# Project Report: MNIST-based Four-Digit Number Recognition Using CNN

Introduction

This project's aim was to develop a deep learning model that can accurately recognize and classify sequences of four digits, using images derived from the MNIST dataset. By combining individual MNIST digits into sequences, we created a more complex classification task that mimics real-world applications such as reading zip codes or account numbers. We employed a Convolutional Neural Network (CNN), a type of deep learning model particularly suited for image recognition tasks, and evaluated its performance using standard metrics.

Dataset Description

The MNIST dataset is a collection of 70,000 handwritten digit images, widely used for training and testing in the field of machine learning. For this task, we concatenated individual digit images to form sequences representing four-digit numbers, generating a new custom dataset with a variety of sequences. This custom dataset maintains the original single-channel grayscale format of the MNIST images, and the four-digit sequences provide a unique challenge that extends the utility of the MNIST dataset. It contained total 10,000 classes, each class representing each number from 0 to 9999. Each of the class had 10 training examples hence the total number of training samples was equal to 100,000. The train to test split was 80/20. The label consisted of 40 units. Sets of 10 from 40 units were one hot encoded to represent presence of a particular digit at the respective location.
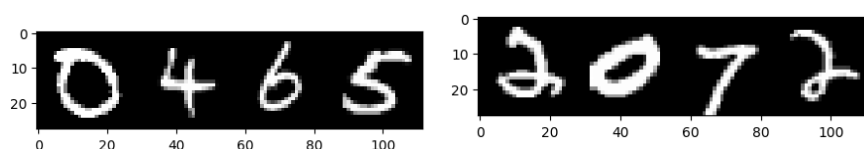
Model Architecture

Our CNN model consists of two convolutional layers, designed to capture spatial hierarchies in the data by learning from the arrangement and patterns of the digits. Following the convolutional layers, two fully connected layers interpret the features and output 40 units representing the probability distribution over the ten classes for each of the four positions in the sequence.

Training Process

The training process involved the use of a custom loss function that calculates the loss for each of the ten classes in all four digit positions independently, before averaging them. We employed the Adam optimizer with a learning rate of 0.001, iterating through the data for

seven epochs. Training and testing were conducted on separate subsets of the data to ensure the evaluation of the model's ability to generalize.

Evaluation Results

The model's evaluation after training yielded a test accuracy of 94.675%. This high accuracy indicates that the model is highly effective at correctly classifying the sequences of four digits, suggesting a robust understanding of the digit features and their sequences. The precision, recall, and F1-score across all classes were all recorded near to 0.98. Below is the loss vs epoch plot for the training and the testing loss.
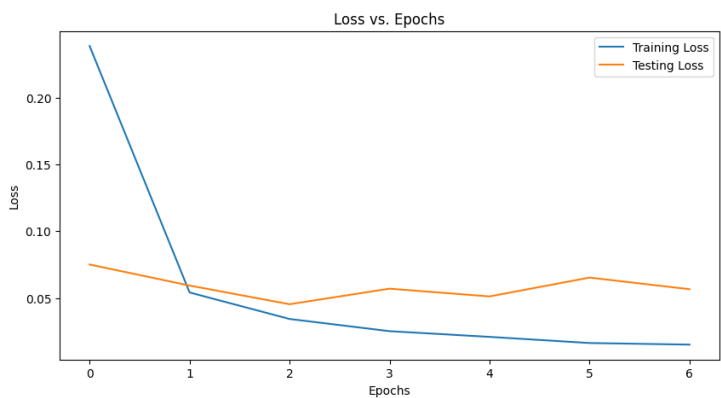
Metric	Score
Precision	0.9865692785914962
Recall	0.9863949914411935
F1 Score	0.986416185893081
Test Accuracy	0.94675
Hamming Loss	0.0027175

Possible Improvements

Improving the model could start with using more varied examples, as we were limited to 4 digits due to computer memory constraints. Secondly, the current model has been trained and evaluated on synthesized sequences from the MNIST dataset, which consists of relatively clean and well-defined images. For practical applications, it would be invaluable to adapt and test the model on a dataset of real-world images that contain digits in various conditions, such as different lighting, angles, and background noise.

Conclusion

The developed CNN model demonstrates outstanding performance in recognizing sequences of MNIST digits. The high precision, recall, and F1-score indicate the model's effectiveness, with potential applications in various domains requiring automated number recognition. Future work may focus on testing the model's robustness against images with noise and distortions, as well as exploring its scalability to sequences of variable lengths.