Analysis of Baseline Code Report

Shachi Shah
ID#: 1217121828
School of Computing and Augmented Intelligence
Arizona State University Online
Azusa, California, United States of America
spshah22@asu.edu

I. Introduction

This report presents an analysis of handwritten digit recognition using a Convolutional Neural Network (CNN) trained on the MNIST dataset. The primary objective was to evaluate the baseline model's performance and analyze the impact of modifying kernel size and the number of feature maps. The MNIST dataset consists of 60,000 training images and 10,000 test images of handwritten digits (0-9), with each image being 28x28 pixels in grayscale.

II. Baseline Model Performance

The baseline CNN was configured with two convolutional layers using a kernel size of 3×3, with 6 and 16 feature maps, respectively. The architecture consisted of:

A. *Convolutional Layer 1:* 6 filters of size 3x3 with ReLU activation.
B. *Max Pooling Layer 1:* 2x2 pooling with stride 1.
C. *Convolutional Layer 2:* 16 filters of size 3x3 with ReLU activation.
D. *Max Pooling Layer 2:* 2x2 pooling with stride 1.
E. *Fully Connected Layer 1:* 120 neurons with ReLU activation.
F. *Fully Connected Layer 2:* 84 neurons with ReLU activation.
G. *Output Layer:* 10 neurons with softmax activation.

After training for 12 epochs, the model achieved a test accuracy of 97.82% with a test loss of 0.0625. The learning curves for training and validation loss and accuracy are provided in Figure 1. The model demonstrated a steady decline in loss while the accuracy improved with each epoch, indicating effective learning without signs of severe overfitting.

III. Results

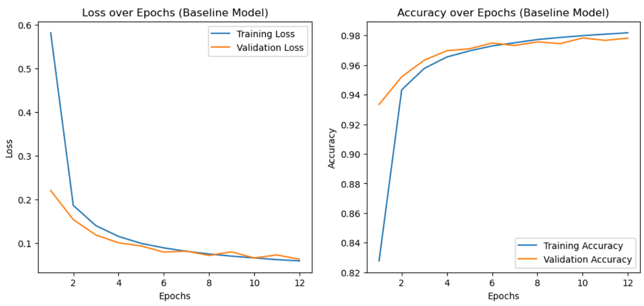

Figure 1: Loss over Epochs (Baseline Model)

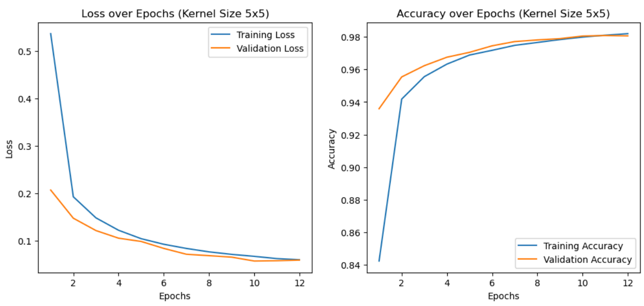

Figure 2: Loss over Epochs (Kernel Size 5x5)

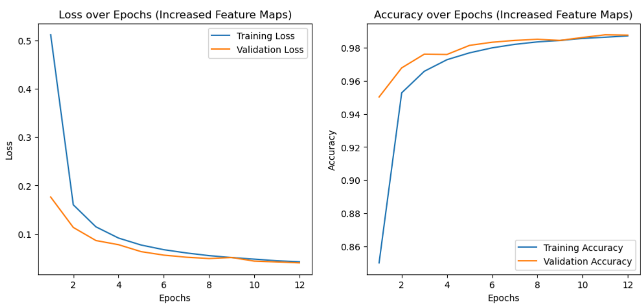

Figure 3: Loss over Epochs (Increased Feature Maps)

A. *Effect of Changing Kernel Size to 5x5*:

The kernel size was increased from 3×3 to 5×5 while keeping the feature maps unchanged. Larger kernels allow the model to capture more spatial information from images but can increase computational complexity. The revised architecture maintained the same layer structure but replaced the 3×3 filters with 5×5 filters in both convolutional layers.

This modification resulted in a slightly improved test accuracy of 98.06% with a test loss of 0.0587. The increase in kernel size enabled the model to recognize patterns more effectively, reducing the loss slightly and leading to marginal accuracy improvement. The corresponding learning curves, shown in Figure 2, illustrate a similar trend to the baseline but with slightly lower validation loss.

B. *Effect of Increasing Feature Maps:*

The number of feature maps was increased from (6,16) to (16,32) while keeping the kernel size at 3×3. Feature maps allow a CNN to learn more complex and diverse patterns within images. This modification resulted in an architecture with:

- Convolutional Layer 1: 16 filters of size 3×3 with ReLU activation.
- Max Pooling Layer 1: 2×2 pooling with stride 1.
- Convolutional Layer 2: 32 filters of size 3×3 with ReLU activation.
- Max Pooling Layer 2: 2×2 pooling with stride 1.

This modification yielded the best results, with a test accuracy of 98.76% and a test loss of 0.0401. The additional feature maps improved the model's ability to detect intricate patterns in the images, enhancing generalization to the test set. The learning curves in Figure 3 indicate a more pronounced decline in loss and an increase in accuracy, supporting the effectiveness of increasing feature maps.

IV. Observations

A. *Comparison and Analysis*:

Model Configuration	Test Accuracy	Test Loss
Baseline (3x3, 6&16 FM)	97.82%	0.0625
Kernel Size 5x5	98.06%	0.0587
Feature Maps 16&32	98.76%	0.0401

From the results, increasing the kernel size provided a minor improvement, whereas increasing the number of feature maps had a more significant positive impact on accuracy and loss reduction. The ability of deeper feature extraction to enhance model performance suggests that feature maps play a more crucial role in improving CNN effectiveness than merely enlarging the receptive field with bigger kernels. The increased number of feature maps allowed for better pattern recognition at different levels, leading to a superior final accuracy.

V. Conclusion

The experiment demonstrated that modifying CNN parameters affects classification performance. Increasing the kernel size slightly improved accuracy, but a more substantial gain was observed when increasing the number of feature maps. This indicates that additional feature maps enable better learning of hierarchical patterns in images, allowing for more detailed feature extraction. Furthermore, these findings highlight the importance of balancing model complexity with computational efficiency.

Future improvements could explore other hyperparameter tuning strategies such as dropout regularization, batch normalization, and deeper architectures to enhance generalization further. Additionally, training the model with different optimizers or adaptive learning rates could be explored to improve convergence speed and overall performance.

References

[1] Yiran, Luo. "Classification Using Neural Networks and Deep Learning Project" *CSE 575 - Statistical Machine Learning*, Ira A. Fulton Schools of Engineering, 13 January 2025.