

Real time handwriting recognition system using CNN algorithms

Maryam Yaseen Abdullah^{1,*} 

¹Department of computer, College of education for women, University of Baghdad, Iraq

*Corresponding Author: Maryam Yaseen Abdullah

DOI: <https://doi.org/10.31185/wjcm.157>

Received: July 2023; Accepted: Septemehr 2023; Available online: Septemehr 2023

ABSTRACT: The growing use of digital technologies across various sectors and daily activities has made handwriting recognition a popular research topic. Despite the continued relevance of handwriting, people still require the conversion of handwritten copies into digital versions that can be stored and shared digitally. Handwriting recognition involves the computer's strength to identify and understand legible handwriting input data from various sources, including document, photo-graphs and others. Handwriting recognition pose a complexity challenge due to the diversity in handwriting styles among different individuals especially in real time applications. In this paper, an automatic system was designed to handwriting recognition using the recent artificial intelligent algorithms, the conventional neural network (CNN). Different CNN models were tested and modified to produce a system has two important features high performance accuracy and less testing time. These features are the most important factors for real time applications. The experimental results were conducted on a dataset includes over 400,000 handwritten names; the best performance accuracy results were 99.8% for SqueezeNet model.

Keywords: Artificial intelligence, handwriting processing, deep learning, CNN models.



1. INTRODUCTION

Over the past two decades, there has been significant research into handwriting recognition, motivated primarily by the fact that humans frequently communicate using handwritten documents, in addition to spoken language. This technology has been put into practical use in a variety of fields, including postal automation, bank check processing, digitization of handwritten documents, and assisting visually impaired readers. It is worth noting that during this time, handwriting recognition systems for character recognition, word spotting, and recognition have made significant advances [1] [2].

Designing and producing an accurate and fast hand-writing system became avital task in many fields and application. An excellent illustration of this is the use of automated processing systems in banks for check processing. Without these systems, banks would need numerous employees who may not be as effective as computerized processing systems.

The development of handwriting character recognition falls under the domains of computer vision, pattern recognition, and artificial intelligence [3, 4]. Its objective is to create computer software capable of identifying and extracting characters from various sources, such as images, paper documents, and screens, and converting them into machine-readable formats. Optical scanning or intelligent word recognition may be utilized to obtain handwriting sources from paper documents, while the movement of a pen tip on a screen may be detected by a designed system. Handwriting recognition can be classified into two categories: offline recognition and real time recognition.

In offline recognition, letter codes are obtained by extracting text or characters from an image, which can be used by a computer [5]. This method involves obtaining digital data from a static representation of handwriting, where the system reads and converts the handwriting document into a digital format. On the other hand, real time recognition involves detecting or converting characters as they are being written on a specialized screen [6, 7]. The system identifies characters

and words by detecting the movement of the pen tip. Various techniques and methods are utilized to ensure that computer systems can accurately recognize characters from handwriting images and documents [8].

Artificial Intelligent algorithms can be utilized to create handwriting recognition systems [4], as the human brain has the capacity to recognize handwriting objects, such as characters, digits, and letters [5]. Nonetheless, human interpretation of hand-writing can be subjective and biased [8], while computer systems are impartial and capable of tackling labor-intensive and challenging tasks. To develop effective handwriting recognition systems, it is crucial to understand how humans perceive and interpret handwriting [8].

The human visual system plays a crucial role in reading handwriting characters, words, letters, and digits, but this process is not as simple as it may seem. Although the brain unconsciously processes information based on prior learning, people may not realize the complexity of recognizing handwriting [6, 9]. Creating a computer system that can accomplish this task is a daunting challenge. Artificial intelligent algorithms, which emulate the functioning of the human brain in a simplified form, are the most effective approach to developing handwriting recognition systems that can match or exceed human performance. As handwriting styles vary among individuals and some are more challenging to read than others, reading handwritten documents can be tedious and time-consuming [8]. Deep learning is ideal for creating such systems because they can extract meaning from complex data and recognize patterns that may be difficult to identify using other methods [7].

The primary aim of this paper is to design a model that uses Convolution Neural Network to identify the person from handwriting digits, characters, and words from images in real time applications.

2. AIMS AND OBJECTIVES

2.1 THE RESEARCH OBJECTIVES

The purpose of this research is to address the following questions:

- What methods and techniques are utilized to recognize handwriting characters?
- How can artificial Convolutional Neural Network (CNN) improve the effectiveness of handwriting recognition systems in real time applications?

2.2 THE RESEARCH CONTRIBUTION

The primary contribution of this research is to create an expert system that utilizes Convolutional Neural Network (CNN) methodology for recognizing Handwriting characters. The study also aims to tackle the problem of low accuracy in Handwriting character recognition systems by developing a system that can efficiently recognize Handwriting characters and words from image media. additionally, the research intends to explore and exhibit the effectiveness of CNN technology in constructing effective handwriting character recognition systems for real time applications.

2.3 THE PAPER ORGANIZATION

The paper is categorized as follows. The related work is explained in section 2. In section 3, the proposed real time handwriting recognition system is described in details. The experimental results and discussion of the proposed system are given in section 4. In section 5, this paper concludes is recorded.

3. RELATED WORK

Extensive research has been conducted on handwriting recognition in the past, and there are comprehensive reviews available on the topic. Despite previous research, the pursuit for an improved and more precise method of handwriting recognition persists to this day.

In traditional machine learning, methods such as SVM and MLP typically use shallow structures with limited computing units and sample diversity. As a result, their performance and generalization ability are often inadequate for complex classification problems involving objects with rich meanings. However, recently developed Convolutional Neural Networks (CNNs) have become widely used in the field of image processing due to their effectiveness in recognizing and classifying images. They have shown remarkable improvements in the accuracy of various machine learning tasks and have become a versatile and powerful deep learning model. Since the introduction of CNN models in deep learning, significant progress has been made in various large-scale recognition tasks in computer vision [1] [2] [9].

In [10], a model based on a fusion strategy was proposed to recognize handwritten Arabic characters in various font types, including SH Roqa, Farsi, Naskh, Igaza, and multi-fonts. They also presented an automatic selection technique for classifiers and features to achieve good segmentation and recognition accuracy. To further improve the accuracy, they

proposed strategies involving Hough Transform and skeletonization methods, as well as modeling a Gaussian mixture framework. Additionally, the authors applied a skew correction strategy to detect and correct skew in the text. In their method, they achieved excellent results for touching characters by using a template matching approach for ligatures that appear among closed consecutive characters or within open characters.

According to the findings reported in reference [11], it is recommended for models to use word-embedding instead of the bag-of-n-grams method. Building on this, a different technique described in reference [12] utilized a ConvNet to calculate the occurrence of n-grams in specific spatial regions of a word present in input images. This approach generated a frequency-based profile, which was then compared to profiles of known words in a dictionary. The results demonstrated an effective attribute-based word-encoding scheme.

The paper [13] proposed a word-spotting mechanism that uses a region proposal network to encode regional features into a distributed word-embedding space for searches. They used a training method based on Connectionist Temporal Classification (CTC) criterion, which was proposed in [14] for training RNNs.

Shi et al. [15] used image features produced by a ConvNet as input to a re-current network such as LSTM [10] or MDL-STM [16] to transcribe words. To improve detection accuracy, the authors in [17] introduced an attention mechanism based on affine transformations to spatially adjust the original images before per-forming sequence-to-sequence transcription. To extend the original dataset, most of the aforementioned methods require different preprocessing techniques, as seen in [17, 18]. Wang and colleagues [19] proposed an adversarial approach that leverages a generator to create challenging examples by introducing occlusions and spatial deformations in the feature space. This technique compels the detector to learn and adjust to infrequent and unusual deformations in real-world input data.

In spite of the state of art researches worked dramatically on designing accurate handwriting system using CNN algorithms, they do not take into account the recognition time which is an important factor in real time applications.

4. THE PROPOSED SYSTEM

In this research an automatic handwriting recognition system is proposed using CNN algorithms. This system is designed to achieve two points, increase handwriting recognition accuracy and reduce the recognition time for real time application. Figure 1 illustrates the proposed system basic blocks, which will have explained in details in the next subsections.

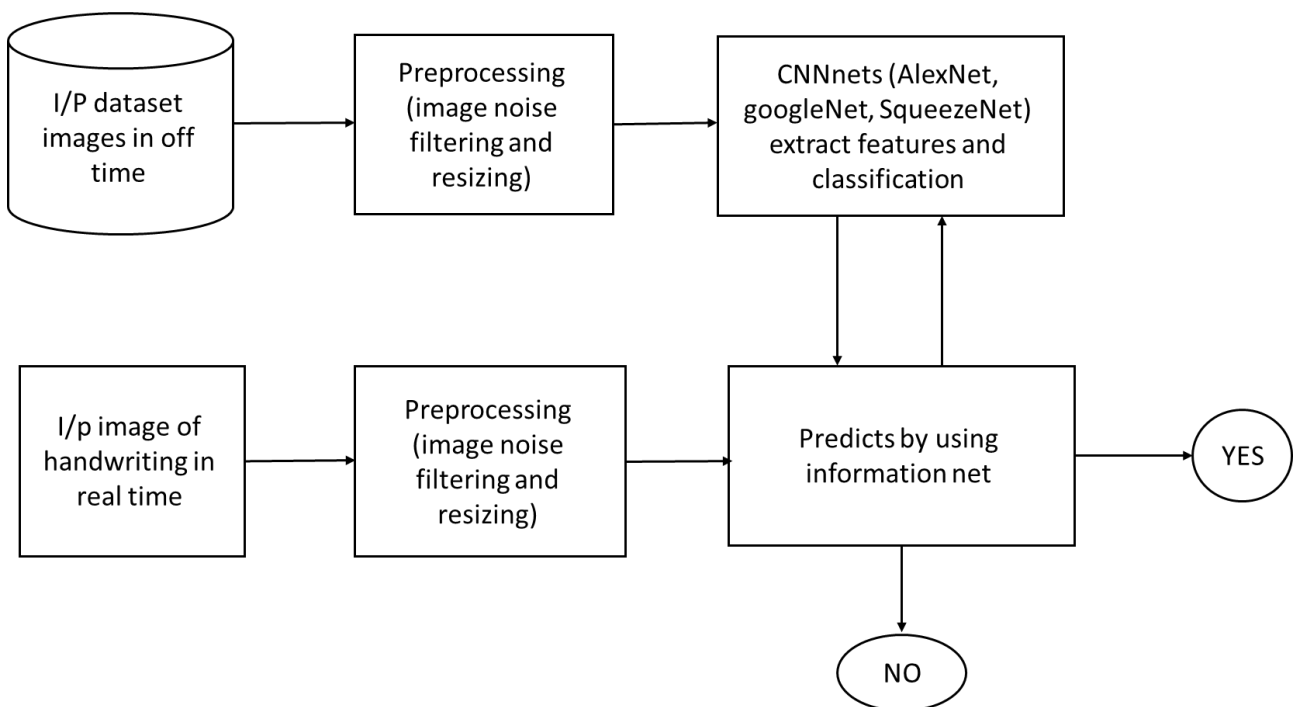


FIGURE 1. The proposed system

4.1 DATASET

This research experiments are connected on a dataset [20] includes over 400,000 handwritten names obtained from charity initiatives. While image processing technologies are often used to convert characters in scanned documents to digital formats, recognizing handwritten characters can still be challenging for machines due to the vast range of individual writing styles. The dataset includes 206,799 first names and 207,024 surnames, which were divided into training (331,059), testing (41,382), and validation (41,382) sets. Figure 2 represents an example of handwriting images.



FIGURE 2. The dataset examples

4.2 PREPROCESSING

In recognition systems, preprocessing is a crucial initial step that can have a significant impact on recognition accuracy. This chapter presents a novel prepro-cessing method for real time handwriting. The proposed approach involves removing stroke hooks using a length threshold with a changed-angle threshold, followed by noise filtering using a smoothing technique that combines the Cubic Spline and equal-interpolation methods. Finally, the handwriting image size is normalized and resized [21].

4.3 DEEP LEARNING USING (CNN)

Deep learning (DL) is implemented using several types of convolutional neu-ral networks (CNNs) [22]. This section presents a selection of DL-CNN architectures that are advantageous due to their reduced requirement for image pre-processing, inherent feature extraction, and classification capabilities.

4.3.1. AlexNet

The CNN architecture known as AlexNet was created in 2012 and contains 61 million learnable parameters. It utilizes max-pooling to achieve down-sampling and typically consists of five convolutional layers and five corresponding ReLU layers. Five max-pooling layers are then added, followed by two fully connected layers. The kernel of filters in AlexNet ranges in size from 11x11 to 3x3 for the five convolutional layers [23]. Figure 3 provides a visual representation of this configuration.

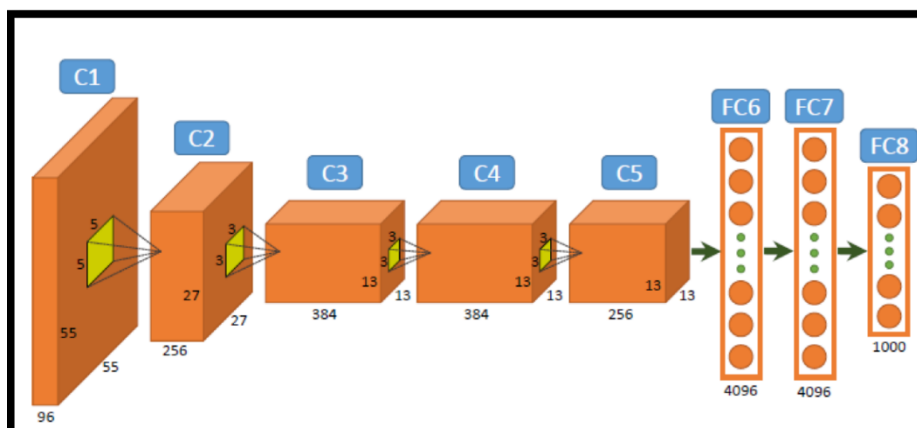


FIGURE 3. AlexNet configuration [23]

4.3.2. GoogleNet

Constructed in 2015, this network contains only five million learnable parameters yet outperforms AlexNet despite being twelve times smaller. It is based on the Inception architecture, which incorporates parallel filters in each layer known as the inception module to increase the number of units [23] [24]. These filters consist of sizes (1x1), (3x3), and (5x5) for every convolution layer. Figure 4 provides a visual representation of this configuration.

4.3.3. SqueezeNet

SqueezeNet is one of the most significant models in image classification due to its unique characteristics and highly accurate results in comparison to other mod-els. Created in 2016, SqueezeNet contains less than one million weights, which is significantly smaller than many other models. It consists of 10 layers, beginning with a convolution layer, followed by 8 fire modules, and ending with another convolution layer. SqueezeNet achieves nearly the same accuracy as AlexNet but with 50 times fewer parameters, making it highly efficient for training due to its small number of weights. This distinctive feature simplifies tasks and speeds up the training process considerably [25]. Figure 5 provides a visual representation of this model.

5. EXPERIMENTAL RESULTS

In this work, three CNNs are utilized: GoogleNet, AlexNet, and SqueezeNet. Each network varies in terms of the number of layers, input image size, memory requirements, output accuracy, and training/testing time.

For AlexNet model, images need to be resized to 227x227x3 pixels, and there are a total of 25 layers in Matlab. To achieve accurate results, the maximum number of epochs is set to 20 for complex background images and 10 for simple black background images. The number of epochs is increased because lower values result in poor accuracy. The Initial Learn Rate is set to 10-4, and the data is divided using a randomized 0.7 ratio for training and a 0.3 randomized ratio for validation.

To use the GoogleNet model, images must be resized to 224x224x3 pixels and processed by 144 across all layers in Matlab. For optimal results, the maximum number of epochs is limited to 10 for complex background images and 5 for simple black background images. The Initial Learn Rate is 10-4, and data is divided using a randomized 0.7 ratio for training, and a 0.3 randomized ratio for validation.

SqueezeNet is a neural network used in computer vision to create a smaller model with fewer parameters, making it easier to fit into computer memory and communicate across networks. It comprises 10 layers, starting with a convolution layer followed by 8 fire modules and ending with another convolution layer. Despite having 50 times fewer parameters than AlexNet, it achieves almost the same accuracy. This is due to the replacement of most 3x3 filters with 1x1 filters, which reduces parameters by a factor of nine. The number of input channels is also decreased using squeeze layers with 3x3 filters, and a down-sample is performed late in the network, which results in large activation maps for the convolution layers [74].

The recorded results are presented in Table 1 which shows the accuracy of the classifier methods and the time taken for the prediction step for validation data. The accuracy of the classifier in (1) is measured by the ratio of correct predictions to total predictions, which is then multiplied by 100 to obtain the accuracy percentage [25].

$$\text{Classification Accuracy} = (\text{correct predictions}) / (\text{total predictions}) \times 100\% \tag{1}$$

Table 1. Comparison The proposed system performance results

CNN type	NO. of layers	NO. of Epochs	Training time/min	Testing time/sec	Acc. %
AlexNet CNN	25	20	127	11.44	98.2
GoogleNet CNN	144	10	354	32.99	97.4
SqueezeNet CNN	68	5	369	11.2	99.8

6. RESULTS DISCUSSION

isplays the performance of three CNN types on different datasets. The outcomes indicate that SqueezeNet is the optimal choice due to its requirement of fewer epochs and the limited number of layers that compresses the input, leading to reduced storage space demand.

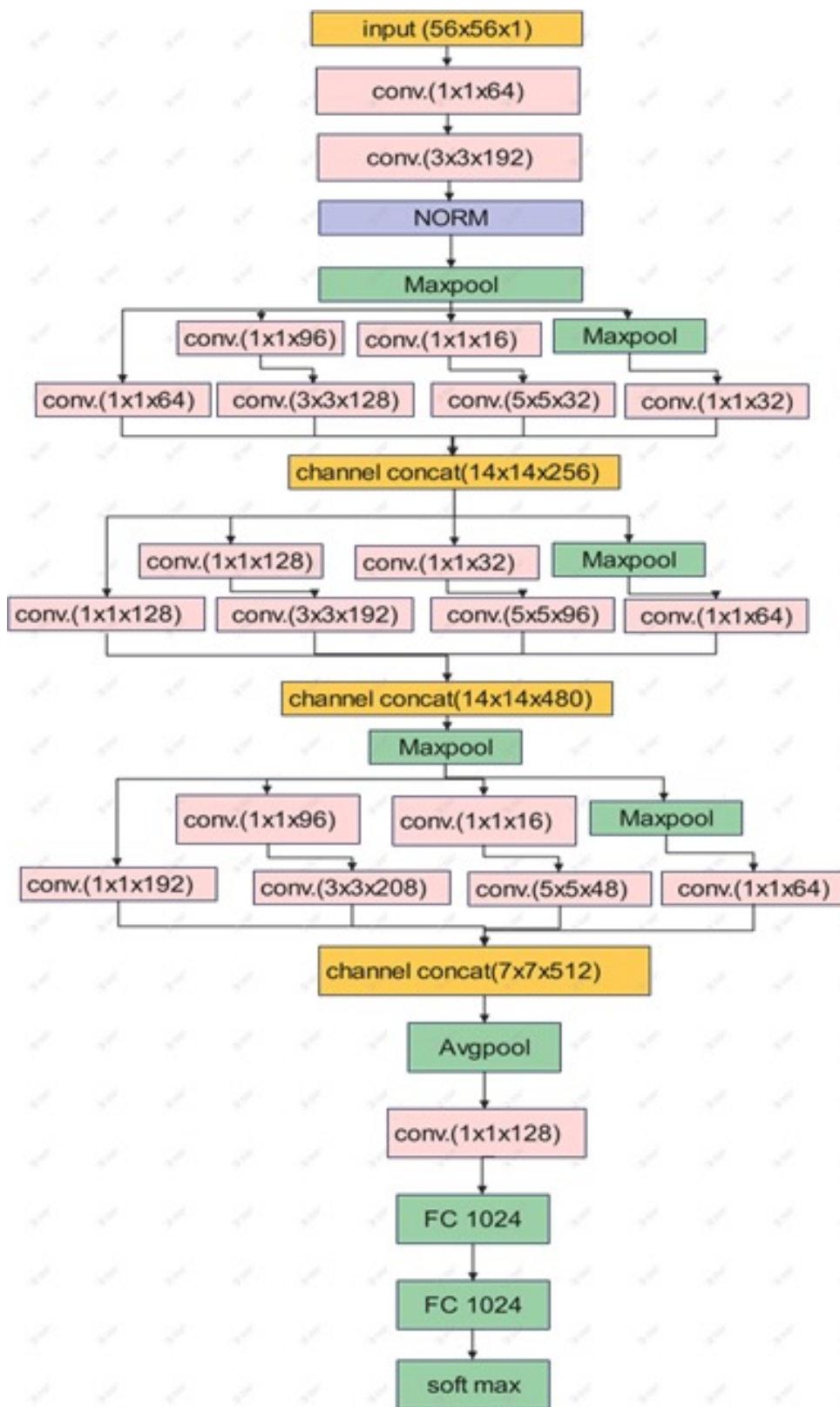


FIGURE 4. GoogleNet configuration [24]

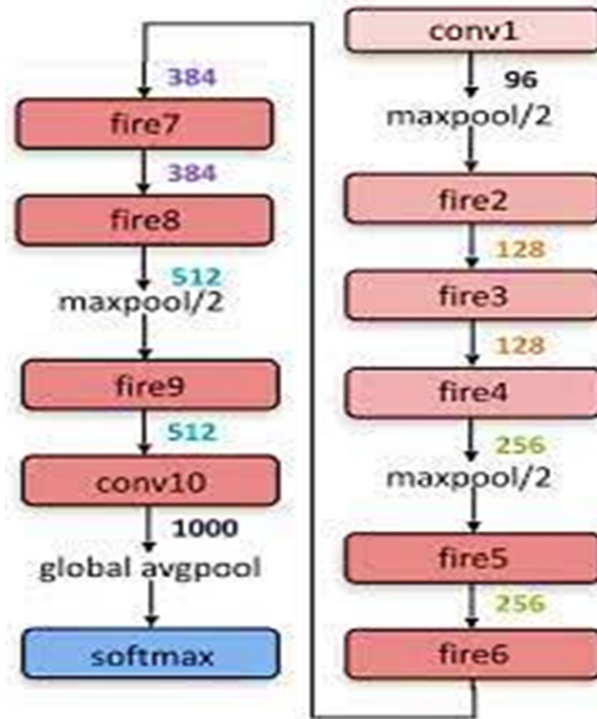


FIGURE 5. SqueezeNet configuration [25]

Illustrates a comparison among the three types of CNNs based on the year of invention, the number of parameters essential for the final classification decision, the ratio of parameter reduction concerning AlexNet as a reference (Ref.), the specific module utilized in each type, and the memory size required for each model.

Table 2. Comparison between three CNN models

CNN	Year of invention	No. of parameters in millions	Reduction ratio	Module implementation	Memory size/MB
AlexNet	2012	60	Ref.	Max pooling	202
GoogleNet	2015	15	14 time smaller	Inception module	20
SqueezeNet	2016	2	40 time smaller	Fire module	5

The acquired results in this research outperformed many recent research such as [26] The authors proposed an intelligent method for recognizing Telugu characters using a multi-objective mayfly optimization with deep learning (MOMFO-DL) model. To extract useful feature vectors, the DenseNet-169 model is used as a feature extractor. The classification of printer characters is carried out using the functional link neural network (FLNN). The recognition performance was of 98.74.

Also, the authors in [27] a method for generating diverse handwritten word images using handwritten characters is proposed. The approach involves training a BiLSTM-CTC architecture with synthetic handwritten words generated in two ways: overlapped and non-overlapped. The experimental setup includes recognition of handwritten documents using deep learning models, with the focus on the Bangla language, which lacks handwritten word datasets. The performance accuracy of the proposed model was 83%.

Moreover, the authors [28] proposed a pipeline system consists of 18 layers, including four layers each for convolution, pooling, batch normalization, dropout, and one Global average pooling and Dense layer. Hyperparameters such as optimizer, kernel initializer, and activation function were carefully examined. The proposed architecture was evaluated on two publicly available datasets, the 'Arabic Handwritten Character Dataset (AHCD)' and 'Modified Arabic handwritten digits Database (MadBase).' The model achieved accuracies of 96.93% and 99.35% respectively for these datasets.

The proposed system in this paper is outperformed the accuracy of many states of art research and the vital the contribution of this paper is specifying the testing time and required memory storage which are essential factors in real time systems applications.

7. CONCLUSION

This study aimed to create a system that can accurately recognize and distinguish handwritten characters and digits. This is crucial in today's digital world, especially for organizations dealing with handwritten documents that need computer analysis. By using advanced handwriting classification and recognition systems, both individuals and organizations can handle complex tasks effectively. We developed a CNN deep learning-based system for handwriting image recognition. Our system tackles key challenges in real-time applications: testing speed, memory usage, and performance. We tested our system on a large dataset and found it to perform comparably to recent state-of-the-art research. For future work, we suggest exploring a new handwriting recognition system that combines natural language processing (NLP) techniques for feature extraction and handwriting image features. This could open up exciting possibilities for further advancements.

FUNDING

None

ACKNOWLEDGEMENT

The authors would like to thank the editors and the anonymous reviewers for their insightful comments and suggestions, which were very helpful in raising the quality of this paper.

CONFLICTS OF INTEREST

The author declares no conflict of interest.

REFERENCES

- [1] L. Chen, S. Wang, W. Fan, J. Sun, and S. Naoi, "Beyond human recognition: A CNN-based framework for handwritten character recognition," *3rd IAPR Asian Conference on Pattern Recognition (ACPR)*, pp. 695–699, 2015.
- [2] A. K. A. Hassan, B. S. Mahdi, and A. A. Mohammed, "Arabic handwriting word recognition based on scale invariant feature transform and support vector machine," *Iraqi Journal of Science*, vol. 60, no. 2, pp. 381–387, 2019.
- [3] W. N. Hussein and H. N. Hussain, "A design of a hybrid algorithm for optical character recognition of online hand-written Arabic alphabets," *Iraqi Journal of Science*, vol. 60, no. 9, pp. 2067–2079, 2019.
- [4] X. Wang, J. Gao, and Y. Wang, "A comprehensive review of the extreme learning machine algorithm," *Neurocomputing*, vol. 338, pp. 48–58, 2019.
- [5] T. S. Gunawan, A. F. R. M. Noor, and M. Kartiwi, "Development of English handwritten recognition using deep neural network," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 10, no. 2, pp. 562–568, 2018.
- [6] A. A. Chandio, M. Leghari, D. Hakro, S. Awan, and A. H. Jalbani, "A Novel Approach for Online Sindhi Handwritten Word Recognition using Neural Network," *Sindh University Research Journal SURJ (Science Series)*, vol. 48, no. 1, pp. 213–216, 2016.
- [7] T. Wilkinson and A. Brun, "Semantic and verbatim word spotting using deep neural networks," *15th International Conference on Frontiers in Handwriting Recognition (ICFHR)*, pp. 307–312, 2016.
- [8] S. Ding, H. Zhao, Y. Zhang, X. Xu, and R. Nie, "Extreme learning machine: algorithm, theory and applications," *Artificial Intelligence Review*, vol. 44, no. 1, pp. 103–115, 2015.
- [9] N. A. Hamid and N. N. A. Sjarif, "Handwritten recognition using SVM, KNN and neural network," *Procedia Computer Science*, vol. 105, pp. 61–68, 2017.
- [10] A. A. A. Ali and M. Suresha, "A novel features and classifiers fusion technique for recognition of Arabic handwritten character script," *SN Applied Sciences*, vol. 1, no. 10, pp. 1286–1286, 2019.
- [11] M. Jaderberg, K. Simonyan, A. Vedaldi, and A. Zisserman, "Synthetic data and artificial neural networks for natural scene text recognition," *Computer Vision and Pattern Recognition*, 2014.
- [12] A. Poznanski and L. Wolf, "CNN-n-gram for handwriting word recognition," *CVPR*, pp. 2305–2314, 2016.
- [13] T. Wilkinson, J. Lindstrom, and A. Brun, *Neural ctrl-f: segmentation-free query-by-string word spotting in handwritten manuscript collections*, pp. 4443–4452, 2017.
- [14] X. Hou, J. Zhang, Q. Gao, and Y. Guo, "Handwriting recognition using convolutional recurrent neural networks and connectionist temporal classification," *Journal of Intelligent & Fuzzy Systems*, vol. 39, no. 4, pp. 5681–5691, 2020.
- [15] B. Shi, X. Bai, and C. Yao, "An end-to-end trainable neural network for image-based sequence recognition and its application to scene text recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 39, no. 11, pp. 2298–2304, 2017.
- [16] P. Voigtlaender, P. Doetsch, and H. Ney, "Handwriting recognition with large multidimensional long short-term memory recurrent neural networks," *ICFHR*, pp. 228–233, 2016.
- [17] P. Krishnan, K. Dutta, and C. V. Jawahar, "Word spotting and recognition using deep embedding," *2018 13th IAPR International Workshop on Document Analysis Systems (DAS)*, pp. 1–6, 2018.
- [18] P. Krishnan, K. Dutta, and C. V. Jawahar, "Deep feature embedding for accurate recognition and retrieval of handwritten text," *ICFHR*, pp. 289–294, 2016.
- [19] X. Wang, A. Shrivastava, and A. Gupta, *A-fast-rcnn: Hard positive generation via adversary for object detection*, pp. 1325–1334, 2017.
- [20] . Dataset and Landlord, *Handwriting Recognition Dataset*.
- [21] M. A. Al-Omari, "Handwritten Arabic character recognition using deep learning techniques without feature extraction," *Journal of Intelligent & Fuzzy Systems*, vol. 37, no. 3, pp. 4243–4253, 2019.

- [22] D. H. Kadhim and N. H. Salman, "A Review on Face Detection Based on Convolution Neural Network Techniques," *Iraqi Journal of Science*, vol. 63, no. 4, pp. 1823–1835, 2022.
- [23] S. Samir, E. Emary, K. El-Sayed, and H. Onsi *Optimization of a Pre-Trained AlexNet Model for Detecting and Localizing Image Forgeries*, vol. 11, pp. 275–275, 2020.
- [24] S. G. Lee, Y. Sung, Y. G. Kim, and E. Y. Cha, "Variations of AlexNet and GoogLeNet to Improve Korean Character Recognition Performance," *Journal of Information Processing Systems*, vol. 14, no. 1, pp. 205–217, 2018.
- [25] A. Wang, M. Wang, K. Jiang, M. Cao, and Y. Iwahori, "A Dual Neural Architecture Combined SqueezeNet with OctConv for LiDAR Data Classification," *Sensors*, vol. 19, no. 22, pp. 4927–4927, 2019.
- [26] V. K. Sonthi, "An Intelligent Telugu Handwritten Character Recognition Using Multi-Objective Mayfly Optimization with Deep Learning-Based DenseNet Model," *ACM Trans. Asian Low-Resour. Lang. Inf. Process.*, vol. 22, no. 3, pp. 1–16, 2023.
- [27] M. S. Akter, H. Shahriar, A. Cuzzocrea, N. Ahmed, and C. Leung, "Handwritten word recognition using deep learning approach: A novel way of generating handwritten words," in *Proc. IEEE Int. Conf. Big Data (Big Data)*, pp. 1–6, 2022.
- [28] M. Kamal, F. Shaiara, C. M. Abdullah, S. Ahmed, T. Ahmed, and M. H. Kabir, "Huruf: An Application for Arabic Handwritten Character Recognition Using Deep Learning," *International Conference on Computer and Information Technology (ICCIT)*, 2022.