Comparison of ANN and CNN for Handwritten Digit Classification

Sai Pranav, CS22B1027

February 8, 2025

1 Introduction

This report compares an Artificial Neural Network (ANN) and a Convolutional Neural Network (CNN) for classifying handwritten digits from the MNIST dataset. The goal is to evaluate how different architectures impact model performance, considering accuracy, loss, and parameter count.

2 Dataset

The MNIST dataset consists of 60,000 training images and 10,000 test images, each of size 28×28 in grayscale.

3 Model Architectures

3.1 Artificial Neural Network (ANN)

The ANN consists of a simple feedforward structure:

- **Input Layer:** Flattened 28×28 image into a 784-dimensional vector.
- **Hidden Layers:** Dense layers with ReLU activation.
- **Output Layer:** 10 neurons with softmax activation for classification.

Model: "sequential_4"

Layer (type)	Output Shape	Param #
flatten_7 (Flatten)	(None, 784)	0
dense_12 (Dense)	(None, 128)	100,480
dense_13 (Dense)	(None, 64)	8,256
dense_14 (Dense)	(None, 10)	650

Total params: 109,386 (427.29 KB)
Trainable params: 109,386 (427.29 KB)
Non-trainable params: 0 (0.00 B)

Figure 1: ANN Architecture

3.2 Convolutional Neural Network (CNN)

The CNN utilizes convolutional and pooling layers for feature extraction:

- **Input Layer:** $28 \times 28 \times 1$ grayscale image.
- **Feature Extraction:** Two convolutional layers (32 and 64 filters) followed by max pooling.
- **Fully Connected Layers:** Dense layers with ReLU activation.
- **Output Layer:** 10 neurons with softmax activation for classification.

Model: "sequential_1"

Layer (type)	Output Shape	Param #
conv2d_2 (Conv2D)	(None, 28, 28, 32)	320
max_pooling2d_2 (MaxPooling2D)	(None, 14, 14, 32)	0
conv2d_3 (Conv2D)	(None, 14, 14, 64)	18,496
max_pooling2d_3 (MaxPooling2D)	(None, 7, 7, 64)	0
flatten_1 (Flatten)	(None, 3136)	0
dense_2 (Dense)	(None, 128)	401,536
dense_3 (Dense)	(None, 10)	1,290

Total params: 421,642 (1.61 MB)
Trainable params: 421,642 (1.61 MB)
Non-trainable params: 0 (0.00 B)

Figure 2: CNN Architecture

4 Comparison of Models

Model	Total Parameters	Test Accuracy	Test Loss
ANN	109,386	97.6%	0.09
CNN	421,642	99.2%	0.027

Table 1: Comparison of ANN and CNN Performance

5 Observations and Conclusion

5.1 Key Findings

- The ANN has significantly fewer parameters but achieves lower accuracy.
- The CNN outperforms the ANN in terms of accuracy due to better spatial feature extraction.
- The CNN achieves a lower test loss, indicating better generalization.

5.2 Conclusion

CNNs are better suited for image classification tasks, as they efficiently capture spatial hierarchies in images. While ANNs can classify images, they do not take advantage of spatial relationships, making them less effective for such tasks.