International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

HANDWRITTEN DIGIT RECOGNITION TECHNIQUES

Ms. Rimpy[1], Akash Raj[2], Dishan Atwal[3]

Maharaja Surajmal Institute

1. – ABSTRACT :

Handwritten digit recognition remains a critical benchmark problem for evaluating machine learning and computer vision techniques. Deep learning methods, particularly Convolutional Neural Networks (CNNs), have revolutionized this domain, offering exceptional accuracy and automation in feature extraction. This research presents a CNN-based approach, discusses preprocessing techniques like normalization and augmentation, and evaluates the model's performance against traditional classifiers like SVM and k-NN. Our results affirm CNNs' dominance with a test accuracy exceeding 99%, paving the way for broader handwritten recognition applications.

2. - INTRODUCTION

The field of handwritten digit recognition has experienced significant growth, thanks to advances in deep learning. Applications span banking, postal services, and digitization of historical documents. Challenges include varied handwriting styles, overlapping strokes, and noise. This research aims to design a robust, scalable model capable of generalizing across such variations, primarily using the MNIST dataset as a benchmark.

3. – LITERATURE REVIEW

1. LeCun et al. (1998) introduced LeNet-5, achieving high accuracy with CNNs.
2. Simard et al. (2003) proposed elastic distortions to augment training data.
3. Ciresan et al. (2012) leveraged GPU-based DNNs achieving state-of-the-art.
4. Hinton et al. (2006) presented Deep Belief Networks for unsupervised feature learning.
5. Wan et al. (2013) proposed DropConnect, improving generalization.
6. Goodfellow et al. (2013) introduced Maxout Networks.
7. Deng (2012) extensively surveyed MNIST application performance.
8. Sermanet et al. (2013) proposed OverFeat, combining classification and localization.
9. Tang (2013) compared CNNs against linear SVMs for MNIST.
10. Dosovitskiy et al. (2020) proposed Vision Transformers for image recognition, inspiring further exploration.

4. – METHODOLOGY

The methodology involves systematic steps:
1) Data collection via MNIST
2) Data preprocessing with normalization and augmentation
3) CNN architecture with convolutional, pooling, and dense layers
4) Model training with Adam optimizer
5) Evaluation through accuracy, confusion matrix, and loss curves.

5. – FEATURE EXTRACTION AND CLASSIFICATION

CNNs automate feature extraction through convolutional layers that detect edges, textures, and shapes. Pooling layers reduce spatial dimensions, while fully connected layers synthesize learned features to classify inputs into ten-digit classes using softmax activation.

6. – DATA PRE-PROCESSING PHASE

Normalization ensures inputs range between 0 and 1, facilitating faster convergence. Reshaping aligns inputs with CNN expectations (28x28x1). Augmentation techniques like random rotations, shifts, and zooms prevent overfitting and enhance generalization.

7. - DATA COLLECTION PHASE

The MNIST dataset offers a standardized collection of 70,000 handwritten digits. Each sample is a 28x28 pixel grayscale image labeled 0-9. Data is split into 60,000 training and 10,000 testing samples.

8. – EXPERIMENTS AND RESULTS

Dense neural networks achieved 97% accuracy, but CNNs reached 99.2% test accuracy. Confusion matrices showed most misclassifications between similar digits (e.g., 4 and 9). Training used a batch size of 128 over 20 epochs with early stopping to prevent overfitting.

9. - COMPARISON WITH OTHER APPROACHES

SVMs and k-NNs achieved around 96%-97% accuracy but lacked the feature learning efficiency of CNNs. Logistic Regression lagged at ~92%. CNNs not only achieved higher accuracy but also faster inference times post-training.

10. – CONCLUSION AND FUTURE SCOPE

Our study validates the effectiveness of CNNs for handwritten digit recognition. Future work includes lightweight model deployment on mobile devices, multilingual handwriting datasets (Arabic, Chinese), and exploration of Vision Transformers for broader generalization.

11. - REFERENCES

1. LeCun, Y., et al. (1998)...
2. Simard, P., et al. (2003)...
3. Ciresan, D., et al. (2012)...
4. Hinton, G., et al. (2006)...
5. Wan, L., et al. (2013)...
6. Goodfellow, I., et al. (2013)...
7. Deng, L. (2012)...
8. Sermanet, P., et al. (2013)...
9. Tang, Y. (2013)...
10. Dosovitskiy, A., et al. (2020)...
11-20: Additional academic sources from deep learning, MNIST benchmarks, and image classification literature.