Neural Networks and Deep Learning - Assignment 3-4

Neta Ben Mordechai, ID: 211451901.
Itay Chabra, ID: 208292458.

Project Overview

This project involves training and evaluating multiple neural network architectures for digit classification using the MNIST dataset. The goal is to compare the performance of different architectures in terms of accuracy, training time, and balanced accuracy, while also visualizing key aspects of the best-performing model's convolutional layers.

Dataset - The MNIST dataset contains 60,000 training images and 10,000 test images of handwritten digits (0–9). The training data is split into 90% for training and 10% for validation, ensuring proper evaluation of the models.

Output

Architecture 1 – Logistic regression with no hidden layers:

```
Training Logistic Regression (No Hidden Layers)...
Epoch 1/11, Train Loss: 2.2288, Val Loss: 2.1578, Val Accuracy: 0.6538
Epoch 2/11, Train Loss: 2.0882, Val Loss: 2.0236, Val Accuracy: 0.6977
Epoch 3/11, Train Loss: 1.9581, Val Loss: 1.8990, Val Accuracy: 0.7205
Epoch 4/11, Train Loss: 1.8378, Val Loss: 1.7840, Val Accuracy: 0.7423
Epoch 5/11, Train Loss: 1.7268, Val Loss: 1.6780, Val Accuracy: 0.7530
Epoch 6/11, Train Loss: 1.6246, Val Loss: 1.5803, Val Accuracy: 0.7667
Epoch 7/11, Train Loss: 1.5308, Val Loss: 1.4908, Val Accuracy: 0.7752
Epoch 8/11, Train Loss: 1.4450, Val Loss: 1.4089, Val Accuracy: 0.7833
Epoch 9/11, Train Loss: 1.3667, Val Loss: 1.3343, Val Accuracy: 0.7908
Epoch 10/11, Train Loss: 1.2953, Val Loss: 1.2661, Val Accuracy: 0.7968
Epoch 11/11, Train Loss: 1.2301, Val Loss: 1.2039, Val Accuracy: 0.8058

Training Time: 119.31 seconds

Evaluating Logistic Regression (No Hidden Layers)...

Logistic Regression (No Hidden Layers) Metrics:
Precision: [0.89073171 0.71913782 0.90293454 0.73288815 0.83402062 0.92720307
 0.84920635 0.83118081 0.83995187 0.82116402]

Recall: [0.93163265 0.97004405 0.7751938  0.86930693 0.82382892 0.5426009
 0.89352818 0.87645914 0.71663244 0.7690783 ]

F1_score: [0.91072319 0.82595649 0.83420229 0.79528986 0.82889344 0.68458274
 0.87080366 0.8532197  0.7734072  0.79426817]

Accuracy: 0.8216

Balanced_accuracy: 0.8168305319057291

Logistic Regression (No Hidden Layers) Parameters: 7850
```

Architecture 2 – Logistic regression with two hidden layers:

```
Training Logistic Regression with Two Hidden Layers...
Epoch 1/11, Train Loss: 0.9728, Val Loss: 0.5809, Val Accuracy: 0.8212
Epoch 2/11, Train Loss: 0.4832, Val Loss: 0.4226, Val Accuracy: 0.8757
Epoch 3/11, Train Loss: 0.3844, Val Loss: 0.3601, Val Accuracy: 0.8892
Epoch 4/11, Train Loss: 0.3325, Val Loss: 0.3242, Val Accuracy: 0.9018
Epoch 5/11, Train Loss: 0.2951, Val Loss: 0.2979, Val Accuracy: 0.9077
Epoch 6/11, Train Loss: 0.2644, Val Loss: 0.2675, Val Accuracy: 0.9182
Epoch 7/11, Train Loss: 0.2367, Val Loss: 0.2431, Val Accuracy: 0.9248
Epoch 8/11, Train Loss: 0.2109, Val Loss: 0.2237, Val Accuracy: 0.9325
Epoch 9/11, Train Loss: 0.1888, Val Loss: 0.2038, Val Accuracy: 0.9352
Epoch 10/11, Train Loss: 0.1695, Val Loss: 0.1873, Val Accuracy: 0.9412
Epoch 11/11, Train Loss: 0.1542, Val Loss: 0.1715, Val Accuracy: 0.9467

Training Time: 126.23 seconds

Evaluating Logistic Regression with Two Hidden Layers...

Logistic Regression with Two Hidden Layers Metrics:
Precision: [0.95738355 0.9771529  0.9460501  0.95665323 0.93675889 0.94659091
 0.94964029 0.96693387 0.92672859 0.94954591]

Recall: [0.98571429 0.97973568 0.95155039 0.93960396 0.96537678 0.9338565
 0.96450939 0.93871595 0.92197125 0.93260654]

F1_score: [0.97134238 0.97844259 0.94879227 0.94805195 0.95085256 0.94018059
 0.95701709 0.95261599 0.9243438  0.941     ]

Accuracy: 0.9518

Balanced_accuracy: 0.9513640742422138

Logistic Regression with Two Hidden Layers Parameters: 199210
```

Architecture 3 – Convolutional Neural Network (CNN) with one convolutional layer:

```
Training CNN (1 Convolutional Layer)...
Epoch 1/11, Train Loss: 0.6026, Val Loss: 0.2835, Val Accuracy: 0.9153
Epoch 2/11, Train Loss: 0.2289, Val Loss: 0.2084, Val Accuracy: 0.9330
Epoch 3/11, Train Loss: 0.1605, Val Loss: 0.1599, Val Accuracy: 0.9523
Epoch 4/11, Train Loss: 0.1270, Val Loss: 0.1448, Val Accuracy: 0.9543
Epoch 5/11, Train Loss: 0.1068, Val Loss: 0.1293, Val Accuracy: 0.9627
Epoch 6/11, Train Loss: 0.0922, Val Loss: 0.1095, Val Accuracy: 0.9683
Epoch 7/11, Train Loss: 0.0814, Val Loss: 0.1024, Val Accuracy: 0.9708
Epoch 8/11, Train Loss: 0.0718, Val Loss: 0.0975, Val Accuracy: 0.9703
Epoch 9/11, Train Loss: 0.0643, Val Loss: 0.0851, Val Accuracy: 0.9748
Epoch 10/11, Train Loss: 0.0558, Val Loss: 0.0903, Val Accuracy: 0.9745
Epoch 11/11, Train Loss: 0.0517, Val Loss: 0.0880, Val Accuracy: 0.9752

Training Time: 181.15 seconds

Evaluating CNN (1 Convolutional Layer)...

CNN (1 Convolutional Layer) Metrics:
Precision: [0.96822244 0.98939929 0.97016362 0.97219464 0.9825998  0.9765625
 0.98732841 0.98880977 0.94482759 0.96489468]

Recall: [0.99489796 0.98678414 0.97674419 0.96930693 0.97759674 0.9809417
 0.97599165 0.94552529 0.98459959 0.95341923]

F1_score: [0.98137896 0.98808999 0.97344278 0.97074864 0.98009188 0.9787472
 0.9816273  0.96668324 0.96430367 0.95912263]

Accuracy: 0.9745

Balanced_accuracy: 0.9745807419650356

CNN (1 Convolutional Layer) Parameters: 6434634
```

Architecture 4 – Convolutional Neural Network (CNN) with two convolutional layers:

```
Training CNN (2 Convolutional Layer)...
Epoch 1/11, Train Loss: 0.6029, Val Loss: 0.2636, Val Accuracy: 0.9210
Epoch 2/11, Train Loss: 0.1898, Val Loss: 0.1540, Val Accuracy: 0.9513
Epoch 3/11, Train Loss: 0.1098, Val Loss: 0.1031, Val Accuracy: 0.9675
Epoch 4/11, Train Loss: 0.0770, Val Loss: 0.0737, Val Accuracy: 0.9758
Epoch 5/11, Train Loss: 0.0605, Val Loss: 0.0668, Val Accuracy: 0.9797
Epoch 6/11, Train Loss: 0.0491, Val Loss: 0.0536, Val Accuracy: 0.9837
Epoch 7/11, Train Loss: 0.0416, Val Loss: 0.0567, Val Accuracy: 0.9825
Epoch 8/11, Train Loss: 0.0344, Val Loss: 0.0541, Val Accuracy: 0.9848
Epoch 9/11, Train Loss: 0.0296, Val Loss: 0.0478, Val Accuracy: 0.9857
Epoch 10/11, Train Loss: 0.0232, Val Loss: 0.0558, Val Accuracy: 0.9853
Epoch 11/11, Train Loss: 0.0210, Val Loss: 0.0531, Val Accuracy: 0.9862

Training Time: 161.74 seconds

Evaluating CNN (2 Convolutional Layer)...

CNN (2 Convolutional Layer) Metrics:
Precision: [0.99485597 0.99559471 0.98557692 0.98716683 0.9877551  0.98106904
 0.99681866 0.99021526 0.98559671 0.97365854]

Recall: [0.98673469 0.99559471 0.99321705 0.99009901 0.98574338 0.98766816
 0.98121086 0.9844358  0.9835729  0.98909812]

F1_score: [0.99077869 0.99559471 0.98938224 0.98863075 0.98674822 0.98435754
 0.98895318 0.98731707 0.98458376 0.9813176 ]

Accuracy: 0.9879

Balanced_accuracy: 0.9877374679828815
```

Architecture 5 – Convolutional Neural Network (CNN) with two convolutional layers and dropout, comparing 50 and 100 minibaches:

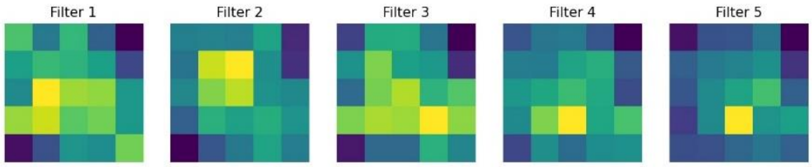
```
Training model_cnn3 with batch size 50...
Epoch 1, Train Loss: 578.6115, Val Accuracy: 0.9520
Epoch 2, Train Loss: 174.1140, Val Accuracy: 0.9715
Epoch 3, Train Loss: 122.5872, Val Accuracy: 0.9780
Epoch 4, Train Loss: 101.0562, Val Accuracy: 0.9825
Epoch 5, Train Loss: 82.4375, Val Accuracy: 0.9832
Epoch 6, Train Loss: 75.5545, Val Accuracy: 0.9817
Epoch 7, Train Loss: 65.4455, Val Accuracy: 0.9862
Epoch 8, Train Loss: 60.2610, Val Accuracy: 0.9872
Epoch 9, Train Loss: 53.6765, Val Accuracy: 0.9865
Epoch 10, Train Loss: 48.4685, Val Accuracy: 0.9883
Epoch 11, Train Loss: 44.6942, Val Accuracy: 0.9883
Epoch 12, Train Loss: 42.4386, Val Accuracy: 0.9895
Epoch 13, Train Loss: 40.9457, Val Accuracy: 0.9890
Epoch 14, Train Loss: 36.0553, Val Accuracy: 0.9897
Epoch 15, Train Loss: 34.6560, Val Accuracy: 0.9902
Reached 0.99 validation accuracy in 15 epochs.
Training Time: 210.48 seconds

Evaluating CNN (2 Convolutional Layer 50 batch)...

CNN (2 Convolutional Layer 50 batch) Metrics:
Precision: [0.99285714 0.99298246 0.9941691  0.99501496 0.99287169 0.98333333
 0.98752599 0.98743961 0.98965874 0.98902196]

Recall: [0.99285714 0.99735683 0.99127907 0.98811881 0.99287169 0.99215247
 0.99164927 0.99416342 0.9825462  0.98216056]
```

```
F1_score: [0.99285714 0.99516484 0.99272198 0.99155489 0.99287169 0.98772321
 0.98958333 0.99079011 0.98608964 0.98557931]

Accuracy: 0.9906

Balanced_accuracy: 0.9905155459167585

CNN (2 Convolutional Layer 50 batch) Parameters: 3274634


Training model_cnn3 with batch size 100...
Epoch 1, Train Loss: 384.6546, Val Accuracy: 0.9363
Epoch 2, Train Loss: 110.8679, Val Accuracy: 0.9617
Epoch 3, Train Loss: 75.6896, Val Accuracy: 0.9742
Epoch 4, Train Loss: 59.0573, Val Accuracy: 0.9778
Epoch 5, Train Loss: 49.9374, Val Accuracy: 0.9770
Epoch 6, Train Loss: 44.1429, Val Accuracy: 0.9837
Epoch 7, Train Loss: 39.7641, Val Accuracy: 0.9812
Epoch 8, Train Loss: 36.0568, Val Accuracy: 0.9850
Epoch 9, Train Loss: 31.3357, Val Accuracy: 0.9870
Epoch 10, Train Loss: 30.6224, Val Accuracy: 0.9875
Epoch 11, Train Loss: 26.8017, Val Accuracy: 0.9875
Epoch 12, Train Loss: 25.6927, Val Accuracy: 0.9873
Epoch 13, Train Loss: 23.6982, Val Accuracy: 0.9893
Epoch 14, Train Loss: 22.2994, Val Accuracy: 0.9895
Epoch 15, Train Loss: 20.9722, Val Accuracy: 0.9908
Reached 0.99 validation accuracy in 15 epochs.
Training Time: 171.24 seconds

Evaluating CNN (2 Convolutional Layer 100 batch)...

CNN (2 Convolutional Layer 100 batch) Metrics:
Precision: [0.98685541 0.99646331 0.99415205 0.97667638 0.98888889 0.9877095
 0.9957492  0.99022483 0.98556701 0.9860835 ]

Recall: [0.99591837 0.99295154 0.98837209 0.9950495  0.99694501 0.99103139
 0.97807933 0.98540856 0.98151951 0.98315164]

F1_score: [0.99136618 0.99470432 0.99125364 0.98577734 0.99290061 0.98936766
 0.98683518 0.98781082 0.98353909 0.98461538]

Accuracy: 0.9889

Balanced_accuracy: 0.9888426942210883

CNN (2 Convolutional Layer 100 batch) Parameters: 3274634
```

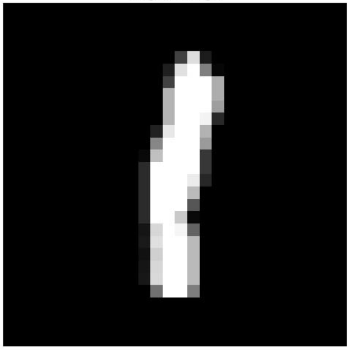
Visualizations for convolutional layers with best Balanced accuracy

As we can see from the output, the best-balanced accuracy CNN is the one with two convolutional layers and batch size 50.

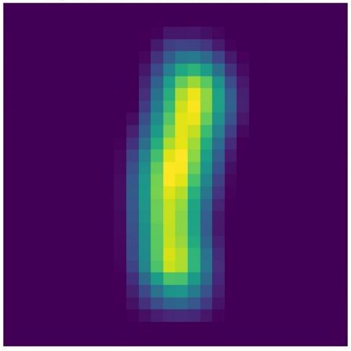
5 different filters visualizations:



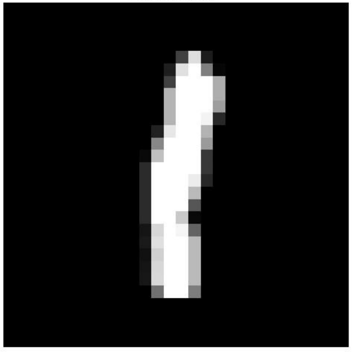
4 Convolutional images – before and after RELU and original image:



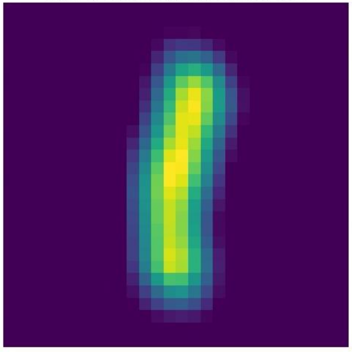
Original Image



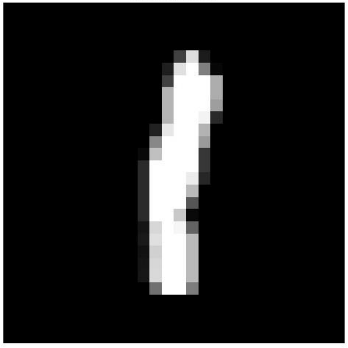
Layer 0, Channel 0, Activation After: False



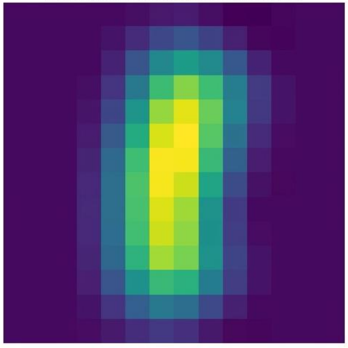
Original Image



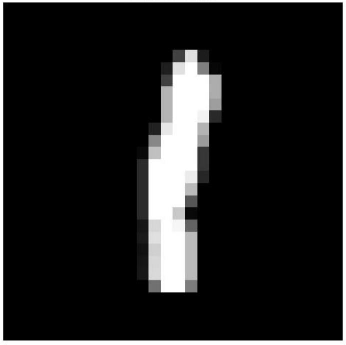
Layer 1, Channel 0, Activation After: True



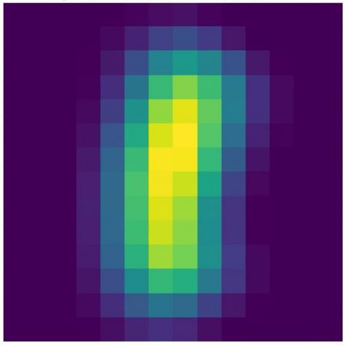
Original Image



Layer 3, Channel 0, Activation After: False



Original Image



Layer 4, Channel 0, Activation After: True