Comparison of MLP and CNN in Training on the MNIST Dataset

Taehyeon Kim
Undergraduate Student, Dept. of Computer Engineering
Kyonggi University
Suwon, Korea
dannykim05@kyonggi.ac.kr

***Abstract*—This paper compares the performance of MLP and CNN using the MNIST dataset. The experiment shows that CNN performs better than MLP in classifying images. The comparison is based on training loss values to evaluate the effectiveness of each model.**

***Index Terms*—CNN, MLP, MNIST, Image Classification, Deep Learning**

I. Introduction

MLP and CNN are types of artificial neural networks. These models both have nodes, weights, activation functions, etc. However, CNN is more commonly used than MLP in image classification. This paper aims to verify whether CNN truly performs better than MLP in image classification.

II. Method

The architectures of MLP and CNN, and training settings such as optimizer, batch size, and epochs are described below.

MLP Architecture:

- Flatten layer: reshapes input from [1, 28, 28] to [784]
- Fully connected layer: $784 \rightarrow 128$
- ReLU activation
- Fully connected layer: $128 \rightarrow 10$

CNN Architecture:

- 2D Convolutional layer: in channels = 1, out channels = 16, kernel size = 3, padding = 1
- Max Pooling layer: kernel size = 2, stride = 2
- Flatten layer: reshapes input from [16, 14, 14] to [3136]
- Fully connected layer: $3136 \rightarrow 10$

Training Settings:

- Optimizer: SGD
- Learning rate: 0.1
- Batch size: 64
- Epochs: 5

III. Results

In Epoch 1, loss values are similar. However, by epoch 5, the training loss of MLP fluctuates around $\pm$ 0.1, while CNN's training loss remains around $\pm$ 0.01. CNN's loss is ten times smaller than that of MLP.

TABLE I
MLP Training Loss (Selected Batches)

Epoch	Batch Index	Loss
1	0	2.2996
1	100	0.5997
1	200	0.4387
1	300	0.4640
1	400	0.1842
5	600	0.0815
5	700	0.0771
5	800	0.0886
5	900	0.1146

TABLE II
CNN Training Loss (Selected Batches)

Epoch	Batch Index	Loss
1	0	2.2996
1	100	0.5997
1	200	0.4387
1	300	0.4640
1	400	0.1842
5	600	0.1059
5	700	0.1204
5	800	0.0337
5	900	0.0096

IV. Discussion

CNN performs better than MLP in image classification. One reason is that the MLP model uses a flatten layer at the beginning, which loses local features. In contrast, the CNN model uses convolutional layers that extract local features through kernel operations. Due to this and other factors, CNN achieves lower loss values than MLP.

One limitation of this experiment is that the loss value of MLP slightly increases at the end of epoch 5. Another limitation is that only the loss value was used to compare the two models. Although loss is useful to evaluate performance, it is not sufficient alone. Other factors should also be considered for a more accurate comparison.

Conclusion

Both MLP and CNN are artificial neural networks. However, CNN is more widely used in image classification. This experiment suggests that this is because MLP loses spatial

information through flattening, while CNN preserves it via convolution. The results demonstrate that the model architecture has a significant impact on performance, and CNN is more suitable for image classification tasks such as MNIST.