



HANDWRITTEN CHARACTER RECOGNITION: A SOFT COMPUTING APPROACH

5881

¹Dr. Sanjiv Kumar Shukla, ²Dr. Balram Timande, ³Dr. Laxman L. Kumarwad, ⁴Dr. Avinash Dhole, ⁵*Sandeep Bhad

^{1,5}Rungta College of Engineering and Technology, Bhilai, (C.G), India

²Tulsiramji Gaikawad Patil College of Engineering and Technology, Mohgaon, Nagpur, (M.S.), India

³Government College of Engineering, Karad, (M.S), India

⁴Shri Shankaracharya Technical Campus, Bhilai, (C.G), India

*Corresponding Author- sandeepbhad@gmail.com

Abstract:

Handwritten characters are extremely challenging to comprehend because people's writing styles differ greatly from each other. Detecting and deciphering legible handwriting input from sources like touch screens, images, paper documents, and other sources is referred to as handwriting character recognition. Handwritten word recognition involves a variety of Soft Computing Techniques. The provision of cost-effective, high-performance, and dependable responses is one of the many advantages that come with employing a Soft Computing strategy. Handwriting character recognition uses various techniques and methods. Few use neural networks. Neural networks are more efficient and robust than other methods for recognising handwriting. This paper intends to report the development of a system for reading digital copies of handwritten text.

Keywords: - Soft Computing, ANN, Recognition, Reliable, Image Processing

DOI Number: 10.14704/nq.2022.20.9.NQ44685

Neuro Quantology 2022; 20(9):5881-5885

I. Introduction

Humans can solve and recognise any problem, but machines can't. It is necessary to employ a large number of strategies or procedures in order to function as a human. Handwriting numbers and recognising characters are becoming more and more important in today's digital world because they are used in so many everyday tasks. Complex problems can be solved by humans with the aid of systems that can read handwritten letters, characters, and numbers, which would otherwise be tedious and expensive to implement interactively. Inferential OCR can be done in two different ways, depending on the text format and document collection process. The optical character recognition (OCR) is classified into "Handwritten Character Recognition" (HCR) and intelligent recognition called Printed Character Recognition.

Variable human handwriting requires high recognition ability. There are many instances in

which the writing style of the same individual can shift dramatically from one occasion to the next. In addition, based on how the documents are acquired, OCR can be divided into two distinct types: offline and online recognition systems as presented in figure 1. Recognizing the pre-written document that was obtained through the many different input methods would be the focus of the offline system. On the other hand, the writing is recognised as soon as it is written by the system that recognises it online. The use of the electric pen can be found online. When using this pen to write on a digitizer, the movement of the pen itself records the input. In character recognition, there was a lot of work, but few papers were reviewed without optimal precision. We taught our computer Neural computation to give more accurate tests. This work uses English handwritten data as training input to recognise characters. This work is confined to identifying English characters and numerals only. Neural networks identify characters easily, which is



difficult for humans and normal systems. A qualified system would be an extraordinary piece of evaluative information. In light of the newly relevant conditions, this expert can then be used to incorporate predictions and respond to ambiguous queries. One of ANN's additional benefits is its capacity to pick up new skills through prior experience. ANN can develop self-organization using the knowledge it acquires during the learning process. The goal is to show how well neural networks work for recognising characters written by hand.

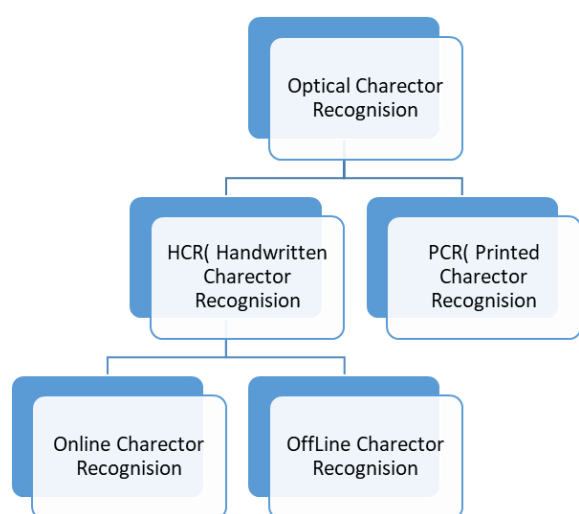


Figure 1 Optical Character Recognition flow

II. Literature review

The earliest attempt to recognize the handwritten character was made in 1959, when word recognition research was being conducted in Grimsdale. This study from the 1960s demonstrated how to use the "examination by combination" method that was put forth by Eden in 1968. He provided evidence that each handwritten character can only include a certain amount of schematic highlights. Later, practically all methodologies for auxiliary techniques in the field of character recognition made use of this hypothesis.

In this work, authors show how to train an artificial neural network to decode 0-9 digits and mathematical operators. The network has been trained to recognize the written digits 0 through 9. ANN proposed for recognition by the authors of a study [1]. In the suggested method, the first step is pattern recognition, and the second step is feature extraction. At first, the fuzzy pattern is

fetches using the fuzzy approach for pattern recognition. Problems arise when attempting to employ this strategy with handwritten characters due to the fact that different writers like to give each character its own unique shape, size, and placement [2].

Using multi-class GANs training techniques and sigmoid cross-entropy, this research recognizes Arabic handwritten letters. The proposed method is tested on 16800 handwritten Arabic characters. Compared to previous techniques, multi-class GANs is 99.7% accurate at recognizing Arabic handwritten characters [3]. Several Neural Network models useful for Optical Character Recognition are discussed, along with their respective behaviours. In their suggested work, they take into account a wide range of variables, including the number of hidden layers, the number of neurons employed in each layer, the number of epochs, and so on. In order to recognize characters, they utilize a network that uses both forward and backward propagation of information. Stages such as pre-processing, character segmentation, normalization, and de-skewing are included in their suggested work [4, 7]. In [5], the authors propose 'diagonal-based feature extraction' to recognize handwritten alphabets. Multilayer feed-forward neural network. Comparing the system's accuracy to other traditional horizontal or vertical based feature extraction methods, it performs at a high level. The authors of developed a system with the help of MATLAB in [6]. It converts the image to grayscale. Preprocessed image used for recognition. Character recognition uses MLP neural network [15].

The similarity of the characters makes it harder because the writing styles are different [8]. Using traditional methods, this research has improved its ability to recognize objects. Even though the number of studies has grown, even with well-designed features and strong classifiers, computers are still not as good at recognizing things as humans are. The problem is that there aren't enough benchmark datasets [9]. A great number of researchers have constructed their own datasets and classifiers in their own laboratories; however, the accuracies of each dataset are unique to that dataset, and the majority of the datasets are not made available to the public [10]. It is significantly more challenging to implement HDCR systems than it is

5882

to print characters, particularly for the various Indian scripts [11].

In this article, a new dataset containing the numerals and vowels written in Devanagari is presented. A total of 38,750 photographs are obtained from 2400 participants spanning a variety of age ranges. This dataset can be obtained through public access [12, 13]. Authors propose a segmentation algorithm for the dataset that was collected, and we also propose HDCR architectures for convolutional neural networks (CNNs) [14].

System is trained by associations [16]. The computational nodes that make up an artificial neural network are fairly simple, and they only have a limited amount of space for storing information [17]. Data and input are processed locally and connected via one-way channels. The Hidden Markov Model, also known as HMM, would be a primary modeling component that has been utilized in a variety of handwriting recognition systems. It is necessary to investigate the theoretical foundations of this model in order to acquire an accurate comprehension of the operation of handwriting recognition systems [18]. It is expected that HMM will assume the Markov process [20]. It is possible to regard it as the dynamic Bayesian network with the least amount of complexity. Hidden Markov Models are a class of probabilistic graphical models that are used for making predictions about a sequence of hidden variables based on a set of observed variables [19]. In [21], the authors suggest using a Hybrid CNN-SVM Classifier to recognize handwritten numbers. The SVM classifier isn't good at recognizing noisy data and isn't recommended for error generalization. Convolutional neural networks are used to generate features and to build the neural network in this method. Similar accuracy to that of proposed digit recognition architecture using Neural Networks for recognizing handwritten characters without feature extraction achieved in this work [22].

III. Methodology

Five distinct stages make up the current OCR system. To decipher handwriting, one must go through the following steps: capturing images, digitizing them, pre-processing them,

segmenting them, extracting features from them, and finally recognizing them.

Handwriting is used as an input method for image acquisition. The image must have the JPEG or PNG file extension. The image was captured using a digital camera, scanner, or one of several other types of equipment. A digital copy is produced when something is scanned. A digital image is required for the pre-processing step.

Pre-processing occurs after the digital image is available. Digital images are pre-processed to remove noise and checked for skew. Pre-processing data is essential for OCR. Pre-processing removes background noise, enhances the image's region of interest, and makes foreground and background clear.

Character recognition requires segmentation. During testing, image segmentation is carried out. By comparing all points to the typical distance between segmentation points, it checks for the inclusion of error points. Characters are separated from an image by it. Super pixels are produced by the process. The goal of segmentation is to simplify image analysis. Improved script recognition results.

In the feature extraction phase, image features are extracted based on attributes such as character height, number of horizontal lines, character width, and number of circles, pixels, feature position, and vertical arc number. In recognition phase, neural networks classify and recognize image characters. OCR systems use MLP and Korhonen's Self Organizing Map neural networks.

IV. Results

In character recognition techniques, segmentation can be argued to be the most important step. In testing, a representational example of the segmentation that can be found in figure 2. Handwritten characters and numeric values were read by an optical character recognition system. An earlier explanation of the implementation details revealed that a neural network was used. The results of the system were tested on the numbers which are already provided to the system for training range from 0-9 as well as alphabets from a-z. Figure 3 presents the output of tested 6 numbers and found accurate predictions for the same. Figure 4

5883



depicts the testing of random alphabets. Finally we have given the hand-written input image to the system and tested it for the correctness, and found satisfied except the letter r written in the text.

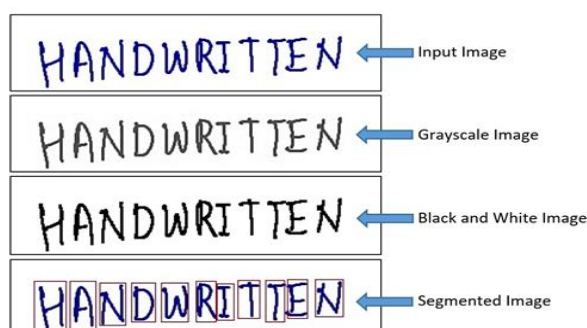


Figure 2 Segmented Image



Figure 3 Digit Prediction

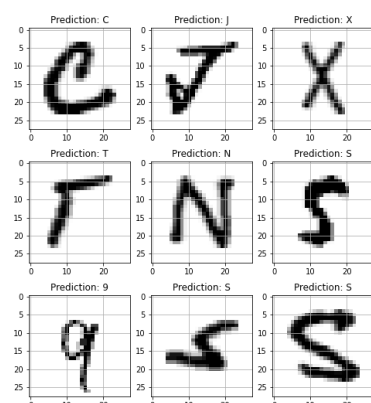


Figure 4 Alphabet Prediction

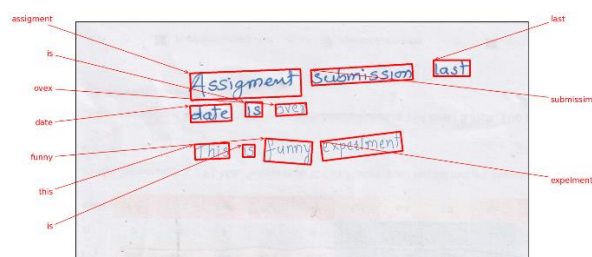


Figure 5 Prediction on new text lines

V. Conclusion

In today's digital world, it is crucial to be able to recognize characters and numbers, particularly in organizations that work with handwritten documents that require computer systems to analyse. Image segmentations still pose challenges because it can be challenging to distinguish between individual letters or numbers when several letters or numbers are overlapping one another. There is evidence of multi-word recognition from the system (sentence recognition). It will also serve as evidence of how well the system was trained. Accuracy and recognition criteria were met by the final system. The work being done by the current research can be expanded to include a great deal more purposes.

References

- [1] Malhotra, A., Gupta, M., Yadav, R., Tokas, P., & Sharma, A. (2022). ANN-Based Handwritten Digit Recognition and Equation Solver. In Lecture Notes on Data Engineering and Communications Technologies (Vol. 90, pp. 77-86). Springer Science and Business Media Deutschland GmbH. https://doi.org/10.1007/978-981-16-6289-8_8
- [2] H. Akouaydi, S. Njah, W. Ouarda, A. Samet and T. Dhieb, "Neural architecture based on fuzzy perceptual representation for online multilingual handwriting recognition," *Journal of Latex Class Files*, vol. 5, no. 3, pp. 1-12, 2019
- [3] Using multi-class GANs training techniques and sigmoid cross-entropy, this research recognizes Arabic handwritten letters. The proposed method is tested on 16800 handwritten Arabic characters. Compared to previous techniques, multi-class GANs is 99.7% accurate at recognizing Arabic handwritten characters.
- [4] L. B. Mahanta and A. Deka, "Skew and slant angles of a handwritten signature," *International Journal of Innovative Research in Computer and Communication Engineering*, vol. 1, no. 9, pp. 2030-2034, 2013.
- [5] J. Pradeep, E. Srinivasan and S. Himavathi, "Diagonal based feature extraction for handwritten character recognition system using neural network," in *Int. Conf. on*

- Electronics Computer Technology*, Kanyakumari, India, pp. 364–368, 2011.
- [6] C. I. Patel, R. Patel and P. Patel, "Handwritten character recognition using neural network," *International Journal of Scientific & Engineering Research*, vol. 2, no. 5, pp. 123–132, 2015.
- [7] Alwaqfi, Y. M., Mohamad, M., & Al-Taani, A. T. (2022). Generative Adversarial Network for an Improved Arabic Handwritten Characters Recognition. *International Journal of Advances in Soft Computing and Its Applications*, 14(1), 176–195. <https://doi.org/10.15849/IJASCA.220328.12>
- [8] Li, Z., Teng, N., Jin, M., & Lu, H. (2018). Building efficient CNN architecture for offline handwritten Chinese character recognition. *International Journal on Document Analysis and Recognition (IJ DAR)*, 21(4), 233–240.
- [9] Kaur, S., & Sagar, B. B. (2019). Brahmi character recognition based on SVM (support vector machine) classifier using image gradient features. *Journal of Discrete Mathematical Sciences and Cryptography*, 22(8), 1365–1381.
- [10] Setlur, S. (2009). Guide to OCR for indic scripts. In V. Govindaraju (Ed.). Springer.
- [11] Bathla, A. K., Gupta, S. K., & Jindal, M. K. (2016, March). Challenges in recognition of Devanagari Scripts due to segmentation of handwritten text. In 2016 3rd International Conference on Computing for Sustainable Global Development (INDIACom) (pp. 2711–2715). IEEE.
- [12] [https:// data. mende ley. com/ datas ets/ pxrnvp4yy8/1](https://data.mendeley.com/datasets/pxrnvp4yy8/1)
- [13] Bharath, A., & Madhvanath, S. (2011). HMM-based lexicon-driven and lexicon-free word recognition for online handwritten Indic scripts. *IEEE transactions on pattern analysis and machine intelligence*, 34(4), 670–682.
- [14] Rajput, S. S., & Choi, Y. (2022). Handwritten Digit Recognition using Convolution Neural Networks. In 2022 IEEE 12th Annual Computing and Communication Workshop and Conference, CCWC 2022 (pp. 163–168). Institute of Electrical and Electronics Engineers Inc. <https://doi.org/10.1109/CCWC54503.2022.9720854>
- [15] Datsi, T., Aznag, K., & El Oirrak, A. (2022). Digit recognition using decimal coding and artificial neural network. *Kuwait Journal of Science*, 49(1). <https://doi.org/10.48129/KJS.V49I1.9556>
- [16] Nair, R. R., Sankaran, N., Kota, B. U., Tulyakov, S., Setlur, S., & Govindaraju, V. (2018, April). Knowledge transfer using Neural network based approach for handwritten text recognition. In 2018 13th IAPR International Workshop on Document Analysis Systems (DAS) (pp. 441–446). IEEE.
- [17] Xu, Y. (2010). Examining the Effects of Digital Feedback on Student Engagement and Achievement. *Journal of Educational Computing Research*, 43(3), 275–291. doi:10.2190/ec.43.3.a
- [18] Shi, W., Caballero, J., Huszar, F., Totz, J., Aitken, A. P., Bishop, R., ... Wang, Z. (2016). Real-Time Single Image and Video Super-Resolution Using an Efficient Sub-Pixel Convolutional Neural Network. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 1874–1883. doi:10.1109/cvpr.2016.207
- [19] Gunawan, T. S., Noor, A. F. R. M., & Kartiwi, M. (2018). Development of english handwritten recognition using deep neural network. *Indonesian Journal of Electrical Engineering and Computer Science*, 10(2), 562–568.
- [20] Zhou, H., Chen, J., Dong, G., & Wang, R. (2016). Detection and diagnosis of bearing faults using shift-invariant dictionary learning and hidden Markov model. *Mechanical systems and signal processing*, 72, 65–79.
- [21] S. Ahlawata and A. Choudhary, "Hybrid CNN-SVM classifier for handwritten digit recognition" *Procedia Computer Science*, Volume 167, 2020, Pages 2554–2560.
- [22] Pradeep J, S. E, Himavathi S. Neural network based handwritten character recognition system without feature extraction. In 2011 international conference on computer, communication and electrical technology (ICCCET) 2011 Mar 18 (pp. 40–44). IEEE.

