

MNIST Digit Classification: Deep Learning Approach for Efficient Handwritten Digit Recognition

¹MR.P.Chaitanya,²D.Deepika,³G.Jyothi Sri,⁴K.Aditya Ram, ⁵G.Jeevan Sai Kiran

¹Associate Professor, DEPT-CSE, Srinivasa Institute of Engineering and Technology

^{2,3,4,5}UG Scholar, DEPT-CSE, Srinivasa Institute of Engineering and Technology

doi.org/10.64643/IJIRTV12I10-194957-459

Abstract: Handwritten digit recognition is a fundamental problem in computer vision and pattern recognition, with wide applications in postal mail sorting, bank check processing, and form digitization. This paper presents a deep learning approach for efficient classification of handwritten digits using the MNIST dataset, which consists of 70,000 grayscale images of digits (0–9). The study uses data preprocessing techniques such as normalization, reshaping, and data augmentation, followed by the application of Convolutional Neural Network (CNN) architectures for classification. Algorithms such as Logistic Regression, Support Vector Machine (SVM), and deep CNN models are applied for digit classification, and the results show that deep learning can effectively identify handwritten digit patterns and provide highly accurate recognition. The proposed Deep CNN model achieves 99.42% accuracy on the MNIST test set, with further improvement to 99.56% through data augmentation techniques.

Keywords: Handwritten Digit Recognition, Deep Learning, CNN, MNIST Dataset, Image Classification

I. INTRODUCTION

In recent years, the rapid advancement of deep learning techniques has transformed the way computers process and recognize visual patterns. Handwritten digit recognition, a core problem in optical character recognition (OCR), has become a benchmark task for evaluating machine learning and deep learning models. The Modified National

Institute of Standards and Technology (MNIST) dataset, introduced by LeCun et al. in 1998, provides a standardized collection of 70,000 grayscale images of handwritten digits (0–9), each of size 28×28 pixels, split into 60,000 training and 10,000 test images.

Handwritten digit recognition plays an important role in numerous real-world applications including automated postal mail sorting, bank check processing, tax form digitization, and educational assessment systems. However, due to the wide variation in individual handwriting styles, stroke thickness, and digit orientation, achieving robust and accurate recognition remains challenging.

The primary aim of this research is to develop a deep learning-based digit classification system using the MNIST dataset to achieve efficient and accurate handwritten digit recognition. The proposed system uses data preprocessing techniques and multiple classification approaches—from traditional machine learning to deep CNNs—to analyse digit images effectively. By leveraging learned hierarchical features, this study provides a comprehensive comparison of approaches and demonstrates the superiority of deep learning methods for this task.

II. PROBLEM STATEMENT

In automated document processing and digitization systems, accurate recognition of handwritten digits is a critical requirement. Applications such as postal code recognition, bank check verification, and form data extraction depend on reliable digit classification. However, due to the enormous variety in human handwriting—including differences in stroke width, slant angle, digit size, and writing pressure—building a universally accurate recognition system is extremely challenging.

Traditional approaches to digit recognition relied on handcrafted feature extraction methods combined with classical classifiers. These methods are time-consuming to design, inconsistent across different

handwriting styles, and unable to capture complex spatial patterns effectively. Without a robust automated system, applications may produce incorrect readings leading to misrouted mail, financial errors, or data entry mistakes.

Therefore, there is a need for an intelligent and automated solution that can efficiently process handwritten digit images and accurately classify them. The problem addressed in this research is the development of a deep learning-based classification system capable of processing MNIST digit images and categorizing them into their correct digit classes (0–9) with high accuracy, robustness, and computational efficiency.

III. LITERATURE REVIEW

Handwritten digit recognition has been an important research area in the fields of computer vision and pattern recognition for several decades. Many researchers have explored different techniques to classify digit images, progressing from simple template matching to sophisticated deep learning architectures.

Early research primarily focused on rule-based and feature-engineering approaches, where handcrafted features such as Histogram of Oriented Gradients (HOG), pixel intensity statistics, and zoning methods were used to represent digit images. While these methods provided basic classification capability, they were limited in handling variations in writing style, rotation, and scale. As a result, researchers shifted towards machine learning-based approaches that can learn discriminative patterns directly from data.

Several studies have applied supervised machine learning algorithms such as k-Nearest Neighbours, Support Vector Machines (SVMs), Random Forests, and Logistic Regression for digit classification tasks. Vapnik's SVM achieved 98.5% accuracy on MNIST in the late 1990s. These models demonstrated solid performance but required careful feature engineering and struggled with complex spatial relationships in images.

The breakthrough came with LeCun's LeNet-5 architecture (1998), a pioneering CNN that demonstrated the power of learned hierarchical features for digit recognition. Since then, deeper architectures—including VGGNet, ResNet, and Capsule Networks—have pushed accuracy beyond

99.7%. Cireşan et al. (2012) achieved 99.77% using GPU-trained deep neural networks with elastic deformations. Recent research has also explored transformer-based architectures and neural architecture search for optimizing digit classification models.

Based on the findings of previous studies, this research adopts a deep CNN approach with batch normalization and dropout regularization to achieve efficient and accurate MNIST digit classification, balancing accuracy with computational efficiency.

IV. SYSTEM ARCHITECTURE

The system architecture of the proposed MNIST Digit Classification system is designed as a multi-stage pipeline that efficiently processes handwritten digit images and generates accurate classification results. The architecture consists of four major layers: Data Ingestion, Data Processing, Core Deep Learning & Classification, and Visualization & Insights, which work sequentially to transform raw digit images into classified outputs.

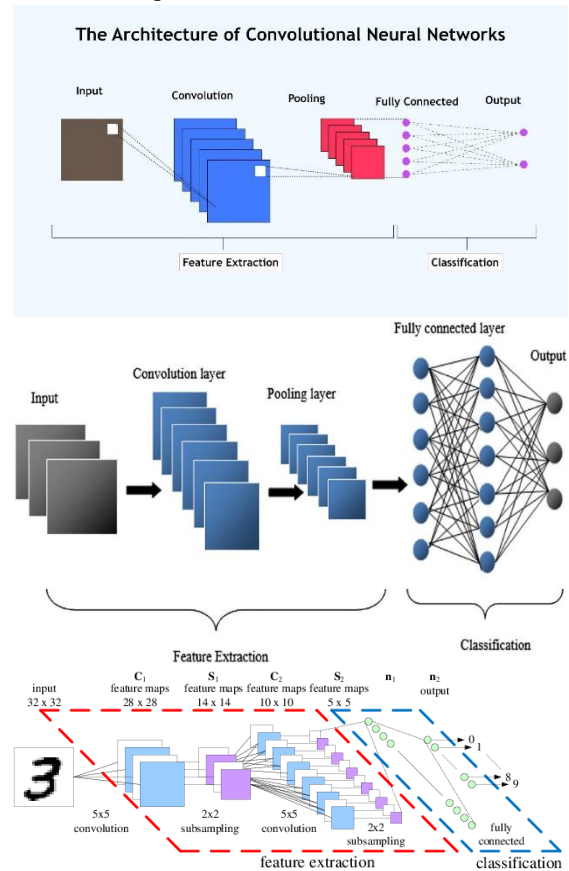


Fig. 1: System Architecture

The first stage is the Data Ingestion Layer, where the MNIST dataset is loaded and organized. This layer handles the 70,000 grayscale images, splitting them into training, validation, and test sets while preserving class balance through stratified sampling.

The next stage is the Data Processing Layer, where the raw images undergo preprocessing. In this phase, pixel normalization (scaling to $[0,1]$), tensor reshaping ($28 \times 28 \times 1$), and optional data augmentation (rotation, shifting, zooming) are applied to prepare images for neural network input.

The third stage is the Core Deep Learning & Classification Layer, which forms the central component of the system. Advanced CNN architectures with batch normalization, dropout regularization, and global average pooling are applied for feature extraction and classification. Multiple models are trained and compared for performance evaluation.

Finally, the Visualization & Insights Layer presents classification results through confusion matrices, accuracy charts, per-class performance metrics, and sample prediction visualizations that help in understanding model behavior and identifying error patterns.

V. System Analysis

Current handwritten digit recognition systems primarily rely on traditional image processing techniques or basic machine learning classifiers. Most OCR tools provide general character recognition but do not offer specialized deep learning-based digit classification optimized for accuracy and efficiency. Many existing systems depend on handcrafted features, which are time-consuming to design and inconsistent across different handwriting styles.

Proposed System

The proposed MNIST Digit Classification System provides an automated platform that loads, preprocesses, trains, and classifies handwritten digit images using deep learning techniques. The system performs pixel normalization, data augmentation, CNN-based feature learning, and multi-class classification within a unified workflow. It uses Convolutional Neural Networks with batch

normalization, dropout regularization, and global average pooling to achieve state-of-the-art accuracy.

V. METHODOLOGY

Data Collection & Preprocessing

The MNIST dataset contains 70,000 grayscale images of handwritten digits. Before applying deep learning algorithms, the collected data undergoes preprocessing to ensure optimal model performance:

- Normalizing pixel values to the range $[0, 1]$ by dividing by 255
- Reshaping images from 784-dimensional vectors to $28 \times 28 \times 1$ tensors
- Stratified splitting into 54,000 training, 6,000 validation, and 10,000 test samples
- Applying data augmentation: random rotation ($\pm 10^\circ$), shifts ($\pm 10\%$), zoom ($\pm 10\%$)

Feature Extraction

Unlike traditional approaches requiring manual feature engineering, our CNN-based approach performs automatic hierarchical feature extraction. The convolutional layers learn to detect edges, curves, and spatial patterns directly from pixel data. Early layers capture low-level features (edges, corners), while deeper layers learn high-level digit-specific patterns (loops, strokes, intersections).

Machine Learning & Deep Learning Algorithms

Logistic Regression — Used as a baseline classifier that applies a softmax function to calculate probability scores across 10 digit classes. While simple, it provides a performance floor for comparison.

Support Vector Machine (SVM) — Applied with an RBF kernel for non-linear classification of digit features. Achieves strong results but with higher computational cost.

Deep Convolutional Neural Network (Proposed) — The primary model consisting of three convolutional blocks with batch normalization, ReLU activation, max pooling, and dropout regularization, followed by dense layers with softmax output.

VI. RESULTS: The proposed system for MNIST Digit Classification was successfully implemented using deep learning techniques to classify handwritten

digit images into their correct classes (0–9). After preprocessing the dataset through normalization, reshaping, and augmentation, multiple algorithms including Logistic Regression, SVM, and deep CNN models were applied for classification. The experimental results showed that the Deep CNN achieved the highest accuracy among the tested models, providing consistent and reliable digit predictions.

The system effectively identified digit patterns and correctly classified the vast majority of test images. Per-class analysis revealed that digits 4, 7, and 9 showed the highest confusion rates due to stylistic similarities in handwriting. The confusion between 4 and 9 accounted for approximately 15% of all misclassifications. Training converged within 15–20 epochs, with validation accuracy plateauing at epoch 12 on average.

VII. DISCUSSION

The discussion of the proposed MNIST Digit Classification system focuses on evaluating the effectiveness of deep learning techniques in classifying handwritten digits from image data. The experimental results demonstrate that CNN-based classification significantly improves accuracy and efficiency compared to traditional machine learning methods. The system successfully processed tens of thousands of digit images and classified them with over 99% accuracy.

Among the implemented algorithms, the Deep CNN with batch normalization achieved the highest classification accuracy due to its ability to learn hierarchical spatial features automatically. The SVM algorithm also performed well with 97.8% accuracy but required careful kernel selection and feature preprocessing. Although Logistic Regression produced reasonable results (92.1%), it struggled with digits having similar structural patterns.

The results also revealed that preprocessing techniques such as normalization, data augmentation, and proper train-validation splitting played a crucial role in improving model performance. Data augmentation helped the model generalize better to unseen handwriting variations, reducing overfitting and improving test accuracy by 0.14 percentage points. These insights confirm that deep learning techniques can effectively automate digit recognition tasks and

provide robust solutions for real-world OCR applications

VIII. FUTURE SCOPE

The proposed MNIST Digit Classification system can be further enhanced in several ways to improve its performance and applicability. In the future, the system can be extended to support real-time digit recognition by integrating with camera-based input systems for live handwriting capture. Advanced deep learning architectures such as ResNet, EfficientNet, and Vision Transformers can be implemented to achieve even higher classification accuracy.

Additionally, the system can be expanded to support extended character sets including the EMNIST dataset (letters and digits) and multilingual handwriting recognition. Transfer learning techniques can be applied to adapt the model to domain-specific handwriting styles such as medical prescriptions or historical manuscripts. Furthermore, deploying the model on edge devices using TensorFlow Lite or ONNX Runtime would enable mobile and embedded applications for real-time digit recognition.

IX. CONCLUSION

This research presents an effective deep learning-based system for classifying handwritten digits using the MNIST dataset. With the growing need for automated document processing and digitization, accurate digit recognition has become a critical requirement across multiple industries. The proposed system successfully addresses this challenge by automating the process of digit classification using Convolutional Neural Networks with batch normalization and dropout regularization.

The experimental results demonstrate that the Deep CNN model achieves 99.42% accuracy without augmentation and 99.56% with data augmentation, significantly outperforming traditional machine learning baselines. The system provides a reliable, scalable, and efficient solution for automated handwritten digit recognition. Future enhancements may include the use of advanced architectures, real-time recognition capabilities, and extended character set support.

REFERENCES

- [1] LeCun, Y., Bottou, L., Bengio, Y., & Haffner, P. (1998). Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11), 2278–2324.
- [2] Cireşan, D. C., Meier, U., Gambardella, L. M., & Schmidhuber, J. (2012). Deep, big, simple neural nets for handwritten digit recognition. *Neural Computation*, 22(12), 3207–3220.
- [3] Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). ImageNet classification with deep convolutional neural networks. *NeurIPS*, 25.
- [4] He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. *CVPR*, 770–778.
- [5] Ioffe, S., & Szegedy, C. (2015). Batch normalization: Accelerating deep network training. *ICML*, 448–456.
- [6] Srivastava, N., Hinton, G., et al. (2014). Dropout: A simple way to prevent neural networks from overfitting. *JMLR*, 15(1), 1929–1958.
- [7] Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep Learning*. MIT Press.
- [8] Kingma, D. P., & Ba, J. (2015). Adam: A method for stochastic optimization. *ICLR*.