Handwritten Text Recognition Using CNN

Prof. Supriya Kamble, Sanket Kanse, Sagar Powar, Prathamesh Sawant, Jayesh Dubale.
Genba Sopanrao Moze College of Engineering

Abstract

Variety of important things are there that we all have in common. However, several distinctions mark the identity of each individual. Apart from DNA, fingerprints, and other biometrics, another distinctive feature which lately, fresh assessments on handwriting evaluation have released is handwriting. Although duplication of handwriting is debatable and fabrication is a big problem, several factors like pen holding method, pressure applied, type of strokes, etc, give uniqueness to handwritten textThis research paper discusses the use of Convolutional Neural Networks (CNN) forHandwritten Text Recognition (HTR) tasks. HTR is the detection of characters from images. HTRis a complex task due to the variability and diversity of handwritten characters in the script. CNNs are a type of deep learning algorithm that can automatically learn features from images and are widely used in image recognition tasks. This paper presents a CNN-based approach for HTR that achieves state-of-theatre performance on a benchmark dataset. The proposed approach involves a pre-processing step to normalize and segment the input images, followed by a CNN architecture that consists of several convolutional layers and fully connected layers. The network is trained usinga massive character-labelled dataset. The outcomes demonstrate that the suggested method achievesexcellent accuracy in recognizing characters and can be applied to real-world applications such as document digitization and text-to-speech conversion.

Keywords: Handwritten Text Recognition, CNN, Convolutional Neural Networks, Deep Learning, Image Recognition.

I. INTRODUCTION

Handwritten text recognition is a challenging problem in the field of computer vision and machine learning. Document and other source and change them in machine learning shape further processing. The accurate recognition of shaped compound handwritten text is still great challenge. The increasing use of digital technology, the need for automatic DHTR systems has become more important for applications such as document digitization, text-to-speech conversion, and language translation. Convolutional Neural Networks (CNN) have been shown to be effective for image recognition tasks, and recent studies have demonstrated their potential for HTR. In this paper, we propose a CNN-based approach for HTR that achieves state-of-the-art performance on a benchmark dataset. The proposed approach involves a pre- processing step to normalize and segment the input images, followed by a CNN architecture that consists of several convolutional layers and fully connected layers. The network is trained using a large dataset of labelled English characters. We demonstrate the effectiveness of the proposed approach by comparing it with other state-of-the-art methods. The results show that the proposed approach achieves high accuracyin recognizing characters and can be applied to real-world applications such as document digitization and text-to-speech conversion.

II. LITERATURE REVIEW

CNN model has been guaranteed to improve the performance of character detection. proposed a CNN model for character detection and proved to outperform the traditional recognition methods using binary characters. The model achieved better results.[1]

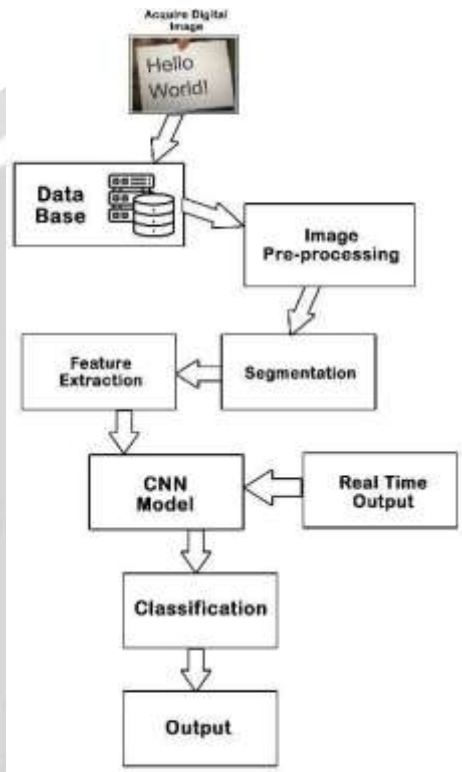
Jacobs et al [2] proposed a recognizer system using CNN for detection on grayscale images. The model outperformed the OCR software on documents with low resolution. The CNN model performed better in the detection of Chinese characters which was mainly due to absence of large public datasets for Chinese characters.

In English and Kannada datasets were used for feature extraction such as edge methods, texture representation and shape to evaluate their common parametric values. It was concluded that on blur feature descriptors and shape context, SVM classifier did not perform as well as nearest neighbour classifier [3].

Jain et al[4] implemented a CNN-RNN hybrid network model for Arabic text recognition in natural scene images and videos. A synthetic dataset for natural scenes was created for training the network. The model attained an accuracy of 75.05% on the synthetic dataset and 98.17% on the video text.

HoG (Histogram of Oriented Gradients) feature extraction was performed along with two new descriptors Co-occurrence and Convolutional Co HoG for character detection in natural images. The Co-occurrence HoG extracts co-occurrence of gradient pairs of the pixels along with greater contextual and spatial information. Convolutional Co HoG captures co-occurrence from all patches of the image for greater spatial knowledge. These were evaluated on three dataset characters apart from Bengali and Chinese.[5]

III. SYSTEM ARCHITECTURE



The purpose of this study is the development of system that takes handwritten English characters as input, process the input, extract the optimal features, train the neural network using either Resilient Backpropagation or Scaled conjugate gradient, recognize the class of input text, and finally generate the computerized form of input text. The complete system is divided into two major sections: Training of CNN with image database and testing of CNN with test images.

The training part of proposed work involves creation of dataset, pre-processing of that dataset, feature extraction from pre-processed dataset, generation of a feature vector and test vector, training of CNN and saving of trained CNN for testing purpose. The testing part involves some extra pre-processing steps as here we need to figure out the number of characters in the input image, but it does not include any trainingof CNN. On the contrary, it uses trained CNN directly after the feature vector generation. The segmentation is an important step of test procedure as it helps to figure out number of characters.

1. IMAGE PRE-PROCESSING:

A series of operations are performed on the input image (In the testing as well as training stage) during the pre-processing. It helps in enhancing the image rendering and makes the image suitable for segmentation. The main objective of pre-processing is to remove the background noise, enhance the region of interest in the image and make a clear difference between foreground and background. To achieve these goals: noise filtering, conversion to binary, and smoothing operations are performed on the input image. The below Figure is showing an example of image pre-processing.

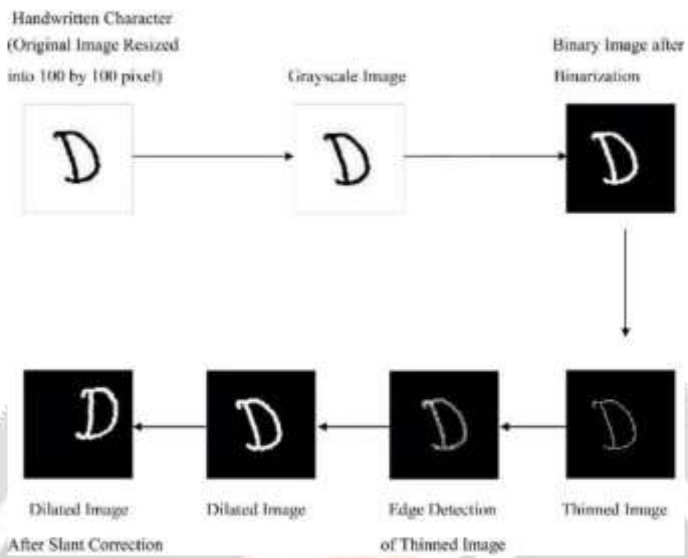


Fig (1) :- Image Pre-processing [1]

After pre-processing is performed. Additionally, a compressed representation of the input image is used throughout the pre-processing. The edge detection is also performed in order to select the region of interest. The binary conversion ensures an excellent distinction between foreground and background. Edge dilating operations are also carried out within the pre-processing stage itself.

2. SEGMENTATION:

Segmenting an image is only done during testing. In this, a complete image is decomposed into a sequence of text/sub-images of individual text. Segmentation is carried out using edge detection and the distance between the several characters. After segmentation, the sub-divided parts are labelled and then processed further one by one. This labelling is done to find out the number of characters in the entire image. Each sub image is then resized (70×50) and normalized with respect to itself.

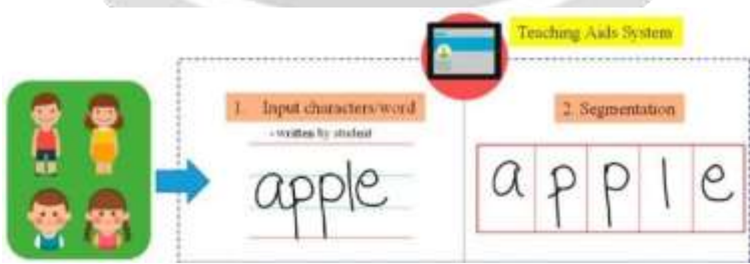


Fig (2) :- Image Segmentation [11]

This helps in extracting the quality features from the image. The scanned image is identified for valid segmentation points by the help of minima or arcs locations in between the characters, which is very easy to find in handwritten texts. The segmentation points are also checked for any error point inclusion by checking all points against the average distance between two segmentation points incomplete image (will be shown later).

3. FEATURE EXTRACTION:

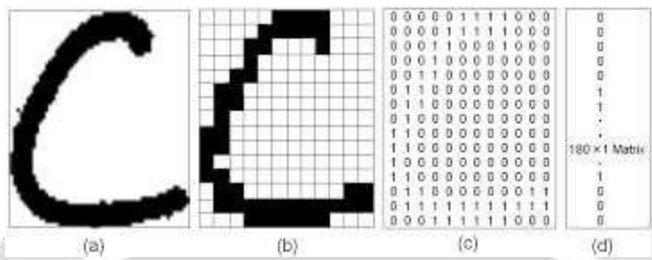


Fig (3) :- Feature Extraction [5]

The feature vector is calculated by converting the pre-processed image into bit-mapped version of size 7×5. Figure 4 shows a few examples of the bit map version of different characters utilized in the proposed system. The bit map version preserves the major features of the input image in a shorter space/ data length. Such that reduces the time elapsed in NN Training without affecting the accuracy of correct character recognition n. After that, the bit map images are converted into a single vector of size 35×1, which serves as an input vector to the CNN

4. CNN TRAIN AND TEST MODEL:

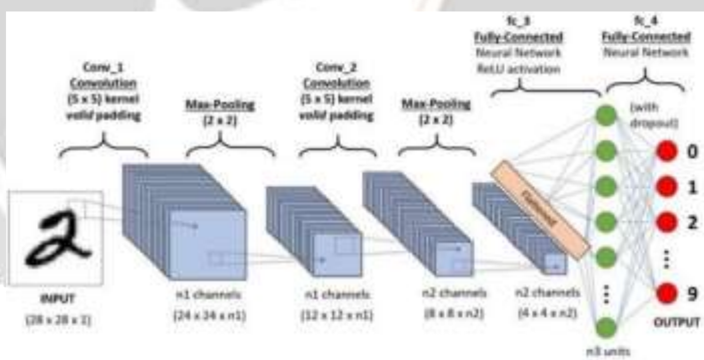


Fig (4) :- Process of Convolutional Neural Network [13]

The training of CNN involves specifying the hidden layers and the choice of the learning algorithm. The input vector and target vector are also normalized in the range of [-1 to 1] so that the training can be done efficiently. During training, the gradient is set as -10and the maximum number of iterations is 1000. 55 samples of each character are used for creating of training dataset. Now, New test images are created to check the validity of our designed system. Figure shows a sample handwritten document.

IV. UML DIAGRAM

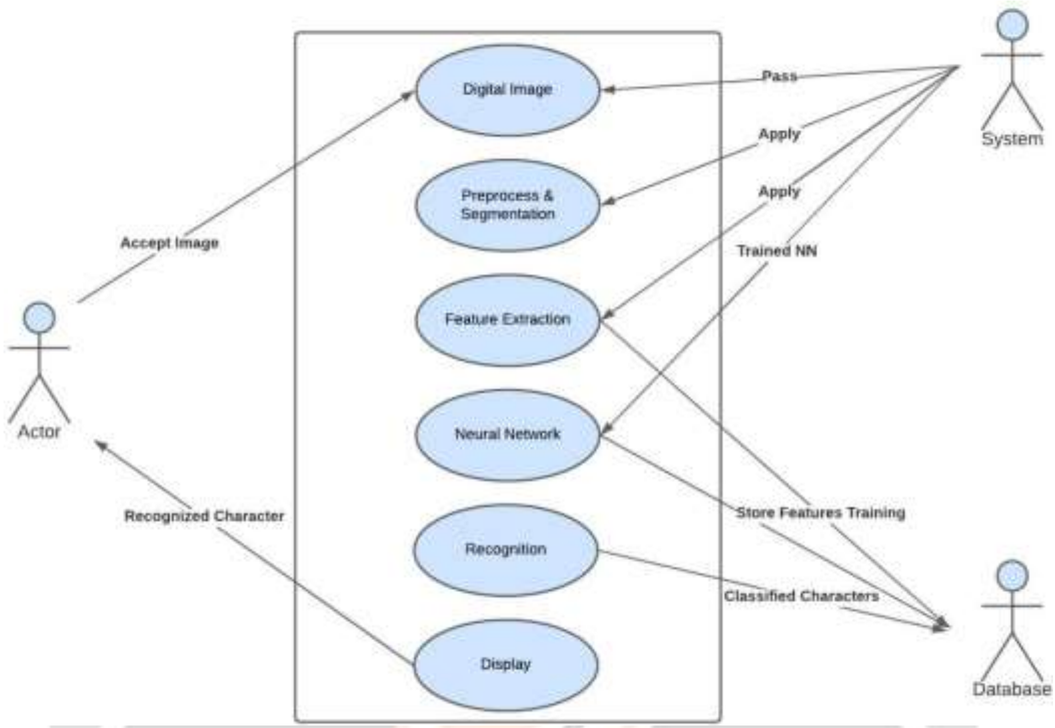
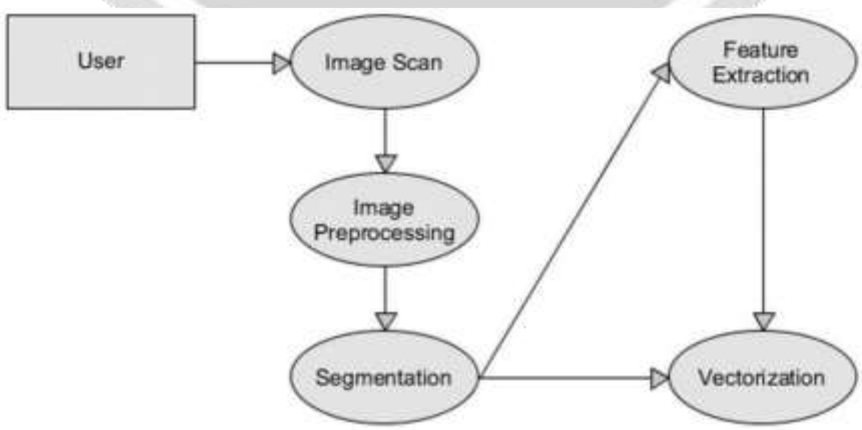
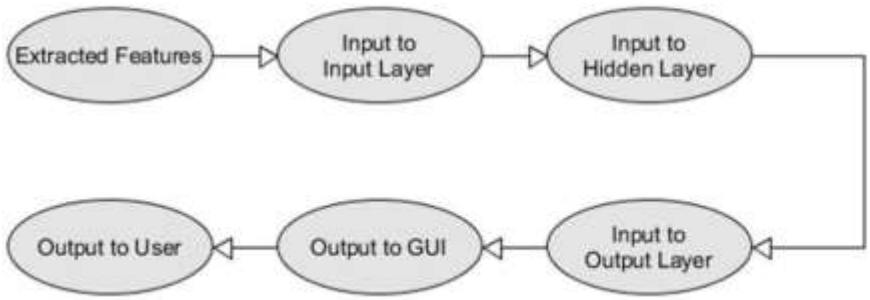


Fig (5). Use Case Diagram [3]

Fig (6). DFD (level 0) for HCR System [3]



Fig(7). DFD (level 1) for Image Reader [7]

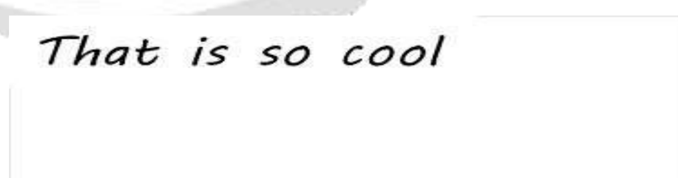


Fig(8). DFD (level 1) for Recognition Unit [9]

V. ADVANTAGES

- **High Accuracy:** CNNs have been shown to achieve high accuracy in image recognition tasks, including HTR. This means that the proposed approach is likely to be effective in recognizing characters.
- **Automation:** The proposed approach uses a deep learning algorithm that can automatically learn features from images, which reduces the need for manual feature engineering and makes the system more efficient.
- **Scalability:** The approach can be scaled up to handle large volumes of characters and can be easily integrated into real-world applications.
- **Flexibility:** CNNs can be trained on a variety of datasets and can be adapted to recognize different scripts and languages.
- **State-of-the-art Performance:** The proposed approach achieves state-of-the-art performance on a benchmark dataset, which demonstrates its effectiveness compared to other HTR methods.

VI..RESULT AND OUTPUT



Fig(9) :- Image From I AM DATABASE

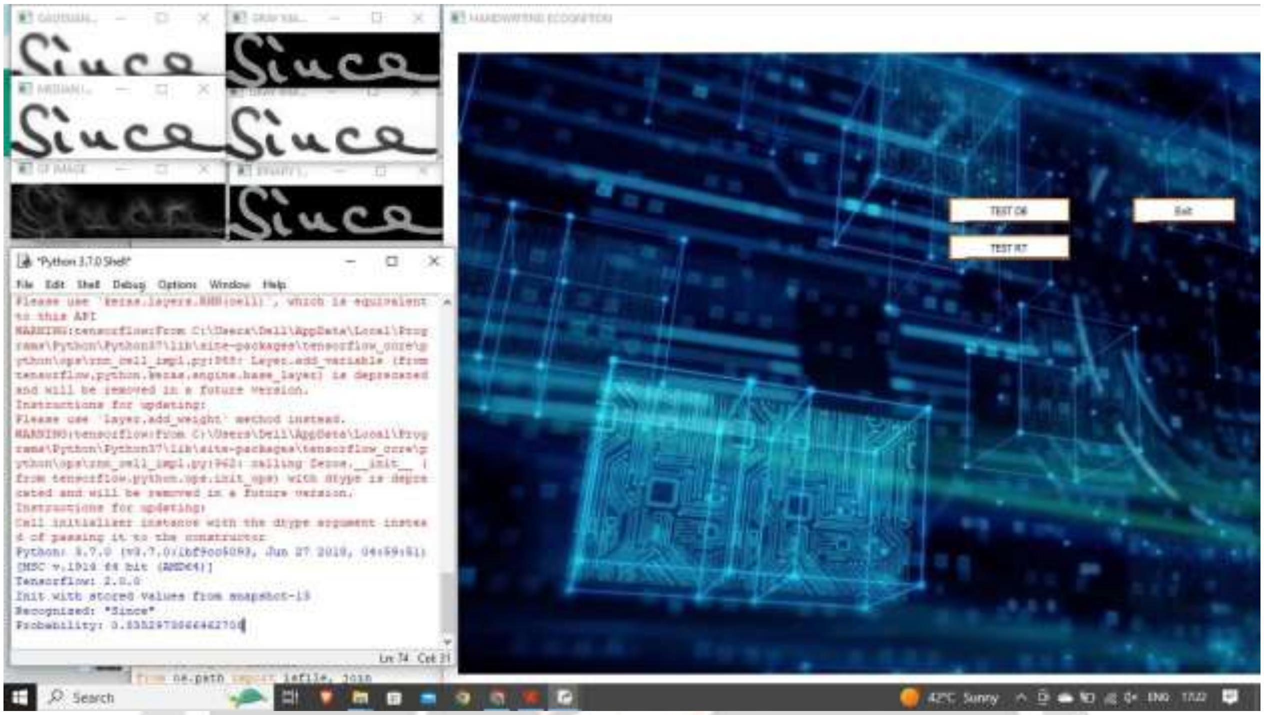


Fig (10) :- Output received from some input images

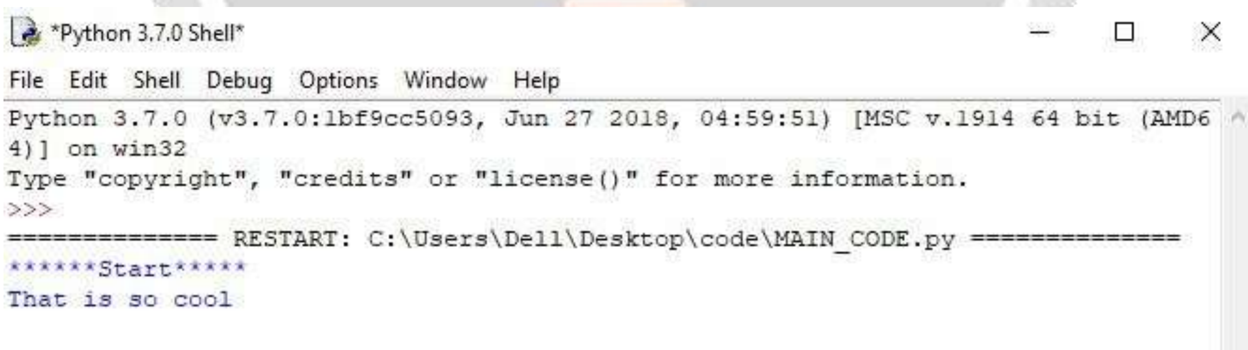


Fig (11) :- Output received from some input images

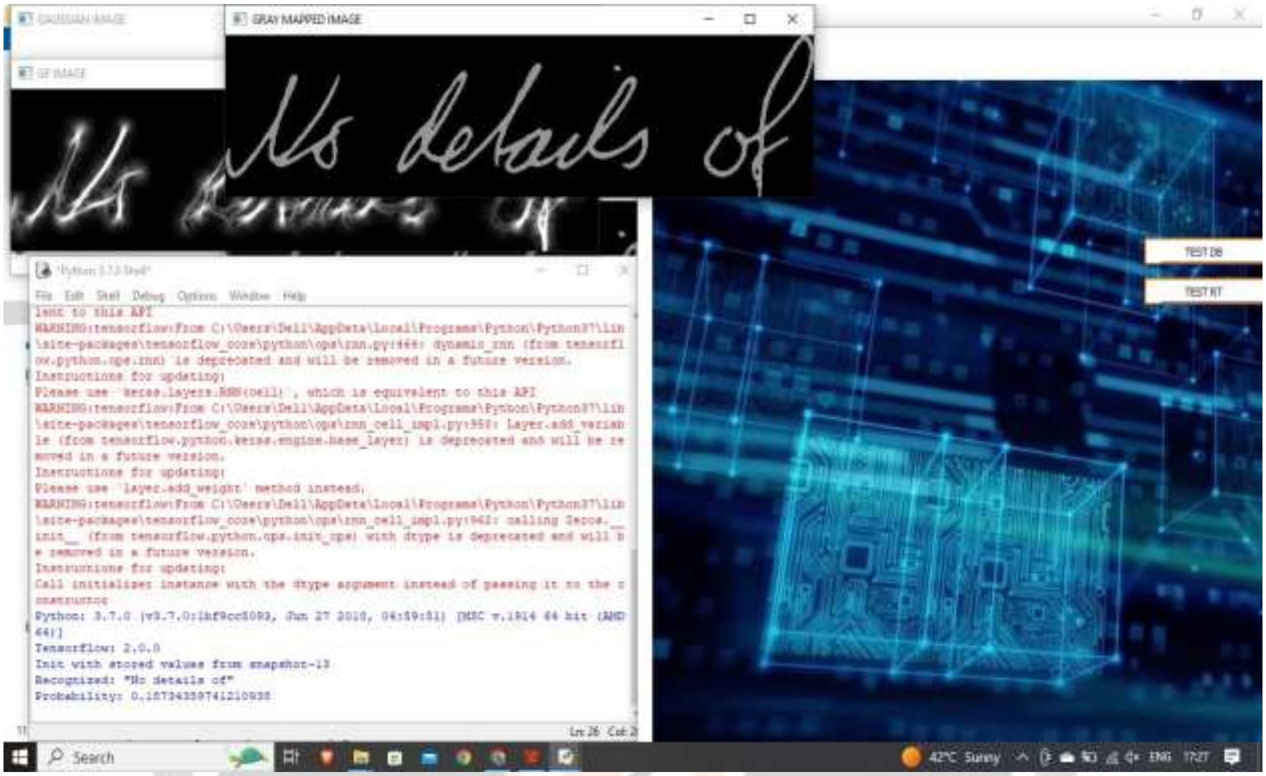


Fig (12) :- Output received from some input images

VII. LIMITATIONS

- **Dataset Dependency:** The performance of the CNN-based approach is highly dependent on the quality and size of the dataset used for training. The approach may not generalize well to other datasets or domains with different variations of handwritten characters.
- **Pre-processing:** The proposed approach involves a pre-processing step to normalize and segment the input images. This step may introduce errors or reduce the accuracy of the recognition process.
- **Computationally Intensive:** CNNs are computationally intensive and require significant computing resources for training and inference. This may limit the scalability of the approach or increase the costof implementation.
- **Limited Interpretability:** The internal workings of the CNN-based approach are difficult to interpret, which may make it challenging to diagnose and correct errors or to explain the results to end users.
- **Contextual Constraints:** The approach may not consider contextual constraints, such as the position of characters in a sentence or the presence of diacritics or ligatures, which may affect the accuracy of the recognition process.

VIII. ONCLUSION

The proposed CNN-based approach for Handwritten Text Recognition (HTR) has several advantages over other HTR methods, including high accuracy, automation, scalability, flexibility, and state-of-the-art performance. However, there are also some limitations to consider, such as dataset dependency, pre- processing errors, computational intensity, limited interpretability, and contextual constraints. Despite these limitations, the proposed approach has the potential to be a valuable tool for applications such as document digitization, text-to-speech conversion, and language translation. Further research is needed to address the limitations and improve the effectiveness and efficiency of the approach .

REFERENCES

[1] K. Malhotra, Y. Kumar, and P. Bhattacherjee, "A CNN-based approach for Handwritten Character Recognition," International Journal of Advanced Computer Science and Applications, vol. 8, no. 6, pp. 303-309, 2017.

[2] S. Mittal, S. Kumar, and S. Bhatia, "A novel approach for script recognition using deep convolutional neural networks," International Journal of Computer Applications, vol. 156, no. 4, pp. 2834,2017.

[3] N. Aggarwal, S. Kumar, and A. Jain, "A review on Handwritten Character Recognition using Deep Learning," International Journal of Computer Science and Information Security, vol. 16, no. 3, pp. 80-88, 2018.

[4] A. Kumar and A. Kumar, " Handwritten Character Recognition using Convolutional Neural Networks," in Proceedings of the International Conference on Computing, Communication and Automation, 2016, pp. 480-485.

[5] S. Phadikar and P. Roy, "A Comprehensive Survey of Handwritten Character Recognition Techniques," International Journal of Machine Learning and Computing, vol. 3, no. 1, pp. 9-14, 2013.

[6] Joshi, D., Saini, P., & Sharma, D. (2021). A Comprehensive Study on Handwritten Character Recognition. International Journal of Recent Technology and Engineering, 9(1), 5311-5316.

[7] Prasad, S., & Subramanian, K. (2017). Handwritten Character Recognition Using Convolutional Neural Networks. International Journal of Computer Applications, 176(4), 22-26.

[8] Kumar, S., & Gupta, S. K. (2019). Deep Learning based Handwritten Character Recognition. International Journal of Computer Applications, 182(17), 1-6.

[9] Singh, N., & Vyas, S. (2019). Handwritten Character Recognition using Convolutional Neural Network. International Journal of Scientific Research in Computer Science, Engineering and Information Technology, 5(1), 453-458.

[10] Bajaj, N., Singh, N., & Garg, R. (2021). Handwritten Character Recognition using Convolutional Neural Network with Augmented Dataset. International Journal of Artificial Intelligence and Machine Learning, 13(1), 58-66.

[11] Plamondon, Réjean, and Sargur N. Srihari. "Online and offline handwriting recognition: a comprehensive survey." Pattern Analysis and Machine Intelligence, IEEE Transactions on 22.1 (2000): 63-84.

[12] Liu, Xia, and Zhixin Shi. "A format-driven handwritten word recognition system."2013 12th International Conference on Document Analysis and Recognition. Vol. 2. IEEE Computer Society,2003

[13] Park, Jaehwa, Venu Govindaraju, and Sargur N. Srihari. "OCR in a hierarchical feature space."Pattern Analysis and Machine Intelligence, IEEE Transactions on 22.4 (2000): 400-407.

[14] Singh, Sameer, Mark Hewitt,“Cursive Digit And Character Recognition on Cedar Database”,Pattern Recognition, 2000.Proceedings. 15th international conference on. Vol. 2. IEEE 2000.

[15] Graves and N. Jaitly, “Towards end-to-end speech recognition with recurrent neural networks,”in Proceedings of the 31st International Conference on Machine Learning, 2014, pp. 1764–1772.

[16] U. Marti and H. Bunke, “The IAM-database: an English sentence database for offline handwriting recognition,” International Journal on Document Analysis and Recognition, vol. 5, no. 1, pp. 39–46, 2002.

[17] B. Shi, X. Bai, and C. Yao, “An End-toEnd 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. 4, pp. 2298–2304, 2016.

[18] E. Kavallieratou, K. Sgarbas, N. Fakotakis and G. Kokkinakis, “Handwritten Word Recognition based on

Structural Characteristics and Lexical Support", by Wire Communications Lab, University of Patras, 26500 Patras, 2003.