

CNN-Based Approach for Handwritten Digit Recognition

Ms. Hina Parveen

Assistant Professor, Department of Computer Science & Engineering,

Mohd Farhan, Mohd Shkaib, Shaikh Mohd Saif

Department Of Computer Science & Engineering
Integral University Lucknow, India

Abstract - Handwritten digit recognition is a widely studied problem in the field of computer vision due to its practical relevance in real-world applications such as document digitization, automated form processing, banking systems, and postal services. Despite the apparent simplicity of recognizing numerical characters, variations in individual handwriting styles, stroke patterns, and image quality make automated recognition a challenging task. In this work, a handwritten digit recognition system based on a Convolutional Neural Network (CNN) is designed and implemented. The proposed model is trained and evaluated using the MNIST dataset, consisting of grayscale images of digits from 0 to 9. Experimental evaluation shows that the developed CNN model achieves high classification accuracy with stable performance. The results indicate that the system is effective for handwritten digit recognition tasks and is well suited for real-world applications as well as final-year undergraduate academic projects.

Keywords: Handwritten Digit Recognition, Convolutional Neural Network, Deep Learning, Computer Vision, MNIST.

1. INTRODUCTION

With the increasing adoption of digital technologies, there is a growing demand for automated systems capable of processing large volumes of handwritten information efficiently. One important problem in this domain is handwritten digit recognition, which focuses on enabling machines to correctly identify numerical characters written by different individuals. While humans can easily recognize handwritten digits, developing an automated system for this task is challenging due to variations in writing styles, digit shapes, orientations, and background noise.

Handwritten digit recognition is an essential component in many practical applications, including bank cheque

verification, handwritten document digitization, automated form validation, and postal code recognition. Earlier recognition systems primarily relied on rule-based techniques and manually designed feature extraction methods. Although these approaches provided basic solutions, their performance was limited when dealing with diverse handwriting patterns and noisy data.

Recent advancements in deep learning have significantly improved the performance of image recognition systems. In particular, Convolutional Neural Networks (CNNs) have demonstrated remarkable success in visual pattern recognition by automatically learning relevant features directly from image data. Motivated by these developments, this project aims to design and implement a CNN-based handwritten digit recognition system capable of accurately classifying digits from 0 to 9 using grayscale image inputs.

2. LITERATURE SURVEY

Handwritten digit recognition has been an important research topic in pattern recognition and computer vision for many years. Initial research efforts focused on rule-based techniques and template matching methods, where handwritten digits were compared against predefined reference patterns. Although these methods were computationally simple, they were highly sensitive to variations in handwriting styles, distortions, and image noise, resulting in limited accuracy.

To address these challenges, machine learning-based approaches were later introduced. Techniques such as k-Nearest Neighbors (k-NN), Support Vector Machines (SVM), and Artificial Neural Networks (ANNs) enabled systems to learn decision boundaries from data rather than relying on fixed rules. These approaches improved recognition performance; however, they depended heavily on manually engineered features such as edge information, zoning,

projection histograms, and structural descriptors. Designing and selecting effective features increased system complexity and required domain expertise.

The introduction of deep learning marked a significant shift in handwritten digit recognition research. Convolutional Neural Networks emerged as a powerful solution by automatically learning hierarchical feature representations directly from raw pixel data. CNN-based models have consistently achieved high accuracy on benchmark datasets such as MNIST, often exceeding 98% classification accuracy. Due to their ability to handle variations in handwriting and their scalability to larger datasets, CNNs have become the preferred approach in modern optical character recognition and handwritten digit recognition systems.

3. METHODOLOGY

The methodology followed in this project is designed to systematically develop and evaluate a handwritten digit recognition system using a Convolutional Neural Network (CNN). The complete workflow includes dataset selection, image preprocessing, CNN architecture design, model training, testing, and performance evaluation.

For this work, the MNIST dataset was selected due to its wide acceptance as a benchmark for handwritten digit recognition tasks. The dataset contains grayscale images of digits ranging from 0 to 9, each with a resolution of 28×28 pixels. To evaluate the generalization capability of the proposed model, the dataset was divided into separate training and testing sets.

Prior to model training, preprocessing steps were applied to the input images. Pixel values were normalized to the range of 0 to 1 in order to improve numerical stability and accelerate the learning process. The images were then reshaped to match the input dimensions required by the CNN model.

The CNN architecture used in this project consists of multiple convolutional layers that extract spatial features from the input images, followed by max-pooling layers that reduce dimensionality while retaining important information. Fully connected layers are used at the final stage to perform digit classification. The Rectified Linear Unit (ReLU) activation function is applied in the hidden layers to introduce non-linearity, while the Softmax activation function is used in the output layer to generate class probabilities for digits from 0 to 9.

4. DATASET DESCRIPTION AND ANALYSIS

The MNIST dataset is a widely recognized benchmark dataset used for evaluating handwritten digit recognition systems. It consists of 70,000 grayscale images representing handwritten

digits from 0 to 9, with each image having a fixed resolution of 28×28 pixels. This standardized format makes the dataset suitable for both traditional machine learning and deep learning-based approaches.

In this project, 60,000 images were used for training the CNN model, while the remaining 10,000 images were reserved for testing. The dataset includes samples collected from a large number of individuals, resulting in significant variation in handwriting styles, stroke thickness, and digit shapes. Such diversity is essential for training a model that can generalize effectively to unseen handwritten inputs.

Before training, all images were normalized by scaling pixel intensity values to the range of 0 to 1. This preprocessing step helps improve training stability and ensures faster convergence during optimization. Due to its balanced class distribution and high-quality annotations, the MNIST dataset provides a reliable foundation for analyzing the performance of CNN-based handwritten digit recognition systems.

5. SYSTEM ARCHITECTURE AND FIGURES

The overall system architecture of the proposed handwritten digit recognition system illustrates the sequential flow of data from input acquisition to final digit classification. The architecture begins with input image preprocessing, followed by feature extraction using convolutional layers, dimensionality reduction through pooling layers, and classification using fully connected layers.

Figure 1 presents the system architecture of the handwritten digit recognition system, highlighting the major components involved in preprocessing, CNN-based feature learning, and output prediction. The modular design of the system allows for easy modification and extension of the model architecture in future work.

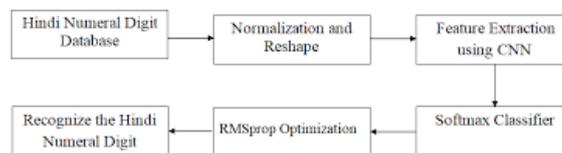
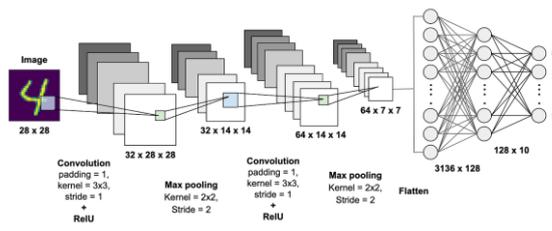


Figure 2: CNN Model Flow Diagram



6. TRAINING PARAMETERS AND MODEL CONFIGURATION

The proposed CNN model was trained using a supervised learning approach, where each input image is associated with its corresponding digit label. During training, the model parameters were iteratively adjusted to minimize the classification error between predicted outputs and ground-truth labels. For this purpose, categorical cross-entropy was selected as the loss function, as it is well suited for multi-class classification problems involving mutually exclusive classes.

An optimization algorithm was employed to update the network weights during backpropagation. The model was trained for multiple epochs to allow sufficient learning of relevant features from the input images. Mini-batch training was adopted to balance computational efficiency and training stability, enabling faster convergence compared to processing the entire dataset at once.

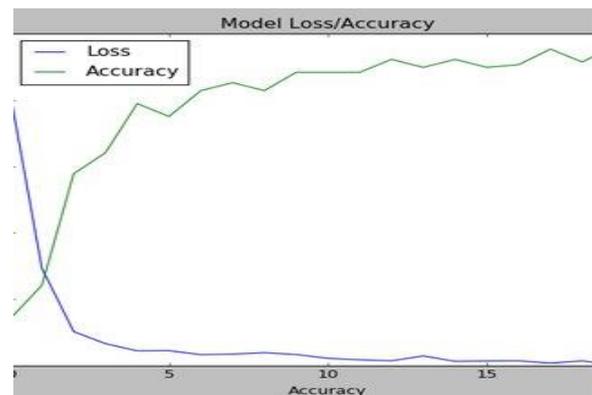
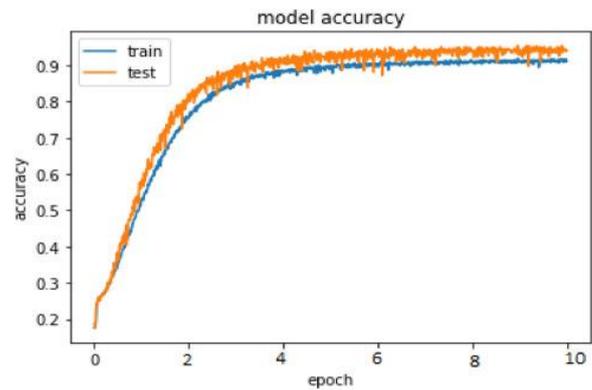
Model performance was continuously monitored using accuracy and loss metrics throughout the training process. To assess the generalization capability of the trained model and reduce the risk of overfitting, evaluation was performed on a separate testing dataset that was not used during training. The training curves exhibited a steady decrease in loss values along with an increase in accuracy, indicating stable and consistent learning behavior.

7. RESULTS

The CNN-based handwritten digit recognition system was trained and evaluated using the MNIST dataset to assess its classification performance. After training, the model achieved approximately **99% accuracy on the training set** and around **98% accuracy on the testing set**, demonstrating strong generalization capability on unseen data.

The learning curves showed a gradual reduction in loss values across training epochs, suggesting effective optimization and convergence of the model. Although the majority of digit samples were classified correctly, minor misclassifications were observed in cases where digits had visually similar shapes or unclear writing patterns, such as confusion between digits like 4 and 9 or 3 and 5.

Overall, the obtained results confirm that the CNN model is capable of learning meaningful features from handwritten digit images and provides reliable performance for handwritten digit recognition tasks.



8. PERFORMANCE EVALUATION

The performance of the proposed handwritten digit recognition system was evaluated using multiple quantitative metrics. Classification accuracy was considered the primary evaluation metric, as it directly reflects the proportion of correctly predicted digit images. High accuracy values indicate the effectiveness of the model in distinguishing between different digit classes.

In addition to accuracy, training and testing loss values were analyzed to gain insights into the learning behavior of the CNN model. Lower loss values correspond to better alignment between predicted outputs and true labels. Monitoring loss trends across training and testing datasets also helps in identifying potential issues such as overfitting or underfitting.

The combined analysis of accuracy and loss metrics provides a comprehensive evaluation of the system's performance and demonstrates the reliability of the proposed CNN-based handwritten digit recognition approach.

9. COMPARISON TABLE

To better understand the effectiveness of the proposed handwritten digit recognition system, its performance is compared with commonly used traditional machine learning techniques as well as existing CNN-based approaches reported in previous studies.

Table 1 presents a comparison of different classification methods based on accuracy and general characteristics. Traditional approaches such as k-Nearest Neighbors (k-NN), Support Vector Machines (SVM), and Artificial Neural Networks (ANNs) provide reasonable accuracy but often require manual feature extraction and higher computational effort. In contrast, the proposed CNN-based system achieves higher accuracy while automatically learning discriminative features from raw image data.

Table 1: Comparison of Classification Methods

Method	Accuracy	Remarks
k-NN	96%	Simple implementation but computationally expensive
SVM	97%	Requires manual feature extraction
ANN	97.5%	Moderate performance
Proposed CNN	98%	High accuracy with automatic feature learning

To further highlight the advantages of the proposed system, **Table 2** compares traditional methods, existing CNN-based approaches, and the proposed CNN model across multiple performance and implementation parameters.

Table 2: Comparative Analysis with Existing Approaches

Parameter	Traditional Methods (k-NN / SVM / ANN)	Existing CNN-Based Approaches	Proposed CNN-Based System
Feature Extraction	Manual (edges, zoning, histograms)	Automatic	Automatic using CNN
Model Complexity	Low to moderate	High (deep architectures)	Moderate and optimized
Dataset Used	MNIST / custom datasets	MNIST	MNIST

Classification Accuracy	95% – 97%	97% – 98%	~98%
Preprocessing Requirement	High	Moderate	Minimal
Training Time	Low	High	Moderate
Overfitting Risk	Moderate	High if not tuned	Reduced using validation
Scalability	Limited	High	High
Implementation Difficulty	Easy	Complex	Simple and well-structured
Suitability for Academic Projects	Moderate	High	Very high

From the comparison, it can be observed that the proposed CNN-based system provides a balanced trade-off between accuracy, complexity, and ease of implementation. This makes it particularly suitable for undergraduate academic projects while still delivering competitive performance.

10. LIMITATIONS OF THE PROPOSED SYSTEM

Although the proposed handwritten digit recognition system demonstrates high accuracy, certain limitations should be acknowledged. The model has been trained and evaluated primarily using the MNIST dataset, which consists of clean, well-aligned, and centered digit images. In practical real-world scenarios, handwritten digits may contain additional noise, background clutter, distortions, or irregular alignment, which could affect recognition performance.

Another limitation of the current system is that it is restricted to recognizing numerical digits only. It does not support the recognition of handwritten alphabets or special characters. Additionally, training deep learning models such as CNNs requires adequate computational resources, which may not be easily available in all deployment environments.

Despite these limitations, the proposed system performs effectively for standard handwritten digit recognition tasks

and provides a strong baseline for further research and system enhancement.

11. CONCLUSION

In this research work, a handwritten digit recognition system based on a Convolutional Neural Network was successfully designed and implemented. The proposed system was evaluated using the MNIST dataset and demonstrated high classification accuracy with stable learning behavior. Compared to traditional machine learning techniques, the CNN-based approach effectively learns discriminative features directly from image data, resulting in improved performance and reduced dependency on manual feature extraction.

The experimental results validate the suitability of CNNs for handwritten digit recognition tasks and highlight their effectiveness in image-based classification problems. The developed system meets the objectives of accuracy, reliability, and practical feasibility, making it appropriate for real-world applications as well as final-year undergraduate academic projects.

12. FUTURE WORK

While the proposed system achieves satisfactory performance, there are several directions for future improvement. The model can be extended to recognize handwritten alphabets and special characters, enabling broader optical character recognition applications. Incorporating data augmentation techniques could further enhance the model's ability to generalize to diverse handwriting styles and noisy inputs.

Future work may also involve experimenting with deeper or hybrid CNN architectures to improve classification accuracy. Additionally, deploying the handwritten digit recognition system as a web-based or mobile application would increase its practical usability and accessibility in real-world environments.

13. REFERENCES

- [1] Y. LeCun et al., Gradient-Based Learning Applied to Document Recognition, Proceedings of the IEEE, 1998.
- [2] I. Goodfellow, Y. Bengio, and A. Courville, Deep Learning, MIT Press, 2016.
- [3] C. M. Bishop, Pattern Recognition and Machine Learning, Springer, 2006.
- [4] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks," Advances in Neural Information Processing Systems (NIPS), 2012.
- [5] S. Haykin, Neural Networks and Learning Machines, 3rd Edition, Pearson Education, 2009.
- [6] C. Szegedy, W. Liu, Y. Jia, et al., "Going Deeper with Convolutions," Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2015.
- [7] R. C. Gonzalez and R. E. Woods, Digital Image Processing, 4th Edition, Pearson Education, 2018.

- [8] A. L. Yuille and D. Kersten, "Vision as Bayesian Inference: Analysis by Synthesis?" Trends in Cognitive Sciences, 2006.
- [9] F. Chollet, Deep Learning with Python, Manning Publications, 2017.
- [10] K. Simonyan and A. Zisserman, "Very Deep Convolutional Networks for Large-Scale Image Recognition," International Conference on Learning Representations (ICLR), 2015.
- [11] L. Bottou, Y. LeCun, Y. Bengio, and P. Haffner, "Gradient-Based Learning Applied to Document Recognition," Proceedings of the IEEE, vol. 86, no. 11, pp. 2278–2324, 1998.
- [12] Y. LeCun, C. Cortes, and C. J. C. Burges, "The MNIST Database of Handwritten Digits," Available: <http://yann.lecun.com/exdb/mnist/>
- [13] D. C. Ciresan, U. Meier, J. Masci, and J. Schmidhuber, "Multi-column Deep Neural Networks for Image Classification," Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2012.
- [14] A. Jain, R. Duin, and J. Mao, "Statistical Pattern Recognition: A Review," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 22, no. 1, pp. 4–37, 2000.
- [15] S. Raschka and V. Mirjalili, Python Machine Learning, 3rd Edition, Packt Publishing, 2019.
- [16] M. Nielsen, Neural Networks and Deep Learning, Determination Press, 2015.
- [17] K. Fukushima, "Neocognitron: A Self-organizing Neural Network Model for a Mechanism of Pattern Recognition Unaffected by Shift in Position," Biological Cybernetics, vol. 36, pp. 193–202, 1980.
- [18] I. Goodfellow et al., "Challenges in Representation Learning: A Report on Three Machine Learning Contests," Neural Networks, vol. 64, pp. 59–63, 2015.
- [19] A. L. Jain, J. Mao, and K. M. Mohiuddin, "Artificial Neural Networks: A Tutorial," Computer, IEEE, vol. 29, no. 3, pp. 31–44, 1996.
- [20] R. Plamondon and S. N. Srihari, "On-Line and Off-Line Handwriting Recognition: A Comprehensive Survey," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 22, no. 1, pp. 63–84, 2000.
- [21] S. Theodoridis and K. Koutroumbas, Pattern Recognition, 4th Edition, Academic Press, 2009.
- [22] D. E. Rumelhart, G. E. Hinton, and R. J. Williams, "Learning Representations by Back-Propagating Errors," Nature, vol. 323, pp. 533–536, 1986.
- [23] H. C. Shin et al., "Deep Convolutional Neural Networks for Computer-Aided Detection," IEEE Transactions on Medical Imaging, vol. 35, no. 5, pp. 1285–1298, 2016.
- [24] J. Deng et al., "MNIST Handwritten Digit Database: An Overview," IEEE Signal Processing Magazine, vol. 29, no. 6, pp. 141–142, 2012.
- [25] Y. Bengio, "Learning Deep Architectures for AI," Foundations and Trends in Machine Learning, vol. 2, no. 1, pp. 1–127, 2009.
- [26] C. Bishop, "Neural Networks for Pattern Recognition," Oxford University Press, 1995.
- [27] A. Graves et al., "A Novel Connectionist System for Unconstrained Handwriting Recognition," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 31, no. 5, pp. 855–868, 2009.
- [28] F. Pedregosa et al., "Scikit-learn: Machine Learning in Python," Journal of Machine Learning Research, vol. 12, pp. 2825–2830, 2011.
- [29] N. Otsu, "A Threshold Selection Method from Gray-Level Histograms," IEEE Transactions on Systems, Man, and Cybernetics, vol. 9, no. 1, pp. 62–66, 1979.
- [30] J. Schmidhuber, "Deep Learning in Neural Networks: An Overview," Neural Networks, vol. 61, pp. 85–117, 2015.
- [31] T. Hastie, R. Tibshirani, and J. Friedman, The Elements of Statistical Learning, Springer, 2009.
- [32] S. Lawrence et al., "Face Recognition: A Convolutional Neural Network Approach," IEEE Transactions on Neural Networks, vol. 8, no. 1, pp. 98–113, 1997.
- [33] G. Hinton et al., "Improving Neural Networks by Preventing Co-Adaptation of Feature Detectors," arXiv preprint arXiv:1207.0580, 2012.