A COMPARATIVE STUDY OF LENET-4 AND LENET-5 FOR CHINESE MNIST DATASET

Moses Sinanta P.W.J.

ABSTRACT

This research evaluates and compares the performance of LeNet-4 and LeNet-5 Convolutional Neural Network architectures on the Chinese MNIST dataset for handwritten digit classification. LeNet-4 comprises two convolutional layers followed by average pooling layers, while LeNet-5 extends this with an additional convolutional layer for enhanced feature extraction. Results indicate that while both models demonstrate strong capabilities in digit classification, LeNet-5 achieves higher accuracy and validation scores, suggesting superior generalization to new data, while LeNet-4 may be preferable in scenarios prioritizing lower prediction errors.

Keywords – *Convolutional Neural Networks, Handwritten digit classification, Chinese MNIST dataset*

INTRODUCTION

This research explores the application of Convolutional Neural Networks (CNNs), specifically LeNet-4 and LeNet-5 [1] to the Chinese MNIST dataset [2], which consists of handwritten Chinese digits. The motivation for this study arises from the need to improve the understanding of handwritten digit classification, a task in various optical character recognition (OCR) systems. The primary objective is to evaluate and compare the performance of these two CNN architectures in terms of accuracy, precision, recall and other relevant metrics. The scope of the study includes preprocessing the dataset, training the models and analyzing the results to determine which architecture provides better classification performance.

METHODOLOGY

The Chinese MNIST dataset was created as part of a project at Newcastle University, involving 100 Chinese nationals. Each participant wrote all 15 numbers with a standard black ink pen on a table with 15 designated regions on a white A4 paper, repeating this process 10 times. Each sheet was then scanned at a resolution of 300×300 pixels, resulting in a dataset of 15,000 images. These images represent one character from a set of 15 characters, grouped into samples and suites, with 10 samples per volunteer and 100 volunteers in total.

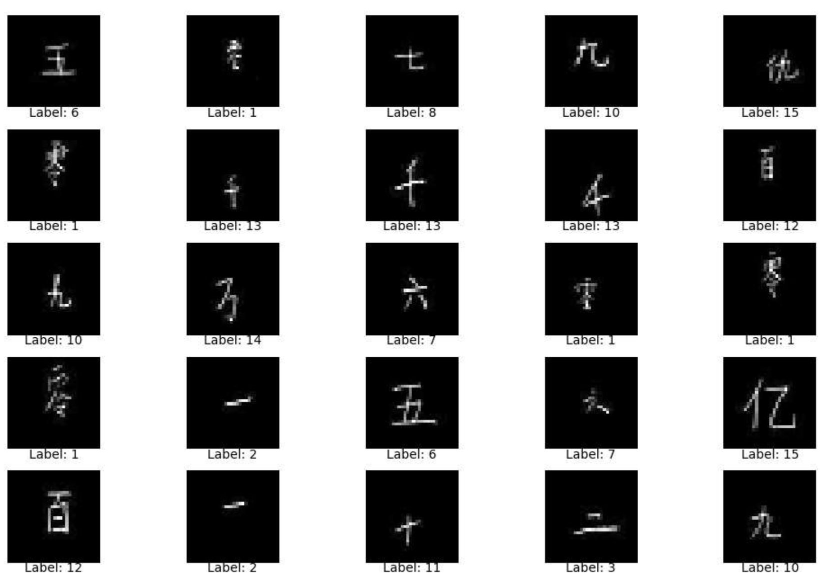

Figure 1 Chinese MNIST dataset sample data

The first model used for this research is the LeNet-4, a CNN designed for digit classification tasks. It consists of two convolutional layers, each followed by average pooling layers, which help in extracting features from the input images. The network includes a convolutional layer with six filters of size 5×5, followed by an average pooling layer, then a second convolutional layer with 16 filters of size 5×5, and another average pooling layer. The final layer is a dense layer with 84 neurons, leading to a softmax output layer for classification, employing the tanh activation function throughout the network.

LeNet-5, an enhancement over LeNet-4, is also a CNN tailored for digit classification tasks. This architecture includes three convolutional layers: the first with six filters of size 5×5, the second with 16 filters of size 5×5 and the third with 120 filters of size 5×5, each followed by average pooling layers. After feature extraction, the network utilizes two dense layers, the first with 120 neurons and the second with 84 neurons, culminating in a softmax output layer. The LeNet-5 architecture uses tanh activation functions, similar to LeNet-4, but with an additional convolutional layer that enhances its feature extraction capabilities and overall classification performance.

RESULTS

The following graphs summarizes the training results for both LeNet-4 and LeNet-5 models over 100 epochs in terms of training and validation loss as well as training and validation accuracy:

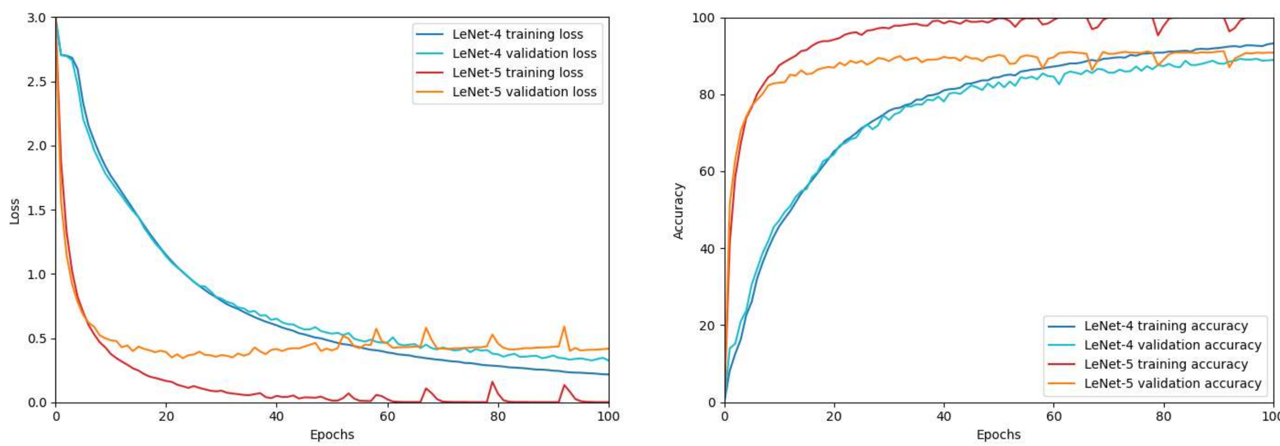

Figure 1 LeNet-4 and LeNet-5's training and validation loss (left) and accuracy (right) over epochs

The following table summarizes the key performance metrics for both LeNet-4 and LeNet-5:

Table 1 Experiment results

Model	Test accuracy	Test loss	Precision	Recall	F1 score
LeNet-4	89.2%	0.313	0.891	0.892	0.891
LeNet-5	91.6%	0.339	0.915	0.916	0.915

EVALUATION AND DISCUSSION

LeNet-5 exhibits a significantly lower training loss compared to LeNet-4, starting at 1.875 and decreasing to nearly zero, while LeNet-4's loss starts at 2.706 and decreases to 0.217. The validation loss for LeNet-5 also shows a more substantial decline, beginning at 1.563 and stabilizing around 0.417, whereas LeNet-4's validation loss starts at 2.704 and ends at 0.326. These findings indicate that LeNet-5 is more efficient in minimizing loss during both training and validation phases. Despite minor fluctuations, LeNet-5 maintains a consistently lower validation loss, suggesting better model performance and generalization.

In terms of training accuracy, LeNet-5 quickly reaches a perfect score of 100%, highlighting its strong capacity to learn from the training data, while LeNet-4's training accuracy improves steadily, peaking at 93.2%. For validation accuracy, LeNet-5 consistently outperforms LeNet-4, achieving a peak accuracy of 90.8% compared to LeNet-4's 88.9%. This indicates that LeNet-5 not only learns more effectively from the training data but also generalizes better to unseen data. The rapid attainment of perfect training accuracy in LeNet-5 suggests potential overfitting, yet its high validation accuracy demonstrates its robust generalization capabilities.

The performance metrics indicate that both LeNet-4 and LeNet-5 models are effective in classifying handwritten Chinese digits but with distinct strengths. LeNet-5 achieved a higher test accuracy of 91.6% compared to LeNet-4's 89.2%, suggesting it is better at generalizing to new data. However, LeNet-5's higher test loss of 0.339 compared to LeNet-4's 0.313 indicates that while it performs well overall, it might be slightly overfitting the training data. This trade-off between accuracy and test loss highlights the importance of balancing model complexity and generalization.

Furthermore, precision, recall and F1 score metrics for both models were closely aligned with their accuracy results, demonstrating consistent performance across different evaluation criteria. LeNet-5's higher precision and recall underscore its ability to correctly identify and classify more instances of handwritten digits. Both models, despite their minor misclassifications, maintained high-performance levels, illustrating their robustness and reliability. The comparative analysis thus underscores that while LeNet-5 generally offers better accuracy, LeNet-4's lower test loss might make it preferable in scenarios where minimizing prediction errors is crucial.

CONCLUSION

Both LeNet-4 and LeNet-5 demonstrate strong capabilities in classifying handwritten Chinese digits, with each model exhibiting distinct strengths. LeNet-5 achieves higher test accuracy and validation accuracy, indicating superior generalization to new data. However, LeNet-4's lower test loss suggests it may be more reliable in minimizing prediction errors, making it potentially preferable in certain applications. The findings highlight the importance of selecting the appropriate model based on specific performance priorities, whether it be higher accuracy or lower prediction error.

Future research could explore optimizing these architectures further or incorporating more advanced techniques such as data augmentation and regularization to enhance model performance and reduce overfitting. Additionally, investigating other CNN architectures or hybrid models could provide deeper insights and potentially yield even better results for handwritten Chinese digit classification.

REFERENCES

[1] LeCun, Yann. et.al. (1998). *Gradient-Based Learning Applied to Document Recognition*. *Proceedings of the IEEE*, 86(11), 2278-2324.

[2] Nazarpour, M.C.K. and M. Chen. (2017). *Handwritten Chinese Numbers*. Accessed on 17 May 2024 from https://data.ncl.ac.uk/articles/dataset/Handwritten_Chinese_Numbers/10280831.

This project is documented at https://github.com/MosesSinanta/Chinese_MNIST_LeNet_Comparison