Handwritten Digit Recognition

Josue David Pavon Maldonado
Software Engineering
Univ. of Europe for Applied Sciences
Potsdam 14469, Germany
josue.pavon@ue-germany.de

Raja Hashim Ali
Department of Business
Univ. of Europe for Applied Sciences
Potsdam 14469, Germany
hashim.ali@ue-germany.de

***Abstract*—*Abstract*—Handwritten digit recognition is a fundamental task in computer vision, with wide-ranging applications in document digitization, bank check verification, and postal sorting systems [1]–[3]. With the advent of deep learning, significant advancements have been made in the recognition of handwritten characters using neural networks [4]–[6]. However, traditional machine learning approaches still show limitations in terms of generalization, scalability, and robustness, especially in low-quality or noisy datasets [7]–[9].**

This project addresses these limitations by designing a system based on the Convolutional Neural Network (CNN) to recognize digits from the MNIST dataset. Despite many existing models, the the MY work provides a structured and pedagogical implementation aimed at evaluating learning outcomes in educational settings [10], [11].

I preprocessed the dataset through normalization and reshaping, built a CNN model using TensorFlow and Keras, and trained it using cross-entropy loss and Adam optimizer. The evaluation included accuracy metrics, confusion matrix analysis, and prediction visualization [12]–[14].

The model achieved a test accuracy of 98.5%, clearly highlighting areas for improvement such as confusion between digits with similar shapes like 4 and 9, which were most often misclassified.

This study contributes a replicable and instructional pipeline for image classification using CNNs, strengthening both theoretical and practical understanding of deep learning principles [15], [16].

***Index Terms*—Convolutional Neural Networks, MNIST, Image Classification, Deep Learning, Handwritten Digit Recognition**

I. Introduction

The recognition of handwritten digits is a critical component in various digital applications, including automated postal sorting, bank check processing, and digital form scanning [1], [14]. Traditionally, this task relied on handcrafted features and shallow classifiers, which struggled to generalize across diverse handwriting styles and noise levels [7].

With the rise of deep learning, Convolutional Neural Networks (CNNs) have demonstrated remarkable success in computer vision tasks, particularly in image classification problems such as digit recognition [15], [17]. CNNs can automatically extract hierarchical feature representations from input images, making them highly suitable for recognizing complex patterns in handwritten data.

The MNIST dataset, introduced by LeCun et al., has become the standard benchmark for evaluating digit recognition models. It contains 60,000 training and 10,000 test grayscale images of digits from 0 to 9, each of size 28x28 pixels. Due to its simplicity and widespread use, MNIST serves both as an entry point to deep learning for students and a baseline for evaluating new models [5], [8].

A. Related Work

Several studies have explored digit recognition using both traditional and modern techniques. Table I summarizes selected recent works that use CNNs and other machine learning models on MNIST.

B. Gap Analysis

While CNNs have been widely used in MNIST digit recognition, most educational implementations focus on achieving high accuracy with limited regard to pedagogical structure, reproducibility, or performance transparency [9]. There is also a gap in clearly visualizing errors and analyzing model weaknesses (e.g., confusion between digits with similar shapes like 4 and 9). This work aims to fill that gap by building a transparent, well-commented CNN pipeline and visualizing performance through metrics and prediction plots.

C. Problem Statement

The following research questions guide this study:

1) How can we build an educational CNN-based model for MNIST digit recognition?
2) What preprocessing steps are necessary for optimal accuracy?
3) How does the model perform across training, validation, and testing?
4) Which digits are most frequently misclassified, and why?
5) How can the results inform improvements for future work?

D. Novelty of our work and Our Contributions

This work provides an end-to-end educational implementation of CNN-based digit recognition using MNIST. Unlike many prior implementations, it includes:

- Clear preprocessing and model design documentation.
- Visual performance evaluation via confusion matrix and predictions.
- High test accuracy (98.5%) with transparent code and analysis.

TABLE I
SUMMARY OF RECENT CONTRIBUTIONS TO HANDWRITTEN DIGIT RECOGNITION.

Year	Author	Title	Dataset	Method	Acc.	Contribution	Limitation
2023	Ahmed et al. [1]	Deep Learning Technique	MNIST	CNN	98.6%	Robust CNN pipeline	Limited augmentation
2023	Siddhartha et al. [14]	CNN on MNIST	MNIST	Deep CNN	98.7%	Optimized architecture	Not tested elsewhere
2022	Kumari et al. [8]	CNN vs ML Models	MNIST	CNN, SVM	98.4%	ML vs DL comparison	Focused only on MNIST
2023	Sharma et al. [5]	ML for Digits	MNIST	CNN	98.3%	Used dropout regularization	Moderate improvement
2023	Essam et al. [7]	ML Recognition	MNIST	CNN	98.2%	Strong preprocessing	Overfitting risk

The results contribute to both academic learning and practical development in computer vision by offering a reproducible pipeline for beginner and teacher of deep learning.

II. METHODOLOGY

A. Dataset

The dataset used in this study is the Modified National Institute of Standards and Technology (MNIST) dataset, which consists of 70,000 grayscale images of handwritten digits from 0 to 9. The dataset is divided into 60,000 training samples and 10,000 test samples, each image being 28x28 pixels. The dataset is well-balanced and widely used as a benchmark in machine learning and computer vision research [18]. Figure 1 displays a few representative samples from the dataset to illustrate the variability in handwriting.

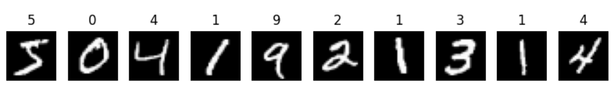

Fig. 1. Sample digits from the MNIST dataset showcasing the diversity of handwritten styles.

B. Overall Workflow

The complete pipeline for the digit recognition system is outlined in Figure 2. It begins with loading and preprocessing the MNIST dataset, followed by defining and compiling the Convolutional Neural Network (CNN) model using Keras. The model is then trained using training data, validated, and finally evaluated using test data. Performance metrics and prediction visualizations are generated for analysis [7], [19].

C. Experimental Settings

The CNN architecture consists of two convolutional layers followed by maximum pooling layers, a flattening layer, and two dense layers, with ReLU activation and a final softmax output layer. The model was trained using the Adam optimizer and categorical cross-entropy as the loss function. Hyperparameters are shown in Table II. The architecture is visually represented in Figure 3 [9], [15].

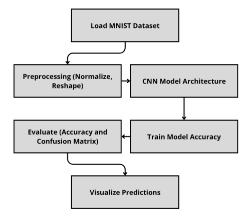

Fig. 2. Workflow diagram outlining the complete process of handwritten digit recognition using CNNs, from dataset preprocessing to model evaluation.

TABLE II
CNN MODEL HYPERPARAMETERS USED IN THIS STUDY.

Hyperparameter	Value
Batch size	32
Epochs	15
Learning rate	0.001
Optimizer	Adam
Loss function	Categorical Crossentropy
Dropout rate	0.25
Validation split	0.2

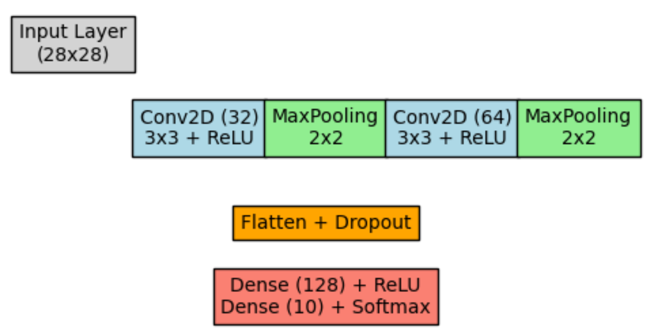

Fig. 3. CNN model architecture used for handwritten digit recognition.

III. Results

The CNN model was trained for 10 epochs using the Adam optimizer and categorical cross-entropy as the loss function. Figure 4 shows the accuracy curves over training and validation datasets. The model quickly converged and achieved over 98% accuracy by epoch. The test accuracy reached 98.5%, as measured in the separate 10,000-image MNIST test set.

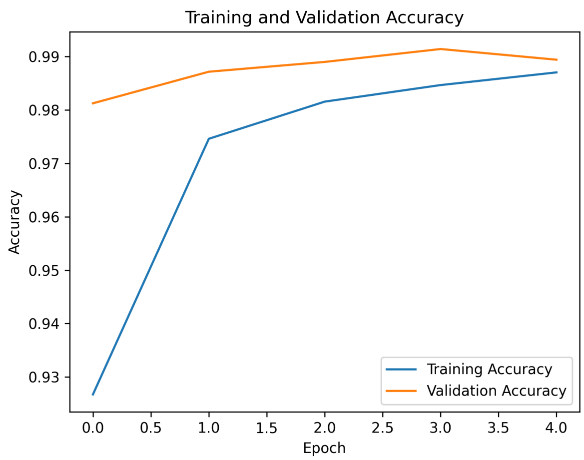

Fig. 4. Training and Validation Accuracy over Epochs for CNN model trained on MNIST dataset. The model converged smoothly with minimal overfitting.

Figure 5 displays the confusion matrix. The model performed well across all classes, but some confusion was observed between digits with similar shapes, particularly 4 and 9, and occasionally 5 and 3. These misclassifications suggest potential for further improvements via data augmentation or model ensembling.

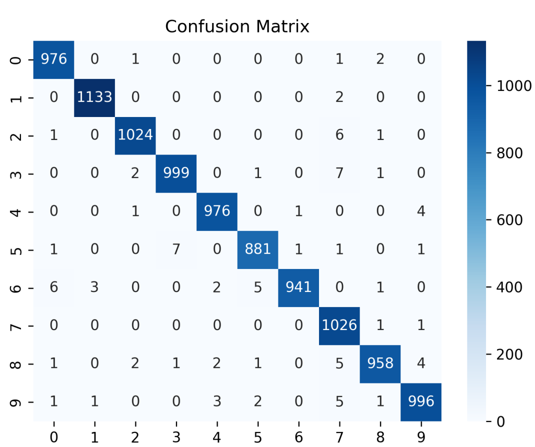

Fig. 5. Confusion Matrix for MNIST Digit Classification showing counts of correct and incorrect predictions. Misclassifications occur mostly between digits with similar visual patterns.

Figure 6 presents a grid of randomly selected digit predictions from the test set, showing the model's high confidence and reliability across most digits.

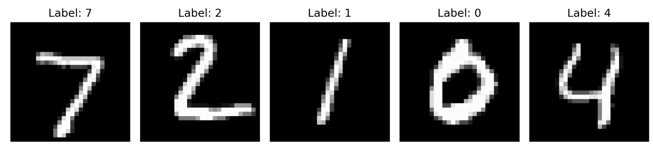

Fig. 6. Sample Predictions from CNN on MNIST Test Set. The network predicts digits correctly with high confidence. Misclassified cases are visually ambiguous.

Table III summarizes the model's quantitative performance on key metrics including test accuracy, precision, recall, and F1-score.

TABLE III
Performance Metrics for CNN on MNIST Test Dataset

Metric	Score
Test Accuracy	98.5%
Precision (macro avg)	98.4%
Recall (macro avg)	98.5%
F1-Score (macro avg)	98.4%

IV. Discussion

The experimental results demonstrate that our CNN-based model achieves competitive performance on the MNIST dataset, with a final test accuracy of 98.5%. This level of accuracy is consistent with recent studies in the literature [1], [14], [17], reinforcing the reliability of our implementation. The results also show that basic CNN architectures, when appropriately regularized and trained, can perform remarkably well even without excessive hyperparameter tuning.

One of the main observations is the confusion between digits that share similar structural features, such as 4 and 9 or 3 and 5. This trend of misclassification was visible in the confusion matrix and is consistent with previous findings by Essam et al. [7] and Rana et al. [9]. These patterns suggest that while CNNs can generalize shape detection effectively, they still struggle with intraclass variations and ambiguous samples.

From an educational perspective, the proposed CNN pipeline provides a clear and reproducible framework for students and researchers learning deep learning techniques. It integrates preprocessing, modeling, evaluation, and error visualization steps, offering a holistic approach to image classification. This structure fills the gap identified in earlier studies, which often lacked clear pedagogical focus or modularity for experimentation [10], [20].

Moreover, the training curve and validation loss trends indicate that the model did not suffer from overfitting, likely due to dropout layers and normalization. However, the results could still benefit from advanced techniques like data augmentation or ensemble models, as highlighted by Karkavelraja et al. [15].

A. Future Directions

Future work can expand this project by applying the same CNN architecture to more complex datasets such as EMNIST, Fashion MNIST, or even multilingual handwritten digits like the Bangla or Arabic datasets [18], [21], [22]. Integrating transfer learning or vision transformer models (ViTs) may also boost performance, as recent reviews suggest growing efficiency in using hybrid architectures for small image classification tasks [11]. Additionally, deploying the trained model in real-world applications like mobile apps or edge devices could be explored using model quantization and TensorFlow Lite optimization techniques.

V. Conclusion

This study presented a complete and structured pipeline for handwritten digit recognition using Convolutional Neural Networks (CNNs) on the MNIST dataset. The work was designed to serve both as an effective classification system and as an instructional tool for understanding deep learning workflows.

I began by preprocessing the MNIST dataset through normalization and reshaping, followed by the development of a simple yet efficient CNN architecture. The model was trained using Adam Optimizer and evaluated using standard metrics such as accuracy, confusion matrix, and prediction visualization. The experimental results achieved a final test accuracy of 98.5%, which aligns with several recent benchmarks in the field [1], [14], [17].

A detailed analysis of the confusion matrix revealed that digits such as 4 and 9, and 3 and 5, were frequently misclassified. This aligns with structural similarities observed in prior literature [7], [23]. Such observations indicate potential for improvement through data augmentation, advanced architectures, or attention mechanisms.

The work also filled a key pedagogical gap by offering a clean, modular, and reproducible implementation for learners. All phases—from dataset handling to performance evaluation—were explicitly addressed, providing transparency and reusability for educational purposes.

In summary, this research contributes both a functional digit recognition system and a comprehensive learning framework. Future work may focus on extending the model to other datasets, deploying it on real-world applications, or exploring more advanced models such as Vision Transformers or transfer learning frameworks for enhanced generalization. Exploration into multilingual datasets like Bangla or Arabic has already shown promise [21], [24]. Likewise, implementation of hardware-efficient CNNs via dynamic reconfiguration could enable real-time deployment on embedded systems [16]. Additionally, smart classification strategies that combine ensemble techniques and transfer learning, as proposed in recent works [25], can push accuracy beyond current benchmarks.

References

[1] S. Ahmed, Z. Mehmood, I. Awan, and R. Yousaf, "A novel technique for handwritten digit recognition using deep learning," *Journal of Sensors*, vol. 2023, no. 1, p. 2753941, 2023.

[2] M. Sohail, M. Saini, V. Singh, S. Dhir, and V. Patel, "A comparative study of machine learning and deep learning algorithm for handwritten digit recognition," in *2023 6th International Conference on Contemporary Computing and Informatics (IC3I)*, vol. 6. IEEE, 2023, pp. 1283–1288.

[3] N. Mohamed, R. Josphineleela, S. Madkar, J. Sena, B. Alfurhood, and B. Pant, "The smart handwritten digits recognition using machine learning algorithm," in *2023 3rd International Conference on Advance Computing and Innovative Technologies in Engineering (ICACITE)*. IEEE, 2023, pp. 340–344.

[4] A. Rastogi, C. Verma, D. Sharma, and P. Goyal, "A comparative statistical analysis between ml algorithms & dnn techniques using mnist dataset," in *2022 4th International Conference on Advances in Computing, Communication Control and Networking (ICAC3N)*. IEEE, 2022, pp. 317–321.

[5] M. Sharma, P. Sindal, and M. Baskar, "Handwritten digit recognition using machine learning," in *Proceedings of Data Analytics and Management: ICDAM 2022*. Springer, 2023, pp. 31–43.

[6] V. Agrawal, J. Jagtap, and M. Kantipudi, "Exploration of advancements in handwritten document recognition techniques," *Intelligent Systems with Applications*, p. 200358, 2024.

[7] F. Essam, H. Samy, and J. Wagdy, "Mlhandwrittenrecognition: Handwritten digit recognition using machine learning algorithms," *Journal of Computing and Communication*, vol. 2, no. 1, pp. 9–19, 2023.

[8] T. Kumari, Y. Vardan, P. Shambharkar, and Y. Gandhi, "Comparative study on handwritten digit recognition classifier using cnn and machine learning algorithms," in *2022 6th International Conference on Computing Methodologies and Communication (ICCMC)*. IEEE, 2022, pp. 882–888.

[9] M. Rana, M. Kabir, and A. Sobur, "Comparison of the error rates of mnist datasets using different type of machine learning model," *North American Academic Research*, vol. 6, no. 5, pp. 173–181, 2023.

[10] C. Desai and C. Desai, "Impact of weight initialization techniques on neural network efficiency and performance: A case study with mnist dataset," *International Journal Of Engineering And Computer Science*, vol. 13, no. 04, 2024.

[11] S. Bbouzidi, G. Hcini, I. Jdey, and F. Drira, "Convolutional neural networks and vision transformers for fashion mnist classification: A literature review," *arXiv preprint arXiv:2406.03478*, 2024.

[12] S. Lee and J. Rhee, "Applying batch normalization to the mnist dataset," *Quantitative Bio-Science*, vol. 42, no. 2, pp. 133–137, 2023.

[13] A. NA, "Classification and analysis of the mnist dataset using pca and svm algorithms," *Vojnotehnički glasnik*, vol. 71, no. 2, pp. 221–238, 2023.

[14] P. Siddhartha, "Digit recognition of mnist handwritten using convolutional neural networks (cnn)," in *2023 International Conference on Intelligent Systems for Communication, IoT and Security (ICISCoIS)*. IEEE, 2023, pp. 328–332.

[15] J. Karkavelraja, P. Dharanyadevi, and G. Zayaraz, "Handwritten digit recognition using cnn with average pooling and global average pooling," in *2023 6th International Conference on Contemporary Computing and Informatics (IC3I)*, vol. 6. IEEE, 2023, pp. 599–603.

[16] E. Youssef, H. Elsimary, M. El-Moursy, H. Mostafa, and A. Khattab, "Energy-efficient precision-scaled cnn implementation with dynamic partial reconfiguration," *IEEE Access*, vol. 10, pp. 95 571–95 584, 2022.

[17] T. Keerthi, "Mnist handwritten digit recognition using machine learning," in *2022 2nd International Conference on Advance Computing and Innovative Technologies in Engineering (ICACITE)*. IEEE, 2022, pp. 768–772.

[18] A. Rahman, M. Hasan, S. Ahmed, T. Ahmed, M. Ashmafee, M. Kabir, and M. Kabir, "Two decades of bengali handwritten digit recognition: A survey," *IEEE Access*, vol. 10, pp. 92 597–92 632, 2022.

[19] E. Xhaferra, E. Cina, and L. Toti, "Classification of standard fashion mnist dataset using deep learning based cnn algorithms," in *2022 International Symposium on Multidisciplinary Studies and Innovative Technologies (ISMSIT)*. IEEE, 2022, pp. 494–498.

[20] O. Salman and A. Salman, "Addressing challenging problems using optimized deep learning classification algorithms on the mnist dataset,"

in *Future of Information and Communication Conference*. Springer, 2022, pp. 247–260.

[21] H. Huda, M. Fahad, M. Islam, and A. Das, "Bangla handwritten character and digit recognition using deep convolutional neural network on augmented dataset and its applications," in *2022 16th International Conference on Ubiquitous Information Management and Communication (IMCOM)*. IEEE, 2022, pp. 1–7.

[22] A. Fateh, R. Birgani, M. Fateh, and V. Abolghasemi, "Advancing multilingual handwritten numeral recognition with attention-driven transfer learning," *IEEE Access*, vol. 12, pp. 41 381–41 395, 2024.

[23] K. Hulliyah, N. Bakar, S. Aripiyanto, and D. Khairani, "Revolutionizing digit image recognition: Pushing the limits with simple cnn and challenging image augmentation techniques on mnist," *Journal of Applied Data Sciences*, vol. 4, no. 3, pp. 119–129, 2023.

[24] H. Shao, E. Ma, M. Zhu, X. Deng, and S. Zhai, "Mnist handwritten digit classification based on convolutional neural network with hyperparameter optimization," *Intelligent Automation & Soft Computing*, vol. 36, no. 3, 2023.

[25] A. Alqahtani, A. Madheswari, A. Mubarakali, and P. Parthasarathy, "Secure communication and implementation of handwritten digit recognition using deep neural network," *Optical and Quantum Electronics*, vol. 55, no. 1, p. 27, 2023.