Implementing A Fully Connected Neural Network on Modified MNIST Datasets

Blaine Swieder
Department of Electrical Engineering and Computer Science
Tickle College of Engineering
University of Tennessee, Knoxville
Knoxville, Tennessee 37996
Email: bswieder@vols.utk.edu

***Abstract*—This project evaluates the performance of a fully connected neural network on four variants of the MNIST dataset, including a small-data regime, pixel distribution modification, and a novel 3-digit classification task. We use PyTorch to train and evaluate the model across all datasets. Our goal is to measure generalization ability while considering model complexity.**

I. Introduction

The MNIST dataset is a classic benchmark for handwritten digit classification. In this project, we implement a fully connected neural network (FCNN) without any convolutional layers to test its performance on four modified versions of MNIST. Each variation presents unique challenges including limited data, altered brightness, and increased output complexity.

II. Experimental Setup

A. Model Architecture

The FCNN includes 2–3 hidden layers with ReLU activations and softmax output. The architecture is defined dynamically depending on the dataset. Weight initialization used Xavier or He methods. The Adam optimizer was used throughout.

B. Metrics

We tracked loss, training accuracy, testing accuracy, and calculated a normalized score:

$$\text{Score Ratio} = \frac{\text{Test Accuracy}}{\sqrt[10]{\text{Total Parameters}}} \tag{1}$$

III. Results

IV. Dataset Analysis

A. Dataset 1: Standard MNIST (50K samples)

1) Performance: The Standard MNIST dataset (50K samples) had a 98.08% test accuracy.

2) Discussion: The model achieved 99.93% training accuracy and 98.08% testing accuracy with a smooth loss curve and minimal generalization gap. No regularization was needed. This dataset demonstrated the ease with which a fully connected network can fit MNIST with sufficient data. See the full-size result in Figure 1.

B. Dataset 2: Low-Data Regime (1K samples)

1) Performance: The Low-Data Regime (1,000 samples) had 88.88% test accuracy.

2) Discussion: This low-data scenario caused some overfitting, though the network still achieved 88.88% test accuracy. The model generalized reasonably well, aided by strong initial convergence and stable optimization. The loss curve showed sharp improvement in early epochs, and test accuracy stabilized after epoch 10. This dataset highlighted the FCNN's ability to extract meaningful features even with minimal data. See Figure 2 in the appendix for the complete output.

C. Dataset 3: Brightened Input (Modified Pixel Values)

1) Performance: The Brightened Input (Modified Pixel Values) dataset had 90.99% test accuracy.

2) Discussion: Despite the altered pixel distributions, the model converged cleanly to 90.99% test accuracy. This shows the robustness of the architecture to changes in data scale. The network quickly adapted to the new brightness pattern, and training remained stable. Final accuracy hovered just under 91%, with minimal fluctuations in validation performance. See Figure 3

D. Dataset 4: Concatenated 3-Digit Classification (1,000 Classes)

1) Performance: The Concatenated 3-Digit Classification (1,000 Classes) dataset had 78.38% test accuracy.

2) Discussion: This dataset introduced the most complexity, with inputs formed by vertically stacking three digits and output labels ranging from 0 to 999. The network architecture was scaled accordingly, with three hidden layers and over 1.4 million parameters. Despite the challenge, the model reached 96.77% training accuracy and 78.38% on the test set. Loss and accuracy curves showed stable convergence, though some overfitting was present. The relatively strong test accuracy confirms the model's capacity to learn high-dimensional representations with enough training data and depth. See Figure 4 for the full output plot.

V. Conclusion

This project explored the ability of fully connected networks to adapt across dataset shifts, limited training data, and more

TABLE I: Summary of Experiment Results Across Datasets

Dataset	Layers	Hidden Units	Params	Train Acc	Test Acc	Score
Dataset 1	2	[256,128]	235K	99.93%	98.08%	28.47
Dataset 2	2	[256,128]	235K	100.0%	88.88%	25.80
Dataset 3	2	[256,128]	235K	92.03%	90.99%	26.41
Dataset 4	3	[512,256,128]	1.50M	96.77%	78.38%	18.91

complex classification tasks. Despite the lack of convolutional layers, the model performed surprisingly well. The results demonstrate that with careful tuning, FCNNs can achieve strong performance even on nontrivial variations of MNIST.

Appendix: Full-Size Result Figures

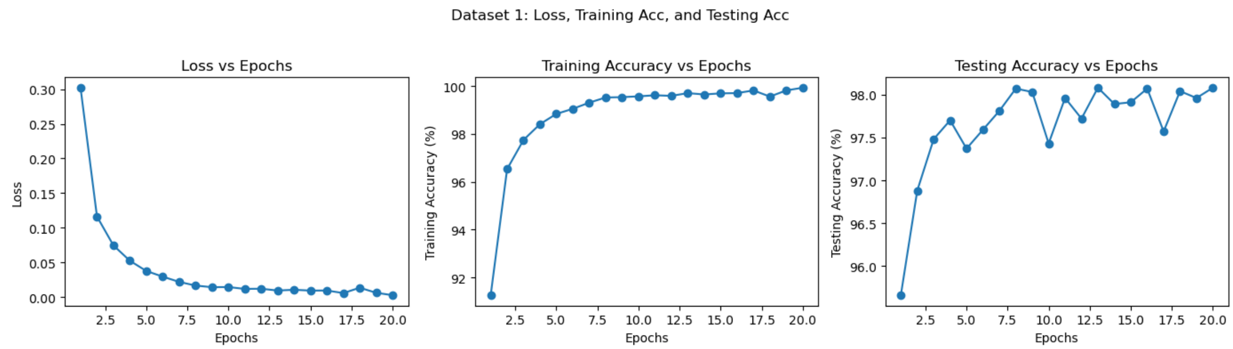

Fig. 1: Dataset 1 – Final Output (full size)

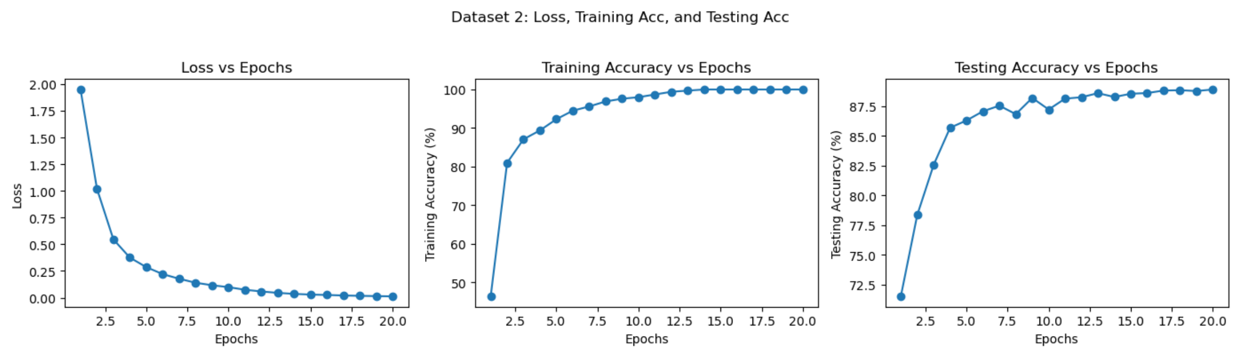

Fig. 2: Dataset 2 – Final Output (full size)

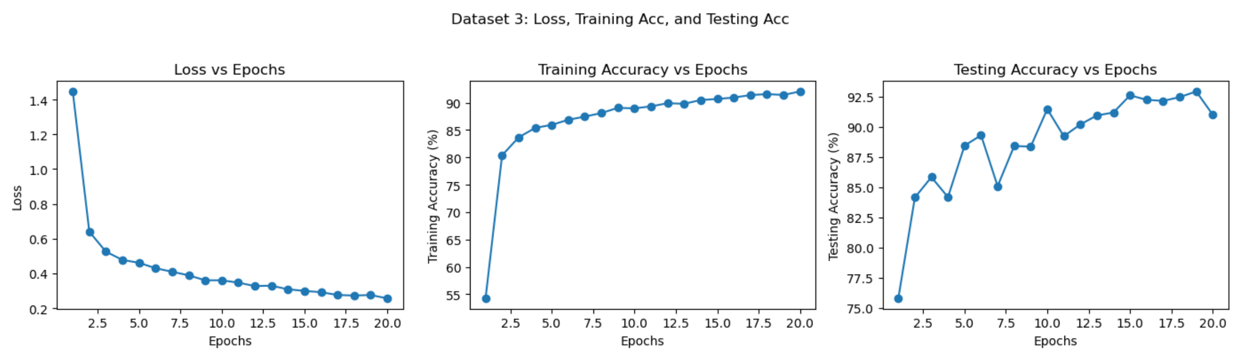

Fig. 3: Dataset 3 – Final Output (full size)

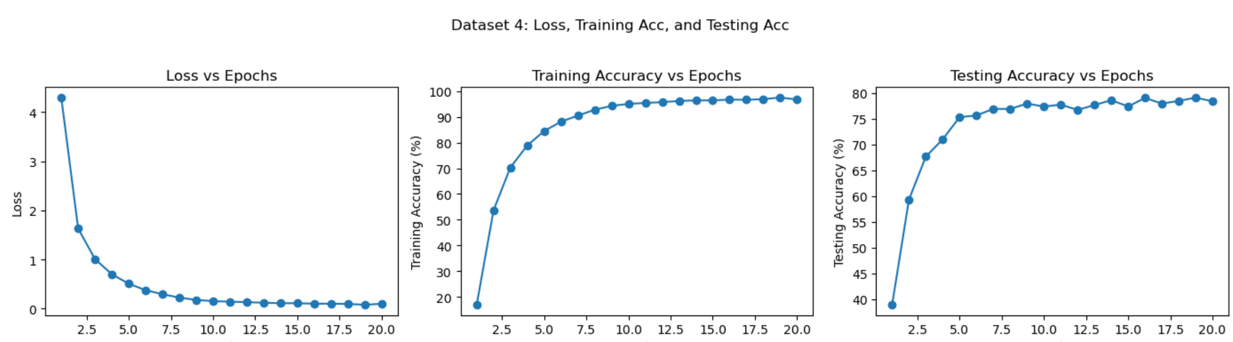

Fig. 4: Dataset 4 – Final Output (full size)