State-of-the-Art Review of Handwritten Digit Recognition

T. Kavitha[1], Karthikeyan Alamuthu[2], Dr. T. Sathis Kumar[3]

[1]Assistant Professor, Department of Computer Science & Engineering, M.A.M College of Engineering and Technology, Siruganur, Trichy, India
[2]Reserach Scholar, Department of Computer Science & Engineering, School of Engineering& Technology, Dhanalakshmi Srinivasan University, Trichy, India
[3]Associate Professor, Department of Computer Science & Engineering, School of Engineering & Technology, Dhanalakshmi Srinivasan University, Trichy, India

ABSTRACT

Handwritten digit recognition is the task of enabling machines to identify and interpret human handwritten digits. The machine has a challenging task because to the imperfect nature of handwritten digits, which exhibit variations between individuals and can be written in numerous different styles. Handwritten digit recognition is a common multiclass classification issue that is typically integrated into the software of mobile banking applications and classic automated teller machines. Its purpose is to enable customers to automatically deposit paper checks. Each data class has many handwritten renditions of a single digit, represented as images. Handwriting character recognition has gained popularity as a research field due to advancements in technology, including the availability of handwriting recording devices and high-performance mobile computers. Handwritten digit recognition has been extensively studied in the realm of deep learning as well. Digit recognition is utilized in several applications such as number plate identification, postal mail sorting, and bank check processing. Handwritten digit recognition has numerous issues due to the diverse writing styles of individuals, as it does not utilize optical character recognition. This review paper offers a comprehensive analysis of the latest algorithms and strategies employed in the field of handwritten digit recognition.

Keywords: handwritten digit recognition, machine learning, artificial intelligence, deep learning, machine learning models

INTRODUCTION

Handwritten digit recognition refers to the capacity of a computer system to accurately identify and interpret handwritten inputs, such as digits and characters, from diverse sources such as emails, documents, photographs, and letters. This subject has been the focus of scientific investigation for several decades. Significant advancements have been made in the field of pattern recognition in recent years. Remarkable achievements have been made in the field of recognizing handwritten digits. The rapid advancement in this field can be attributed to several factors, namely the widespread availability of affordable high-performance computers, the development of novel algorithms optimized for these computers, and the accessibility of extensive character databases for training and evaluation purposes. Handwritten digit recognition is a prominent subject in the field of Optical Character Recognition (OCR) applications and pattern classification/learning research. Digit recognition is applied in various OCR applications such as postal mail sorting, bank check processing, and form data entering. The performance of digit recognition is vital for the overall performance of these applications, as it directly affects both accuracy and speed. The challenge of recognizing handwritten digits is often used as a benchmark in pattern classification and machine learning communities to evaluate classification performance (Liu et al., 2003).

Due to the swift advancement of electronic information, computer input has gained widespread usage. However, handwriting remains an indispensable method for individuals to convey information. Handwriting recognition, which combines handwritten characters and computer input, has gained increasing attention due to its practicality. Handwriting recognition technology serves as the foundation for handwriting interpolation and handwriting identification. Over the last ten years, advancements in machine learning and pattern recognition have expanded the capabilities of handwriting recognition systems. These advancements include the use of artificial neural networks (ANN), support vector machine (SVM), modified quadratic discriminant function (MQDF), and hidden Markov model (Li et al., 2020).

RATIONALEFOR THE RESEARCH

The field of handwritten digit recognition (HDR) is of great importance for both academic and commercial purposes. The problem of HDR is a complex issue that has been the subject of research by utilizing machine learning methods. HDR is designed to receive and interpret handwriting input in the form of images or physical documents. Extracting text from real photos is a challenging operation due to significant changes in font size and shape, texture, backdrop, and other factors. Handwriting character recognition is widely utilized in many study domains such as bank check processing, automatic number-plate recognition, postal address verification from envelopes, ID card recognition, and zip code identification. The essential stages of character recognition (CR) include segmentation, feature extraction, and classification. The rapid advancement in character recognition is a testament to the progress made in learning algorithms and the accessibility of extensive databases (Ali et al., 2019). Several databases such as MNIST, CENPARMI, CEDAR, and others have facilitated advanced study in the domain of pattern recognition. Among these tasks, MNIST is often regarded as the benchmark for pattern recognition. Various classifiers, such as restricted Boltzmann machines (RBMs) and neural networks (NN), have been evaluated using the MNIST dataset. Handwritten digit recognition using CNN as a classifier has recently gained attention in the field of deep learning due to its wide range of applications. Deep learning is a rapidly expanding field in machine learning that offers superior performance in character recognition and pattern recognition. This is mostly due to its exceptional feature extraction capabilities and its effectiveness as a classifier.

Several schemes/algorithms have been presented to address the demanding challenge of handwritten digit recognition, which is considered one of the most difficult problems in this sector. Automated recognition of handwritten numbers poses numerous hurdles for researchers due to the diverse orientations and styles that these digits can have.

Categories of Handwritten Digit Recognition

Handwritten digit recognition is categorized into online recognition and offline recognition. Online recognition involves utilizing computer algorithms to identify and interpret handwritten characters based on their stroke patterns and arrangement. This technology has reached a high level of maturity in terms of theoretical research. Offline recognition refers to the computer's ability to recognize characters that are handwritten on paper. The stroke and stroke order are not considered as the foundation. Picture recognition only offers limited information; hence it continues to encounter significant obstacles.

Deep Learning

Deep learning enables the training of computational models consisting of numerous layers of processing, which can learn abstract representations of data at various levels. These techniques have significantly enhanced the current level of performance in speech recognition, visual object recognition, object detection, as well as in other fields such as drug discovery and genomics. Deep learning utilizes the back propagation technique to uncover complex patterns in extensive data sets. This algorithm guides the machine in adjusting its internal parameters, which are responsible for computing the representation in each layer based on the representation in the preceding layer. Deep convolutional neural networks have revolutionized the field of image, video, speech, and audio processing, while recurrent neural networks have provided valuable insights into analysing sequential data, such as text and speech. Deep-learning methods are a type of representation-learning methods that involve many levels of representation. These methods use basic but non-linear modules to change the representation at each level, starting from the raw input, into a little more abstract representation at a higher level. By combining a sufficient number of these transformations, it is possible to learn highly intricate functions. In classification tasks, the upper layers of representation enhance the significant features of the input that are necessary for differentiation and reduce the impact of unimportant differences. An image is represented as an array of pixel values.

The features learned in the initial layer of representation usually indicate the presence or absence of edges at specific orientations and positions in the image. The second layer commonly identifies motifs by identifying specific configurations of edges, irrespective of minor deviations in the placements of the edges. The third layer is responsible for organizing motifs into bigger combinations that represent specific components of recognizable things. Subsequent layers then identify objects by recognizing these combinations of elements. Deep learning is characterized by the fact that the layers of features are not created by human engineers, but rather acquired from data through a versatile learning process. Deep learning is achieving significant progress in solving problems that have proven challenging for the artificial intelligence industry for a long time (LeCun et al., 2015). It has proven to be highly effective in identifying complex patterns in data with a large number of dimensions, making it useful in various fields including as science, industry, and government. Several classification algorithms utilizing Machine Learning have been created and employed, such as K-Nearest Neighbours, SVM Classifier, and Random Forest Classifier, among others. However, despite their 95% accuracy, these approaches are insufficient for real-world applications. The precision in these applications is of utmost importance, but these methodologies fail to deliver the necessary level of precision due to a limited understanding of the task's topology.

METHODOLOGY

In this study, a complete State-of-the-Art (SotA) literature review is favored over a standard general review because of its replicable, transparent, unbiased, rigorous, and scientific methodology. State-of-the-Art reviews primarily focus on recently released material to assess present-day concerns. It involves the process of identifying, selecting, evaluating, extracting, and combining research studies that are accessible in the literature. Well-crafted state-of-the-art review articles provide a complete summary to scholars in a certain topic (Celeste, 2024). This study utilizes the state-of-the-art (SotA) technique, as explained by Barry et al. ., (Barry et al., 2022), to conduct the review.

SotA Process

The belowsection summarizes the 6-step process of SotA technique used in this review.

Stage 1: Establish the primary research issue and area of investigation.

During Stage 1, the researcher(s) develop a basic description of the issue that needs to be summarized. They must also decide which sector of knowledge (and/or practice) the search will focus on.

Step 2: Establish the time line

In this stage, the task is to establish the specific timeframe that will be considered as the state of the art (SotA) for the issue being summarized. The researcher(s) should conduct a comprehensive review of the literature, examining a wide range of sources to gain a deep understanding of the historical progression of knowledge on the topic, including the pivotal moments that have influenced the present perspectives on the subject.

Stage 3: Refine research question(s) to align with the specified period

After analyzing the findings in Stage 2, it is probable that the researcher(s) will have to modify their initial description of the issue that needs to be described. The research question(s) that structure the state-of-the-art review are determined in Stage 3.

Stage 4: Formulate a search strategy to locate pertinent articles

In Stage 4, the researcher(s) develops a search strategy to identify the literature that will be included in the SotA review. The researcher(s) must ascertain which literature databases encompass papers from the specific subject of interest.

Stage 5: Analysis

The literature analysis will incorporate the researcher(s)' subjective perspectives, while being guided by the fundamental principles of inductive research.

Stage 6: Reflexivity

It refers to the ability to reflect on one's own thoughts, actions, and beliefs. Considering the subjective nature of a state-of-the-art review, it is crucial for the article to provide a deep understanding of the researcher(s)' subjectivity. This reflexivity description should clearly explain how the personal perspectives of the researcher(s) influenced their understanding and analysis of the data.

Literature Search

The articles to be included for the review were scrutinized using the SotA technique. During the identification phase, the ScienceDirect (www.sciencedirect.com) website was utilized to search for journal articles. This website offers access to a comprehensive bibliographic database of scientific and medical publications published by Elsevier. 2024 was selected as the year in which research articles were published.

The search was limited to the journal articles that were indexed in the database. It guarantees the quality of the journal articles included in the review and indexes a substantial portion of peer-reviewed literature. The keywords were chosen in accordance with the review's scope. The review's primary objective is to comprehend the latest techniques uded in Handwritten Digit Recognition. The abstract, keywords, and title sections of the paper were searched for keywords. A total of 54 articles were obtained in the computer science field. Six of the articles were review articles, while 46 were research articles, one was mini review and one was a software application.

3. Publication Details

The below table (Table 1) summarizes the publication details of all the 54 journals that were retrieved from the ScienceDirect Database.

Table 2 ***List of54 Publication Details Retrieved:***

S.No	Journal Title	Number of Research Articles
1.	Procedia Computer Science	9
2.	Neurocomputing	7
3.	Pattern Recognition	5
4.	Intelligent Systems with Applications	4
5.	Future Generation Computer Systems	3
6.	Machine Learning with Applications	3
7.	Chip	3
8.	Knowledge-Based Systems	2
9.	Computers & Security	2
10.	Journal of King Saud University - Computer and Information Sciences	2
11.	Computer Aided Geometric Design	1
12.	Pattern Recognition Letters	1
13.	Image and Vision Computing	1
14.	Computer Networks	1
15.	Information Fusion	1
16.	Ad Hoc Networks	1
17.	Computer Science Review	1
18.	International Journal of Cognitive Computing in Engineering	1
19.	Patterns	1
20.	AI Open	1
21.	Blockchain: Research and Applications	1
22.	High-Confidence Computing	1
23.	Natural Language Processing Journal	1
24.	Journal of Information and Intelligence	1

REVIEW OF RESEARCH ARTICLES

Out of the 54 research articles, only 4 articles qualified for the review. This section summarizes the research work of those 4 articles along with their key findings.

Research conducted by (Raquib et al., 2024) introduced a VashaNet model for the purpose of recognizing basic letters in handwritten Bangla. The VashaNet model utilized a 26-layer deep convolutional neural network (DCNN) structure comprising of nine convolutional layers, six max pooling layers, two dropout layers, five batch normalization layers, one flattening layer, two dense layers, and one output layer. The experiment was conducted using two datasets: a primary dataset called CMATERdb 3.1.2, which contained 5750 photos. This dataset was used for training and assessing the model. The proposed character recognition model achieved a test accuracy rate of 94.60% for the primary dataset and 94.43% for the CMATERdb 3.1.2 dataset.

Ravi, (Ravi, 2024) created three CNN models, namely THAC-CNN1, THAC-CNN2, and THAC-CNN3, with the purpose of identifying and categorizing Tamil Handwritten Alphabets. This model utilized both a benchmark dataset and a bespoke dataset, consisting of a total of over 2800 images depicting various Tamil alphabets. These datasets were augmented using several strategies to enhance the data.THAC-CNN1 demonstrated a training dataset accuracy of 97% and a test dataset accuracy of 92.5%, surpassing the 72% and 73.5% accuracy achieved by pre-trained models on their respective datasets.

Bappi et al., (Bappi et al., 2024) conducted a study on the utilization of Convolutional Neural Networks (CNNs) to identify Bangla handwritten characters. The study specifically aimed to expand the range of identified character categories. In order to accomplish this, a new and demanding dataset for handwriting recognition was provided. This dataset was gathered from a large number of pupils. This research developed a novel approach called BNVGLENET, which is a convolutional neural network, to recognize handwritten Bangla letters. The approach modifies the LeNet-5 model and combines it with the VGG architecture. This research obtained an exceptional recognition accuracy of 98.2% on our specialized vowel-consonant class during testing, and attained a high accuracy of 97.5% on the bespoke individual class.

Work of Lamaakal et al., (Lamaakal et al., 2024) showcased a cutting-edge TinyML model designed specifically to identify Arabic hand motions performed in the air, with a particular focus on accurately classifying Arabic digits. The main foundation of this model is the incorporation of Convolutional Neural Networks (CNNs), highlighting their remarkable significance in attaining an impressive accuracy rate of 93.8% in classifying various Arabic Numbers gestures.

DISCUSSION AND CONCLUSION

They key findings of the State-of-the-Art review of applications is summarized in this section. The findings show that in the state-of-the-art techniques deep learning is widely applied to improve the accuracy of hand written digit recognition. The below chart (Figure 1) depicts the comparison of accuracy of models reviewed.

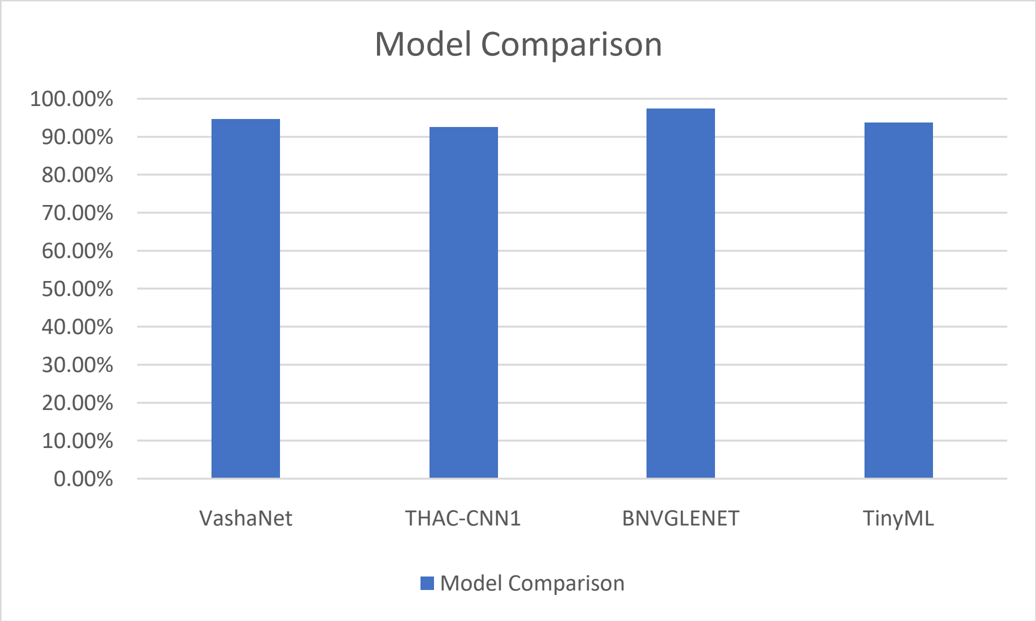

Figure 1 ***Comparison of accuracy of Models***

It is evident from the review that in recent techniques researchers tweak the existing architecture and add more handwritten images to test their models. The new techniques are then compared with other techniques or with standard architectures to evaluate the model performance. Researchers highly rely on accuracy of the model for evaluation. From this review it was found that BNVGLENET, a modified architecture of LeNet-5 and VGG gave highest accuracy of 98.2%. Thus, it can be concluded that deep learning architectures promise to give better accuracy for handwritten digit recognition.

REFERENCES

[1]. Ali, S., Shaukat, Z., Azeem, M., Sakhawat, Z., Mahmood, T., & ur Rehman, K. (2019). An efficient and improved scheme for handwritten digit recognition based on convolutional neural network. *SN Applied Sciences*, *1*(9), 1125. https://doi.org/10.1007/s42452-019-1161-5

[2]. Bappi, J. O., Rony, M. A. T., & Islam, M. S. (2024). BNVGLENET: Hypercomplex Bangla handwriting character recognition with hierarchical class expansion using Convolutional Neural Networks. *Natural Language Processing Journal*, *7*, 100068. https://doi.org/10.1016/j.nlp.2024.100068

[3]. Barry, E. S., Merkebu, J., & Varpio, L. (2022). How to Conduct a State-of-the-Art Literature Review. *Journal of Graduate Medical Education*, *14*(6), 663–665. https://doi.org/10.4300/JGME-D-22-00704.1

[4]. Celeste, T. (2024, June 5). *Library Guides: Which review is that? A guide to review types: State of the Art Review*. https://unimelb.libguides.com/whichreview/stateoftheartreview

[5]. Lamaakal, I., Makkaoui, K. E., Ouahbi, I., & Maleh, Y. (2024). A TinyML Model for Gesture-Based Air Handwriting Arabic Numbers Recognition. *Procedia Computer Science*, *236*, 589–596. https://doi.org/10.1016/j.procs.2024.05.070

[6]. LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep Learning. *Nature*, *521*, 436–444. https://doi.org/10.1038/nature14539

[7]. Li, Z., Wu, Q., Xiao, Y., Jin, M., & Lu, H. (2020). Deep Matching Network for Handwritten Chinese Character Recognition. *Pattern Recognition*, *107*, 107471. https://doi.org/10.1016/j.patcog.2020.107471

[8]. Liu, C.-L., Nakashima, K., Sako, H., & Fujisawa, H. (2003). Handwritten digit recognition: Benchmarking of state-of-the-art techniques. *Pattern Recognition*, *36*(10), 2271–2285. https://doi.org/10.1016/S0031-3203(03)00085-2

[9]. Raquib, M., Hossain, M. A., Islam, M. K., & Miah, M. S. (2024). VashaNet: An automated system for recognizing handwritten Bangla basic characters using deep convolutional neural network. *Machine Learning with Applications*, *17*, 100568. https://doi.org/10.1016/j.mlwa.2024.100568

[10]. Ravi, J. (2024). Handwritten alphabet classification in Tamil language using convolution neural network. *International Journal of Cognitive Computing in Engineering*, *5*, 132–139. https://doi.org/10.1016/j.ijcce.2024.03.001