International Journal of Computational Intelligence Research.
ISSN 0974-1259 Volume 22, Number 1 (2026), pp. 45-48

https://dx.doi.org/10.37622/IJCIR/22.1.2026.45-48

HAND-WRITTEN CHARACTER RECOGNITION: ADVANCED MACHINE LEARNING TECHNIQUES

Mohamad Hallak*, Vaishnavi Nampally ** and Bhatracharyulu N.Ch**
Department of Statistics, Osmania University, Hyderabad-7, Telangana, India

E-Mail- mohamad.hallak@osmania.ac.in *, nampallyvaishnavi@gmail.com ***and* dwarakbhat@osmania.ac.in **

Abstract

Handwritten character recognition (HCR) remains a crucial application in the domains of Pattern Recognition and Artificial Intelligence. With the advancement of machine and deep learning, various models have been developed to achieve high accuracy in interpreting handwritten characters. This study explores a range of traditional machine learning and advanced deep learning architectures, including CNN, ResNet50, LSTM, Capsule Networks, and a hybrid CNN-LSTM model analyse their performances using classification metrics.

Keywords: Hybrid model, CNN, LSTM.

1. Introduction

Hand-written character and its automatic conversion to text font is one of the challenging task. This technology plays a vital role in various applications, including automated postal sorting, bank cheque verification, form digitization, document archiving, and intelligent text input systems. Many researchers developed various methods for same and developed many software's. For any randomly selected complex hand-written text no software will provide 100% accuracy in conversion. Some traditional machine learning techniques such as Support vector machines (SVM), k-Nearest neighbour (KNN) and Naive Bayes relying on manual feature extraction methods like zoning, projection histograms, or geometric moments to handle handwriting variability. Some advanced machine learning techniques like, Convolutional Neural Networks (CNN), Recurrent Neural Networks (RNN), long-short-term memory (LSTM), Capsule Networks and Residual Networks (ResNet) etc. are significantly advanced in many fields by enabling automatic extraction of hierarchical and spatially invariant features. These models demonstrate robustness against noise, distortions, and style variations, while hybrid models like CNN-LSTM architectures effectively capture both spatial and temporal dependencies.

The Modified National Institute of Standards and Technology (MNIST) dataset collected from from www.kaggle.com which consists of 70,000 grayscale images handwritten digits from 0 to 9, with 60,000 images for training and 10,000 images for testing, constructed from NIST's Special Database-1of 30,000 patterns and 30,000 patterns from Special Database-3 which contain binary images of handwritten digits to train and composed of 5,000 patterns from SD-1 and 5,000 patterns from SD-3. Each image is of size 28×28 pixels and contains a single digit cantered and size-normalized. To align with the scope of recognizing only numeric characters excluding zero (i.e., digits 1 through 9), the dataset was filtered to remove all instances of the digit '0'. This results in a more focused classification problem involving nine classes, representing digits from 1 to 9.

This paper presents to evaluate the performance of advanced deep learning models for hand-written character recognition for text font conversion. It presents to analyse model performance using metrics, accuracy, precision, recall, F1-score.

2. Machine Learning Models

Traditional classifiers such as Support Vector Machines (SVM) [1], k-Nearest Neighbors (KNN) [2], Logistic Regression [3], Random Forest [4], and Naive Bayes [5] have served as foundational approaches. Convolutional Neural Networks (CNNs) [6] have become standard in image-based tasks due to their ability to capture local spatial patterns through convolutional filters. Advanced architectures like Residual Networks (ResNet50) [7] address the degradation problem in deep networks by introducing skip connections, significantly improving recognition accuracy in complex datasets. Recurrent Neural Networks (RNNs), especially Long Short-Term Memory (LSTM) networks [8], model sequential dependencies and have been adapted for handwriting recognition to capture temporal context in strokes. Capsule Networks [9] introduce the concept of capsules to preserve spatial hierarchies between features, resulting in improved robustness to affine transformations and superior performance in HCR. Hybrid models combining CNNs and LSTMs (CNN-LSTM Hybrid) [10] exploit the strengths of spatial and temporal feature learning, achieving state-of-the-art results by effectively modelling both character shapes and sequence dependencies.

3. Model Fiiting & Its Analysis

The pre-processing pipeline involves normalizing pixel intensities to a [0,1] range to standardize inputs and facilitate model convergence. Target labels are transformed using one-hot encoding to enable effective multi-class classification. Data augmentation methods, including controlled rotations, zooming, and horizontal shifts, are applied to enhance dataset variability and reduce overfitting. These augmentations improve the model's ability to generalize across diverse handwriting styles and positional variations. Collectively, these steps contribute to improved model accuracy and robustness. The accuracy, precision, recall, F1 score for the dataset is given below

Table-1: Comparison of performance metrics of various learning models				
Model	Accuracy (%)	Precision (%)	Recall (%)	F1 Score (%)
SVM	91.4	91	91.3	91.1
KNN	93.2	93	93.2	93.1
Logistic Regression	92.1	91.5	92.3	91.7
Random Forest	96.8	96.7	96.9	96.8
Naive Bayes	85.6	84.9	85.4	85.1
CNN	98.5	98.4	98.5	98.4
ResNet50	99.6	99.6	99.6	99.6
LSTM	96.2	96.0	96.3	96.1
Capsule Network	98.7	98.7	98.8	98.7
CNN-LSTM Hybrid	99.1	99.1	99.1	99.1

Remarks

1. **Traditional simple algorithms** like **SVM**, **KNN**, and **Logistic Regression** demonstrated moderate performance, with accuracies ranging from **91.4% to 93.2%**. These models rely heavily on feature engineering and may struggle with the variability inherent in handwritten characters.
2. The **Random Forest** classifier achieved a notable accuracy of **96.8%**, indicating its effectiveness in handling complex datasets through ensemble learning.
3. With an accuracy of **85.6%**, **Naive Bayes** underperformed relative to other models, likely due to its strong independence assumptions, which may not hold true for image data.
4. Advanced models such as **CNN**, **ResNet50**, **Capsule Network**, and **CNN-LSTM Hybrid** outperformed traditional methods, achieving accuracies between **98.5% and 99.6%**. These models excel at capturing spatial hierarchies and complex patterns in image data.
5. **CNN-LSTM Hybrid** model achieved **99.1%** accuracy, showcasing the benefits of combining spatial feature extraction with temporal sequence modelling for improved recognition performance.

Conclusions

The comparative analysis underscores the superior performance of deep learning models over traditional machine learning approaches in handwritten character recognition tasks. While traditional models offer simplicity and interpretability, their performance is limited by their reliance on manual feature extraction and assumptions that may not align with the complexities of image data. In contrast, deep learning models, particularly **ResNet50** and **CNN-LSTM Hybrid**, demonstrate exceptional accuracy and robustness by automatically learning hierarchical features and capturing intricate patterns. These findings suggest that for applications requiring high precision in character recognition, adopting advanced deep learning architectures is highly advantageous.

Acknowledgements
The authors are thankful to the editorial board for improving the manuscript. The first author is thankful for Osmania University for their continues support to pursue this research work under Ph.D. program and the second author is thankful to DST-INSPIRE for sanctioning of fellowship to pursue the work.

References

[1] Breiman, L. (2001): "Random Forests", *Machine Learning*, 45(1): 5–32.
[2] Cortes, C., & Vapnik, V. (1995): "Support-Vector Network", *Machine Learning*, 20(3): 273–297.
[3] Cover, T., & Hart, P. (1967): "Nearest Neighbor Pattern Classification", *IEEE Transactions on Information Theory*, 13(1): 21–27.
[4] He, K., Zhang, X., Ren, S., & Sun, J. (2016). "Deep Residual Learning for Image Recognition", Conference on Computer Vision and Pattern Recognition, 770–778.
[5] Hochreiter, S., & Schmidhuber, J. (1997): "Long Short-Term Memory", *Neural Computation*, 9(8):1735–1780.
[6] Hosmer, D. W., Lemeshow, S., & Sturdivant, R. X. (2013): "*Applied Logistic Regression"*, Wiley.
[7] LeCun, Y., Bottou, L., Bengio, Y., & Haffner, P. (1998): "Gradient-Based Learning applied to Document Recognition", *Proceedings of the IEEE*, 86(11): 2278–2324.
[8] Rish, I. (2001): "An Empirical Study of the Naive Bayes Classifier", Proceedings of *International Joint Conference on Artificial Intelligence*, 41–46.
[9] Sabour, S., Frosst, N., & Hinton, G. E. (2017): "Dynamic Routing Between Capsules", Proceedings of Neural Information Processing Systems, 3856–3866.
[10] Sharma, M. (2024): "Handwritten Text Recognition using Hybrid CNN-GRU Model and CNN-LSTM Model on Parzival Database: A Novel Approach", In International Conference on Signal Processing and Computer Vision , Atlantis Press, 383-390.