch 5/10 - Loss: 0.1249, Acc: 96.15%
Epoch 6/10 - Loss: 0.1104, Acc: 96.69%
Epoch 7/10 - Loss: 0.1018, Acc: 96.53%
Epoch 8/10 - Loss: 0.0945, Acc: 97.00%
Epoch 9/10 - Loss: 0.0928, Acc: 96.83%
Epoch 10/10 - Loss: 0.0785, Acc: 97.46%

Training model 2 of 3...
Epoch 1/10 - Loss: 0.5635, Acc: 83.17%
Epoch 2/10 - Loss: 0.2355, Acc: 92.85%
Epoch 3/10 - Loss: 0.1883, Acc: 94.50%
Epoch 4/10 - Loss: 0.1455, Acc: 95.78%
Epoch 5/10 - Loss: 0.1262, Acc: 96.34%
Epoch 6/10 - Loss: 0.1154, Acc: 96.50%
Epoch 7/10 - Loss: 0.0970, Acc: 96.87%
Epoch 8/10 - Loss: 0.0914, Acc: 97.00%
Epoch 9/10 - Loss: 0.0777, Acc: 97.42%
Epoch 10/10 - Loss: 0.0885, Acc: 97.05%

Training model 3 of 3...
Epoch 1/10 - Loss: 0.5327, Acc: 83.87%
Epoch 2/10 - Loss: 0.2416, Acc: 92.81%
Epoch 3/10 - Loss: 0.1760, Acc: 94.86%
Epoch 4/10 - Loss: 0.1484, Acc: 95.53%
Epoch 5/10 - Loss: 0.1203, Acc: 96.52%
Epoch 6/10 - Loss: 0.1099, Acc: 96.68%
Epoch 7/10 - Loss: 0.1001, Acc: 96.79%
Epoch 8/10 - Loss: 0.0887, Acc: 97.22%
Epoch 9/10 - Loss: 0.0800, Acc: 97.50%
Epoch 10/10 - Loss: 0.0830, Acc: 97.28%

Evaluating one model...
Single Model - Test Loss: 0.2576, Test Accuracy: 93.62%
```

Comment: All three networks performed well on their own, but when we combined their predictions (the ensemble), the test accuracy went up from about 93.6% to 94.8%. This shows that averaging the outputs of several models can smooth out their individual errors, resulting in a better overall performance.