iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 10 **Issue:** III **Month of publication:** March 2022

DOI: https://doi.org/10.22214/ijraset.2022.40876

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 10 Issue III Mar 2022- Available at www.ijraset.com

Handwriting to Text Conversion for English Language Using Deep Learning

Ketaki G. Dhotre[1], Harshali K. Ghumate[2], Mayuri Mane[3], Prof. Savita Lade[4]
[1, 2, 3]*UG Student,* [4]*Assistant Professor, Department of Computer Engineering, Rajiv Gandhi Institute of Technology-University of Mumbai*

Abstract*: Because of the rising use of digital technology in all businesses and in all day-to-day activities to store and communicate information, recognition systems in writing have become a prominent study topic and development. Humans still require handwriting copies to be converted into digital copies that can be shared and preserved electronically. Handwriting recognition is one of the most active study areas, and deep neural networks are being used in it. Humans find it simple to recognise handwriting, but computers find it tough. Nowadays, technologies that detect handwriting letters, characters, and figures assist people in doing more sophisticated activities that would otherwise take a long time and be costly. The purpose of this project is to turn handwritten notes into typed documents. We aim to transform handwritten English characters into a computer-readable format using a paragraph as an input, process the paragraph with cursive writing and symbols support, and then train a neural network algorithm to recognize and display the text. CNN is the neural network model that we used. The image can be uploaded by the user. To eliminate background noise, the system pre-processes the input. The machine then looks for text sections in the image. The system then displays the text that is contained in the image to the user. To conduct horizontal-vertical segmentation, we used OpenCV.

Keywords: Handwriting, Bi-LSTM, Convolutional Neural Network, Text Conversion, Deep Learning, OpenCV

I. INTRODUCTION

Handwriting digits and character recognitions have become increasingly important in today's digitized world due to their practical applications in various day to day pursuits. [1]. The text recognition plays an important role in several areas. For recognition, we've to train the system to acknowledge the text. The character recognition involves several steps like acquisition, feature extraction, classification, and recognition. [3]. Handwriting recognition systems can be classified as Online and Offline. The information like the order in which the user has made the strokes is available in online recognition systems [2]. But in offline handwriting recognition system, the user provides the handwritten data in the form on an image. For a variety of reasons, handwriting recognition has proved problematic. Different writing styles are one of them. Another significant factor is the large number of characters, which include capital letters, tiny letters, digits, and special symbols. Thus, we need a large dataset to train a near-accurate neural network model. To develop a good system an accuracy of at least 99% is required. However even the most modern and commercially available systems have not been able to achieve such a good accuracy. The objective of our project is - The user will be able to upload the object image, the system must pre-process the input to eliminate the background, the system should detect text regions in the image, and the system should extract and display the text found in the image to the end-user. We have built a Neural Network (NN) which is trained on word-images from the IAM dataset. The IAM dataset includes 657 writers contributed samples of their writings, 1532 pages of scanned text, 5685 isolated and labelled sentences, and 115320 isolated and labelled words [5]. The main goal of this research is to create a version that will be used to analyse handwriting numbers, characters, and words from images utilizing the Convolution Neural Network concept. [1]

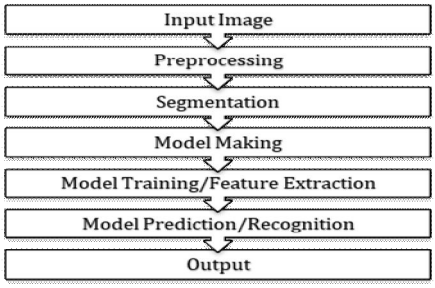


Fig 1: System Flow

©IJRASET: All Rights are Reserved

Section II briefly explains about the past work done in this particular discipline. Sections II and III delineate the datasets and models used for handwriting recognition, and the results obtained for each. Finally, Section V presents our conclusions and briefly illustrates the potential for further improvements in our methodology.

II. STATE OF ART

A lot of research has been done and is going on in the field of handwritten character recognition. Many people have developed systems for handwritten character recognition. A lot of research has been done and is going on in the field of handwritten character recognition. Many people have developed systems for handwritten character recognition. We have studied some of the systems and tried to implement their limitations in our project.

In 2016 Ahmed Mahdi Obaid et al. [4] published a paper under Handwriting to Text Conversion domain. In this paper, they have used 3-layer Artificial Neural Network (ANN) using supervised learning approach. 55 samples of each English alphabet are used as ANN training process in order to make sure the general applicability of system towards new inputs. Two different learning algorithms are being used in this paper. They have implemented additive image processing algorithms in order to deal with the multiple characters as input in a single image or rotated image. The trained system provides an average accuracy of more than 95 % with the unseen test image [4].

In 2018, Rohan Vaidya et al. [2] published a paper under the handwriting recognition domain. In this paper, the image is first segmented into individual lines. After that one single line is segmented into separate words. Lastly, individual word is segmented into individual letters. Then, it predicts using CNN model trained using NIST dataset for characters. As of now their system can't recognize cursive handwritten text and special symbols [2]. In 2020, the paper published by Sara Aqab et al. [1], converts handwritten text to single character and then predicts the image using CNN model which is trained using NIST dataset for characters and Kaggle dataset for digits. The machine learning model was trained with a dataset that contains 42,000 rows and 720 columns, which the result shows 83.4% accuracy [1]. In 2020, Sri. Yugandhar Manchala et al. [3] predicted the handwritten words using Neural Network consisting of (CNN) layers, recurrent neural network (RNN) layers, and a final Connectionist Temporal Classification (CTC) layer trained using IAM words dataset. This model supports the classification of characters. The project is trained and made using the conventional neural network. The accuracy they obtained in this model was above 90.3%.

Existing Systems doesn't support special characters and also doesn't support multiple languages. Cursive handwriting, which is a connected set of letters, has more issues with evaluation. Also, existing system has no implementation for cursive writing.

III.METHODOLOGY

This section is divided into various parts Our Project can be understood by the following steps-

1) Image Segmentation and Processing.
2) Multi-scale feature Extraction which consists of Convolutional Neural Network -7 Layers
3) Sequence Labelling (BLSTM-CTC) that includes Recurrent Neural Network (2 layers of LSTM) with CTC
4) Transcription-Decoding the output of the RNN (CTC decode).
5) Prediction of proper words based on Auto-Correct (Spell Checker)

A. Image Segmentation and Processing

Segmentation is the process of assigning names to pixels. Image segmentation can be used as a pre-processing step before applying a machine learning set of rules to reduce the amount of time it takes to process the image.

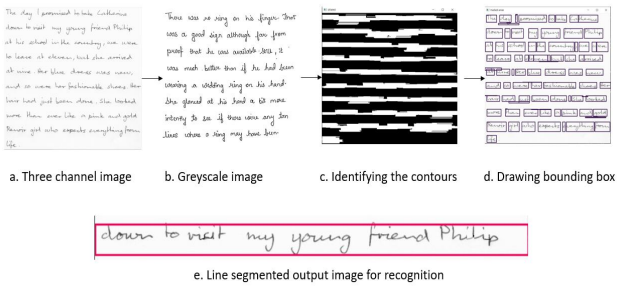


Fig 2: Segmentation of a given paragraph into lines

Segmentation techniques include threshold-based segmentation, edge-based segmentation, region-based segmentation, clustering-based segmentation, and Artificial Neural Network-based segmentation [9]. In our project, we used the Region Based Segmentation method. A region is a collection of connected pixels having comparable properties. Pixels may resemble each other in terms of intensity, colour, or other characteristics. There are several recognized rules that a pixel must satisfy in order to be grouped into related pixel regions in this type of segmentation. We prefer region-based segmentation to edge-based segmentation when working with noisy images [9]. This methodology is further separated into two types based on different methodologies: region growing and region splitting and merging.

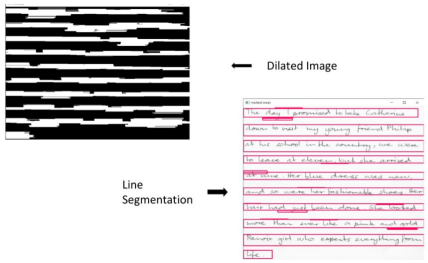


Fig 3: Segmentation of a given paragraph.

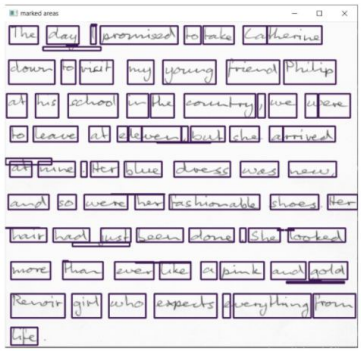


Fig 4: Segmentation of individual words in a line.

B. CNN- Feature Extraction

Machine Learning needs early Feature Extraction as features, whereas Deep Learning operates as a "black box," extracting and classifying functions on its own. When an image is given as an input to Convolutional Neural Networks, the feature is taken by each layer of CNN, and the Fully Connected Layer does categorization. First, we extract the key elements from the handwritten line text image using a Convolutional Recurrent Neural Network. The output of the CNN FC layer (128x64) is sent to the BLSTM, which is used to perform sequence dependence and time-sequence operations. The RNN is then trained using CTC LOSS, which eliminates the Alignment problem in handwritten because each writer's alignment differs. We have trained it on what is written in the image (Ground Truth Text) and BLSTM output. It then calculates the loss simply as a log("gtText"); to minimize the negative maximum probability path. Then, CTC will find out all the possible ways from the given labels. Loss for (X, Y) pair can be expressed as:

$$p(Y \mid X) = \sum_{A \in \mathcal{A}_{XY}} \prod_{t=1}^{T} p_t(a_t \mid X)$$

The CTC conditional **probability** | **marginalizes** over the set of valid alignments | computing the **probability** for a single alignment step-by-step.

Fig. 5: CTC Loss Formula

The segmented images obtained from the 1st step are fed into our trained model which displays the predicted sentence.

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 10 Issue III Mar 2022- Available at www.ijraset.com

C. Use Case Diagram

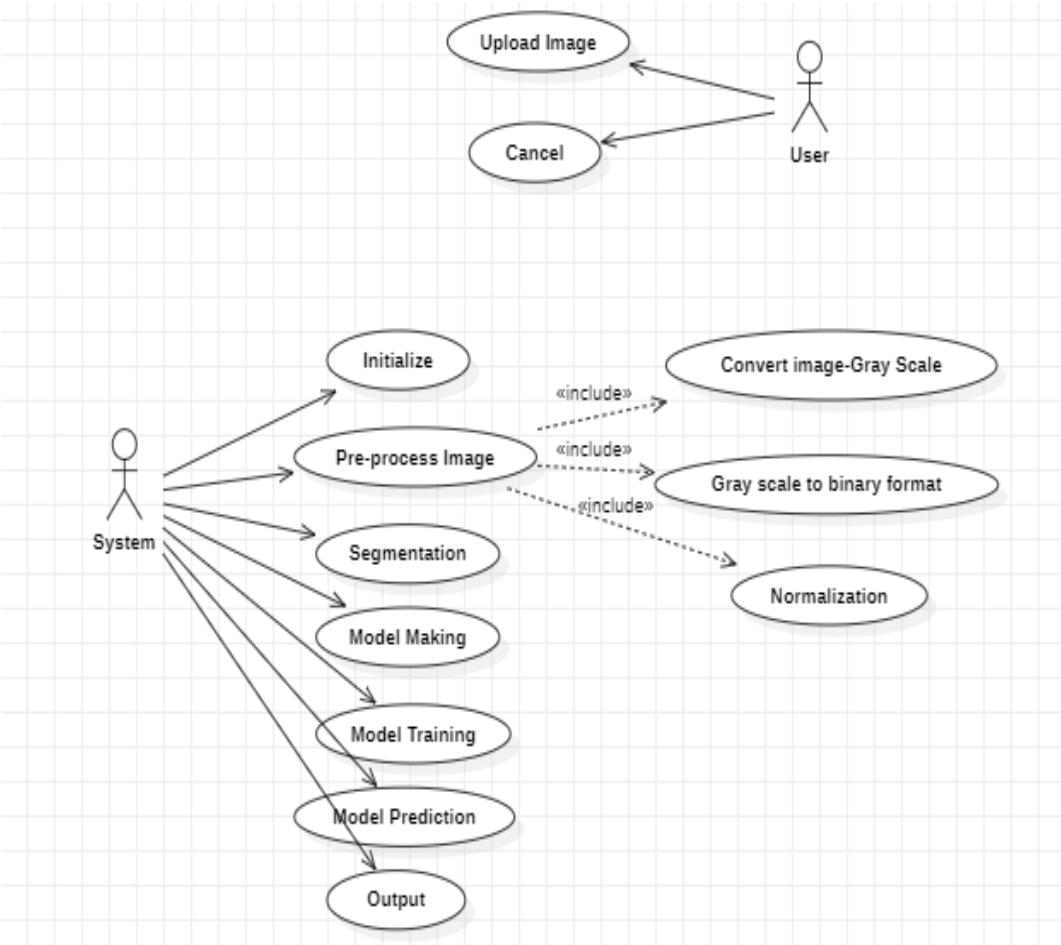


Fig. 6: Use-Case Diagram for Handwriting to Text Conversion.

D. AutoCorrect using Dictionary

The predicted word is passed through the autocorrect library. In this, it is autocorrected to the nearest word using the English dictionary. It is a matching done at word level to find the closest match. This helps us to correct the predicted word in case if any word is predicted wrong.

```
Predicted Word: rrived
Predicted Word using Autocorrect: arrived
```

Fig 7: Prediction using AutoCorrect

E. Dataset Description

The dataset used for training the model is IAM dataset. We have built a Neural Network (NN) which is trained on word-images from the IAM dataset. Characteristics of IAM Dataset - 657 writers contributed samples of their writings, 1532 pages of scanned text, 5685 isolated and labelled sentences, and 115320 isolated and labelled words [5]. The IAM dataset is split into 95% training data and 5% validation data.

Table I-Dataset

Training Data	87292
Testing Data	4316
Validation Data	4316

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 10 Issue III Mar 2022- Available at www.ijraset.com

IV. ARCHITECTURE OF PROPOSED WORK

A. CNN (Convolution Neural Networks)

In this section, we study the usage of convolutional neural networks (CNNs) in the context of handwritten recognition. The machine will use the convolutional neural network (CNN). Fig. suggests an underlying structure of CNN with a purpose to be used in the recognition model. The structure suggests specific kinds of layers, with the primary layer being the input layer and the final layer being the output layer. The 2nd layer is known as the convolutional layer and is accompanied via way of means of pooling layers and convolutional layers.

The architecture of the CNN structure is as follows:

1) *Input Layer:* The input layer in CNN must incorporate image information. Image information is represented via way of means of a 3-dimensional matrix. We need to reshape and convert it into a single column. Suppose you've got got a picture of size 28 x 28 =784, you need to transform it into 784 x 1 before feeding it into the input. If you've got got "m" training examples then the dimension of input will be (784, m).
2) *Convolution layer*: The convolution layer is the constructing block of the entire network. Most of the computational work that is required to recognize characters from the input is completed on this layer. The layer includes a set of learnable filters called parameters of the convolution layer.
3) *Pooling Layer*: The pooling layer is used to reduce the spatial volume of the input picture after convolution. It is used between convolution layers. If we apply FC after the Convo layer without making use of pooling or max pooling, then it is going to be computationally high-priced and we don't need it. So, max-pooling is the best manner to lessen the spatial extent of an enter picture.
4) *Fully Connected Layer* (FC): Fully connected layer includes weights, biases, and neurons. It connects neurons in a single layer to neurons in another layer. It is used to categorize images among different classes via way of means of training.

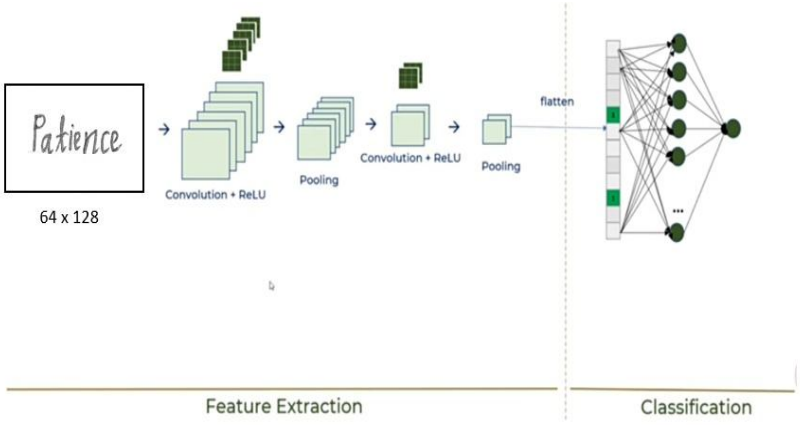


Fig. 8. CNN Architecture

B. Long-Short term Memory

LSTM networks have contextual state cells that operate as long-term or short-term memory cells, to put it another way. The status of such cells is used to regulate the output of the LSTM network. This is a critical quality when we want the neural network's prediction to be based on the historical context of inputs rather than just the final input. Long Short-Term Memory (LSTM) requires the recognition of a Recurrent Neural Network (RNN), which is a special type of RNN. The output from the previous step is provided as input to the current step in an RNN, which is a type of Neural Network (NN).

V. RESULT

The results are promising. The validation accuracy we obtained is 98.20%. We have obtained a test accuracy of 98.42% which is better than most other models. The accuracy was calculated by using the formula-

(Correct Predictions/ Total Predictions) X 100.

We found the correct predictions by checking if the original string is a substring of the predicted word/string. The accuracy obtained of the application makes it suitable for most practical purposes, Although the accuracy can be improved.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |



Fig 9: Input and Output of Handwritten paragraph

VI. CONCLUSION AND FUTURE SCOPE

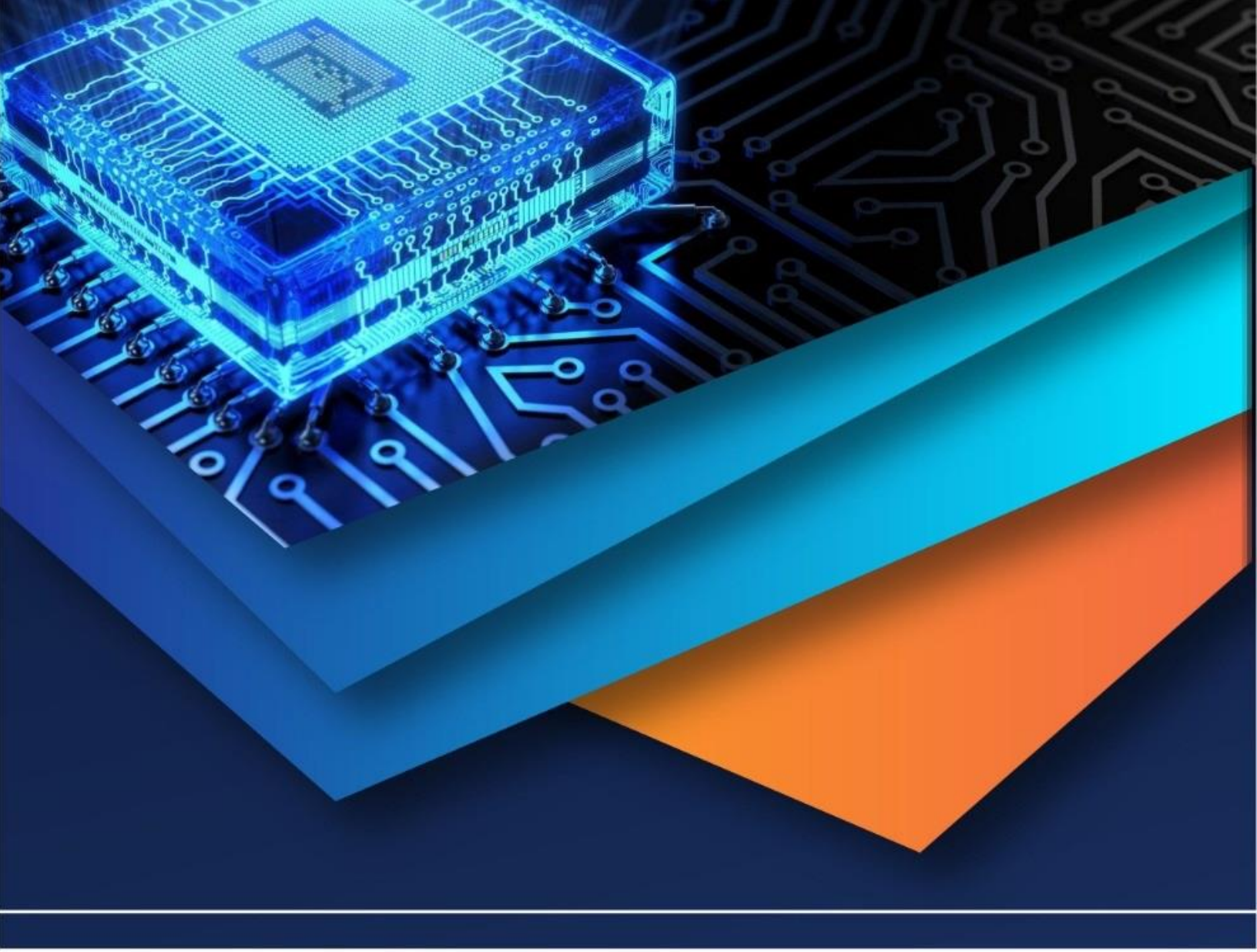
Thus, our project will take handwritten English Characters in the form of a paragraph as an input, process the paragraph which will have support for cursive writing and symbols, then the CNN model trained on IAM dataset will predict the text. In this project we have presented an innovative method for handwritten character detection using deep neural networks for enhancing existing systems. Healthcare start-up where they can digitalize the health records of patients.

Digitizing the prescription would help mitigate these concerns to a great extent and ease the patient's task of understanding the doctors handwriting. As we all know, the online exams have increased due to COVID-19. Hence, this system can be used in correction of papers that are handwritten and uploaded. The institute can provide a model answer and the system can read the handwritten text and semantically compare with the model answer and grade it.

We can use MDLSTM to recognize entire paragraph at once Scan, Attend and Read: End-to-End Handwritten Paragraph Recognition with MDLSTM can be done. The future is completely based on technology and no one will use the paper-pen for writing. In this scenario, people will write on touch pads. We can build an inbuilt software which can automatically detect text while they are writing and convert into digital text.

REFERENCES

[1] Handwriting Recognition using Artificial Intelligence Neural Network and Image Processing in (IJACSA) International Journal of Advanced Computer Science and Applications, Vol. 11, No. 7, 2020.
[2] Handwritten Character Recognition Using Deep-Learning in Proceedings of the 2nd International Conference on Inventive Communication and Computational Technologies (ICICCT 2018) IEEE Xplore Compliant - Part Number: CFP18BAC-ART; ISBN:978-1-5386-1974-2.
[3] Handwritten Text Recognition using Deep Learning with TensorFlow in International Journal of Engineering Research & Technology (IJERT) ISSN: 2278-0181 IJERTV9IS050534 Vol. 9 Issue 05, May-2020.
[4] Handwritten Text Recognition System Based on Neural Network in International Journal of Advanced Research in Computer Science & Technology (IJARCST 2016) Vol. 4, Issue 1 (Jan. - Mar. 2016)
[5] Dataset:https://fki.tic.heia-fr.ch/databases/iam-handwriting-database
[6] "OpenCV" https://en.wikipedia.org,[Online] Available: https://en.wikipedia.org/wiki/OpenCV/.
[7] https://en.wikipedia.org/wiki/Region_growing
[8] https://towardsdatascience.com/image-segmentation-part-2-8959b609d268
[9] Manoj Sonkusare and Narendra Sahu "A SURVEY ON HANDWRITTEN CHARACTER RECOGNITION (HCR) TECHNIQUES FOR ENGLISH ALPHABETS" Advances in Vision Computing: An International Journal (AVC) Vol.3, No.1, March 2016.
[10] J.Pradeep, E.Srinivasan and S.Himavathi ––Diagonal based feature extraction for handwritten alphabets recognition system using neural network‖ - International Journal of Computer Science & Information Technology (IJCSIT), Vol 3, No 1, Feb 2011.
[11] Anita Pal and Davashankar Singh,"Handwritten English Character Recognition Using Neural Network", International Journal of Computer Science and Communication, pp: 141- 144, 2011.
[12] U.-V. Marti and H. Bunke. Text line segmentation and word recognition in a system for general writer independent handwriting recognition. In Proc. 6th Int. Conf. on Document Analysis and Recognition, pages 159– 163



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)