



e-ISSN: 2319-8753 | p-ISSN: 2347-6710

IJRSET

International Journal of Innovative Research in
SCIENCE | ENGINEERING | TECHNOLOGY

A futuristic digital interface with a glowing Earth globe in the center. The interface features various data visualizations including bar charts, line graphs, and pie charts. Text elements like 'BUSINESS', 'NETWORK SEARCH', and 'WORLD' are visible. The background is a dark blue with light blue and white data points and lines. A hand is visible at the bottom, interacting with the interface.

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

Volume 12, Issue 4, April 2023

ISSN INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA

Impact Factor: 8.423

 9940 572 462

 6381 907 438

 ijrset@gmail.com

 www.ijrset.com



Optical Character Recognition Notes Application

Saloni Khatu¹, Kevin Parmar², Aditya Mane³, Rajat Singh⁴, Priya Karunakaran⁵

U.G. Student, Department of Computer Engineering, St. Francis Institute of Technology, Mumbai,
Maharashtra, India^{1,2,3,4}

Associate Professor, Department of Computer Engineering, St. Francis Institute of Technology, Mumbai,
Maharashtra, India⁵

ABSTRACT: Even currently, handwritten documents have an omnipresence that makes it difficult to disregard them. A great deal of people still prefers to note things down in a more conventional fashion. In a manner that is almost habitual, people favour the use of pen and paper to note things down. However, this poses a problem of longevity and security. A key objective in our solution to this problem is developing a Handwritten Text Recognition (HTR) System that will allow users to convert their notes into digital text and integrate in the system various useful functionalities that will enable users to store, share and manage their notes in a user-friendly manner. Handwritten Text Recognition or Character Recognition is enabling a computer to derive and recognize the characters from an image. Much research has been conducted on this technology in the past decade, making it a highly discussed topic. Even so it remains a big challenge to correctly implement it in such a way that it provides substantial precision in a diverse range. HTR is a technology that is proving to be of great use currently. Even though Optical Character Recognition systems for the English language are in existence, yet HTR systems that process handwritten text are not as common and often provide sub-par precision. Our goal is to create a Notes Application with Handwritten Text Recognition capabilities. We aim to achieve this using Artificial Neural Networks for Handwritten Character Recognition along with additional functionalities to store, sort, share and manage the notes.

KEYWORDS: Handwritten Text Recognition, Optical Character Recognition, Artificial Neural Networks.

I. INTRODUCTION

We aim to develop an OCR Notes Application which will provide Handwritten Text Recognition capabilities to help people store and manage their notes anytime and anywhere while avoiding damaging or loss of their notes. We plan to use Artificial Neural Networks (ANN) to recognize handwritten characters. Apart from this, many other functionalities such as segregating, storing, sharing, and managing notes will also be included in the application. We plan to create an OCR Notes Application to manage the users notes. Even currently, many people still prefer to note down things in a traditional way using pen and paper as it may be convenient in some ways. Still there is a need to manage these notes and the quantity of these may quickly get out of hand. As such our goal is to create an OCR Notes Application which will deliver Handwritten Text Recognition functionalities enabling people to store and manage their at any time and place while eliminating a chance of damage or loss of their notes. Thus, we plan to create an application that converts the user's handwritten notes into digital text that can be then stored and managed together. The user can also edit this scanned text and share it with others. We plan to use Artificial Neural Network (ANN) to scan and recognize handwritten characters accurately and convert them into digital text.

II. RELATED WORK

Inpaper [1], the researchers implemented a CNN based system to classify and recognize characters from the NIST Dataset. In paper [2] ussing AutoEncoder and KNN, SVM, in another research, the authors present four methods using a vanilla Autoencoder and a 6 Convolutional Autoencoder. For classification purpose they have used KNN, SVM based classifiers such as hybrid KNN-SVM and m-SVM. In paper [4], a handwritten text recognition system based on Artificial Neural Network is demonstrated. The paper portrays the effectiveness of Artificial Neural Networks in tasks that involve pattern recognition, and that ANN can prove to be a life saver in the development of Handwritten Character Recognition Systems.Paper [5] proposes an efficient fuzzy method for handwritten character recognition.



Comparison among OCR, ANN, ICR and IWR in paper [6] gives a terse analysis of different handwritten recognition techniques, such as Optical Character Recognition (OCR), Artificial Neural Network (ANN), Intelligent Character Recognition (ICR), and Intelligent Word Recognition (IWR). Accuracy of all the mentioned methods were compared, and the relative analysis demonstrates that OCR method is the best among them for the 5 recognitions of English handwritten characters.

III. METHODOLOGY

A. Datasets

1. NIST Database - More than 800,000-character pictures of handwriting from 3600 authors are published by the US National Institute of Science.
2. MNIST Database - The MNIST dataset is considered as one of the most used/cited datasets for handwritten digits. It is the part of the NIST dataset, and that is why it is called modified NIST or MNIST. 10,000 test photos and 60,000 training images make up the dataset. Samples are normalized into grayscale images with reserved aspect ratio, and they are normalized. The dataset hugely decreases the time required for pre-processing and formatting, because it is already in a normalized form.
3. IAM Database - The IAM is a handwritten database of English language. The IAM database consists of 13,353 images of handwritten lines of text created by 657 writers. It includes contributions from 657 authors aggregating 1,539 handwritten pages consisting of 115,320 words and is classified as part of modern collection. The database is labelled at the word, line and sentence levels.

B. Requirements

The functional requirements are classified as follows:

- Choosing the desired text image: The user must be able to choose the images from the directories present in the system.
- Recognition of the text: The text will be recognized from any chosen image from the computer's directory.
 - Recognition of the text from the image.
 - Text from the image should be ready to use.
- Copying the text for different uses: Once the text is recognized and ready to be used, the user will be able to copy, edit, and modify it. He/she may also be able to retrieve the data from the image and store it directly.
- Creating a note: User should be able to create a note and write some text in it.
- Sharing a note: User should be able to share a note to other users.
- Searching for a note: User should be able to search for a particular note using keywords appearing in the note. System should retrieve the note for the user.
- Note Highlighting and Tagging: User should be able to highlight the text using various colors and can add tags to notes for better searching.

The non-functional requirements are:

Product Requirements:

- Usability Requirements: The application will be built in such a way that it is user friendly and doesn't require any guidance to be used. In other words, the application should be access-friendly so its users shall use it easily. The interface should be quite simple and straight forward so that anyone can understand it. The user should only click on the import picture button.
- Reliability Requirements: To prevent any failure's occurrence, the specifications have been thoroughly followed correctly. The only issue is that the application may not recognize 100% of the characters accurately.
- Security Requirements: Any of the data collected about the user would be kept anonymous and the privacy of the person would be not compromised.
- Efficiency Requirements: The application response time shall be adequate and sufficient, that's why the time required for this application to response to its user's actions must be managed and controlled. But to maintain the performance of the application, the user must follow the required steps to get the desired result.
- Portability Requirements: The application will be built to be compatible enough to run on any Windows version.



C. Training

The first part of this project is the training part. It starts by importing the dataset. The dataset images are pre-processed. 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 image and make a clear difference between foreground and background. After pre-processing the features are extracted. The feature vector is computed by converting the pre-processed image into bit mapped. The bit map version preserves the major features of input image in shorter space/data length. Feature Vectors are generated. The next step is to train the neural network. The training of ANN involves specifying the hidden layers and choice of 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. The model is saved after training.

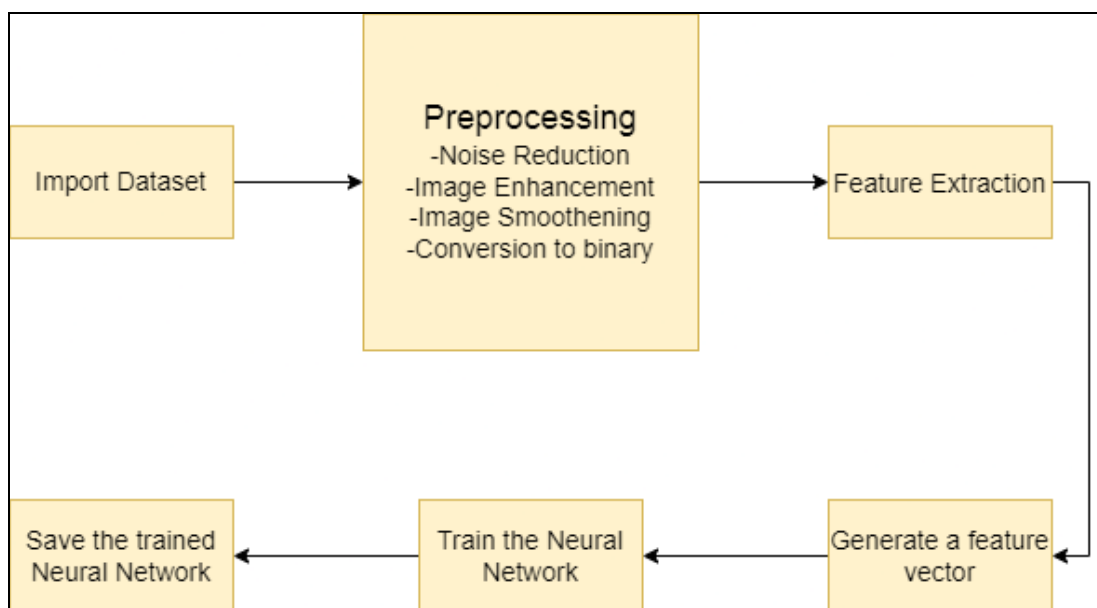


Figure 3.1 – Training Flow Chart

D. Testing

The second part of the system is the testing part. In this phase an image is selected. The pre-processing done in the training part is done on the input image and the image is segmented. In this, a complete image is decomposed into a sequence of text/sub images of individual text. The segmentation is done based on edge detection and gap between the different characters. After segmentation, the sub-divided parts are labelled and then processed further one by one. For each subdivided part, the features are extracted, and it is classified using the saved model. After all sub parts are classified, the output text is generated. the original content as displayed in fig. 3.1. Handwritten notes have a lot of problems to begin with but despite these shortcomings, people still make use of pen and paper to take notes. To provide the users with the essential functionalities of digital notes while allowing them to use the traditional way of note taking, we proposed a system that will allow users to make use of Handwritten Text Recognition functionality to extract text from images and store them safely. This will help users to take notes in the traditional way and take advantage of computerized system to edit, store and share them.

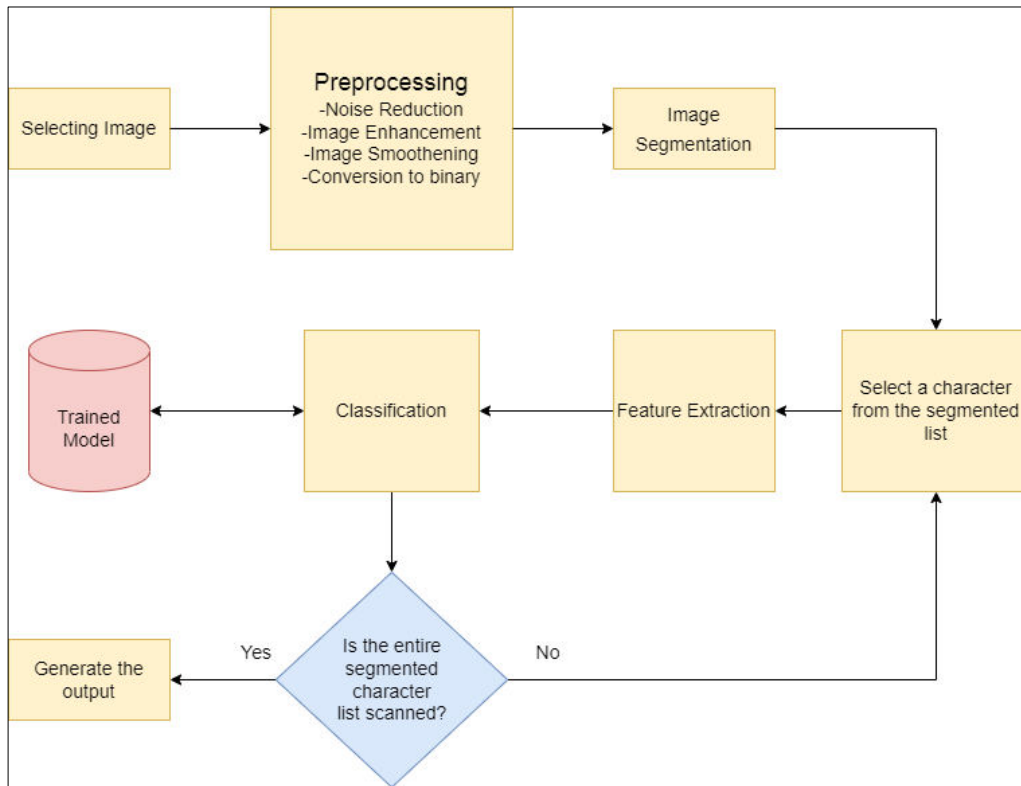


Figure 3.2: Testing Flow Chart

IV. EXPERIMENTAL RESULTS

A. Experimental Setup:

The fig. 3.2 denotes that the model is a Convolutional Neural Network (CNN) that is used for image recognition tasks. It has 7 layers, where the first 6 layers are convolutional layers with different filter sizes and activation functions, followed by pooling layers to down sample the feature maps. The last layer is a fully connected layer that outputs the final predictions.

The output from the CNN layers is fed into an RNN layer composed of 2 layers of LSTM cells. The RNN is used to capture the sequential dependencies in the input data by processing the input in a time-dependent manner. The output of the RNN layer is a sequence of character probabilities corresponding to the input image.

The Connectionist Temporal Classification (CTC) loss function is used to train the RNN by aligning the predicted text with the ground truth text, which corresponds to the input image. The aim of the CTC loss function is to minimize the negative maximum likelihood path by updating the weights of the RNN in a backpropagation manner, which helps the RNN to learn to predict the correct text for a given input image. By using CTC, the model can learn to recognize and label sequential data, such as handwritten text, which is often difficult to recognize due to variations in handwriting styles. The output of the RNN is a sequence of character probabilities that correspond to the input image, which can be further processed to obtain the final text output.

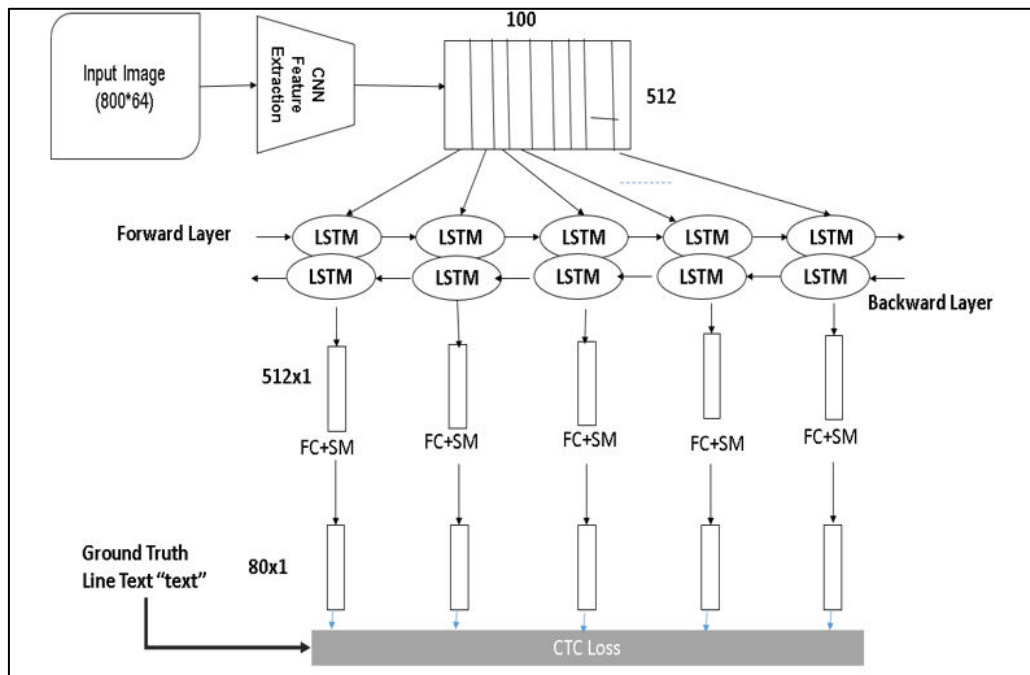


Figure 4.1 – Architecture detail

B. Dataset:

The dataset used in this project is the IAM dataset. It consists of two main parts: a training set and a test set. The training set contains 9,862 lines of handwritten text images with corresponding transcriptions, while the test set contains 3,859 lines of handwritten text images with corresponding transcriptions. Additionally, in fig. 4.1, a validation dataset of 400 grayscale images of handwritten text lines with corresponding ground truth transcriptions was used to validate the accuracy of the model on new, unseen data.

C. Results:

The trained model achieved a CER of 8.32% on the test set, indicating its accuracy in recognizing handwritten text. The validation dataset was used to evaluate the accuracy of the model on new, unseen data and achieved a CER of 14.4%. These results demonstrate the effectiveness of the CNN-RNN-CTC model architecture in recognizing handwritten text and its ability to generalize well to new, unseen data. Further experimentation and optimization can be done to improve the performance of the model on this task.

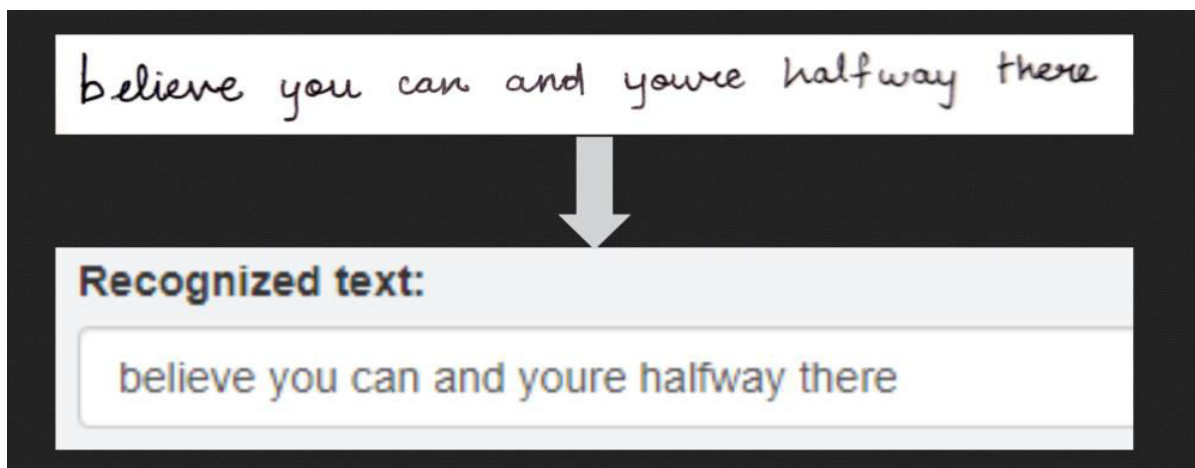


Figure 4.2 – Sample HTR result on handwritten data.



V. CONCLUSION

In fig. 4.2, the CNN-BLSTM-CTC algorithm has several strengths when used for handwritten text recognition. The CNN component is effective at feature extraction, making it possible to identify relevant patterns and structures in the input data. The BLSTM component allows the model to capture sequential dependencies and model the context of the text, which is important in recognizing handwriting. Finally, the CTC loss function enables the model to learn to align the predicted text with the ground truth text, which helps in minimizing recognition errors.

However, limitations of handwritten text recognition systems include difficulty in recognizing cursive writing, and variability in writing styles, which can make it challenging to train models on different handwriting styles. Additionally, handwriting recognition is dependent on the quality of the input image, so noise or blurriness can impact the accuracy of the recognition.

An HTR (Handwritten Text Recognition) system extracts and repurposes data from scanned documents, camera images and image-only pdfs. HTR software singles out letters on the image, puts them into words and then puts the words into sentences, thus enabling access to and editing of the original content. Handwritten notes have a lot of problems to begin with but despite these shortcomings, people still make use of pen and paper to take notes. To provide the users with the essential functionalities of digital notes while allowing them to use the traditional way of note taking, we proposed a system that will allow users to make use of Handwritten Text Recognition functionality to extract text from images and store them safely. This will help users to take notes in the traditional way and take advantage of computerized system to edit, store and share them.

REFERENCES

- [1] I Khandokar, M Hasan, F Ernawan, S Islam, and MN Kabir. Handwritten character recognition using convolutional neural network. In *Journal of Physics: Conference Series*, volume 1918, page 042152. IOP Publishing, 2021.
- [2] Dibyakanti Mahapatra, Chandrajit Choudhury, and Ram Kumar Karsh. Handwritten character recognition using knn and svm based classifier over feature vector from autoencoder. In *International Conference on Machine Learning, Image Processing, Network Security and Data Sciences*, pages 304–317. Springer, 2020.
- [3] Jamshed Memon, Maira Sami, Rizwan Ahmed Khan, and Mueen Uddin. Handwritten optical character recognition (ocr): A comprehensive systematic literature review (slr). *IEEE Access*, 8:142642–142668, 2020.
- [4] [Ahmed Mahdi Obaid, Hazem M El Bakry, MA Eldosuky, and AI Shehab. Handwritten text recognition system based on neural network. *Int. J. Adv. Res. Comput. Sci. Technol*, 4(1):72–77, 2016.
- [5] Romesh Ranawana, Vasile Palade, and GEMDC Bandara. An efficient fuzzy method for handwritten character recognition. In *International Conference on Knowledge-Based and Intelligent Information and Engineering Systems*, pages 698–707. Springer, 2004.
- [6] Usha Tiwari, Monika Jain, and Shabana Mehruz. Handwritten character recognition and analysis. In *Advances in System Optimization and Control*, pages 207–212. Springer, 2019.



SJIF Scientific Journal Impact Factor

Impact Factor: 8.423



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

 9940 572 462  6381 907 438  ijirset@gmail.com



www.ijirset.com

Scan to save the contact details