IJNRD.ORG ISSN : 2456-4184
INTERNATIONAL JOURNAL OF NOVEL RESEARCH AND DEVELOPMENT (IJNRD) | IJNRD.ORG
An International Open Access, Peer-reviewed, Refereed Journal

Handwritten to Text Conversion System

N.Sai Sree
Department of Computer Science
GITAM University
Visakhapatnam,India

U.Sai kiran
Department of Computer Science
GITAM University
Visakhapatnam,India

Uday
Department of Computer Science
GITAM University
Visakhapatnam,India

Sandeep Varma
Department of Computer Science
GITAM University
Visakhapatnam,India

Abstract*—The goal of this paper is to categorize every distinct handwritten word, making it possible to digitize handwritten images. We primarily used character segmentation and direct categorization of words techniques to finish this study. To create a model that can accurately characterize words, we can create a convolutional neural network for the former(CNN) with a variety of alternative topologies. For the latter, we use convolutional and long Short-Term Memory networks(LSTM) to generate bounding boxes for each character. The characters are segmented before being sent to CNN for classification. Finally, using the data from classification and segmentation, we recreate each word.

***Keywords*—Hand written text,CNN,LSTM,RNN,RESNET**

I. Introduction

This paper objectives are to similarly look at the process of identifying handwritten text and converting it to a virtual representation. Because the phrase "handwritten textual content" is sort of standard, we sought to the consciousness of the venture's objectives by giving it a particular definition. In this challenge, we set ourselves the hard task of categorizing any picture of a handwritten word, whether or not it's in the block or cursive writing. This paper may be paired with algorithms that separate the words from traces in a given picture, which can then be used with algorithms that separate strains from photographs of a whole handwritten page.

Handwritten-to-text conversion is the process of transforming handwritten documents or notes into digital text that can be easily edited, searched, and shared. This technology has gained popularity in recent years as it provides an efficient and time-saving way of converting handwritten notes into digital format. Handwritten-to-text conversion software uses optical character recognition (OCR) technology to scan and recognize characters from an image of the handwritten text. The software then converts the recognized characters into digital text that can be saved, edited, and shared in various formats. This technology has several applications in fields such as education, healthcare, finance, and legal industries, where handwritten notes and documents are still prevalent. With those extra elements, our mission can have the appearance of a running prototype that a stop user might use. This version might be fully purposeful and assist the consumer in addressing the trouble of changing handwritten files into a virtual layout by prompting the consumer to take a picture of a page of notes. Observing that at the same time as our version requires positive other layers to be constructed on top of it to produce an output for a ceaseless person that is functioning, we chose cognizance on categorization because we find it to be the most thrilling and difficult thing of the mission.

Due to the fact, CNNs normally carry out a good deal better on raw input pixels rather than features or qualities of a photo, we deal with this task with entire phrase photographs. In light of the outcomes that we received utilizing entire phrase images, we tried to enhance by removing characters from each word photograph and then classifying every character separately to reconstruct a whole phrase. In the end, the fashions utilized by each of our strategies take a phrase's picture as input and output its name

II. Related Works

Handwritten text Recognition has a wide area of research due to its vast applications. The recognition systems can be divided into two major steps. The first step is feature extraction from handwritten images and the second one is classification. Alex Graves[1,2] and researchers suggested different methodologies for feature extraction and this paper presents a novel method for training RNNs to label unsegmented sequences directly, thereby solving both problems Selection of a feature extraction method is probably the single most important factor in achieving high recognition performance in character recognition systems. Different feature extraction methods are designed for different representations of the characters, such as solid binary characters, character contours, skeletons (thinned characters), or gray-level sub-images of each character. The feature extraction methods are discussed in terms of invariance properties, constructability, and expected distortions and variability of the characters. Harold Schiedl[3] used an implementation using TF and some important parts of the code were presented. S. Johansson[4] hints to improve the recognition accuracy were given. The true "representativeness" of the present corpus arises from the deliberate attempt to include relevant categories and subcategories of texts rather than from blind statistical choices. Elie Krevat[5] Performance of our HMM is compared to a baseline Naive Bayes classifier. Experimental results are given for different variants of our HMM algorithm, with optimizations for ignoring the effects of inter-word transitions when applying the Viterbi algorithm to return the most likely character sequence. Yann LeCun[6] as we all know that accuracy mainly depends on the Gradient Descent this Researcher has proved that with the best example of a successful GradientBased Learning technique Given an appropriate network architecture GradientBased Learning algorithms can be used to synthesize a complex decision surface that can classify high-dimensional patterns such as handwritten characters with minimal preprocessing This paper reviews various methods applied to handwritten character recognition and compares them on a standard handwritten digit recognition task Convolutional Neural Networks that are specifically designed to deal with the variability of D shapes are shown to all other technique. Marcus Liwicki [7] This system is based on the combination of several individual classifiers of diverse nature. Recognizers based on different architectures(hidden Markov models and bidirectional LSTM.

III. MATERIALS AND METHODS

A. LSTM

Recurrent neural networks (RNNs), which were first developed to address circumstances when RNNs fail, were preceded by LSTM networks. An RNN is a type of network that takes into account (feeds back) earlier outputs for the current input and retains them in short-term memory. Among the various applications, the most popular are in the areas of composition, on-Markovian control, and speech processing. Still, RNNs have drawbacks. First off, data cannot be kept for a very long time. To predict current performance, it may be necessary to refer to specific information stored long ago. But RNNs can never handle such "long-term dependencies"

B. RNN

To predict future value, we need a mechanism to retain past or historical information. RNNs, short for "recurrent neural networks," are a kind of conventional feedforward artificial neural networks which can handle sequential data and can be taught to retain knowledge about the past. These deep learning algorithms are frequently utilized for order or timing issues, such as speech translation, natural language processing (NLP), speech recognition, and image annotation. They integrate with popular apps like Google Translate, Siri, and voice searches. Recurrent neural networks (RNNs) learn from training input in a manner akin to feedforward and convolutional neural networks (CNNs).

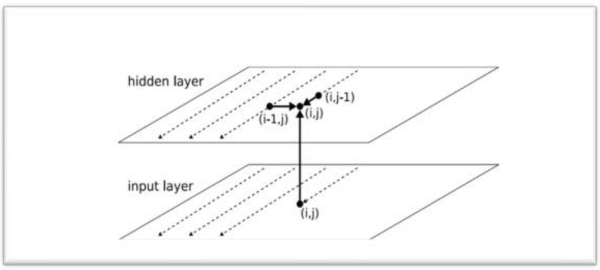


Fig 1. RNN Layers

C. RESNET:

Here is anticipated the development of deeper architecture networks to increase model accuracy since deep learning has advanced tremendously over the past few years. These methods are frequently employed in the classification,

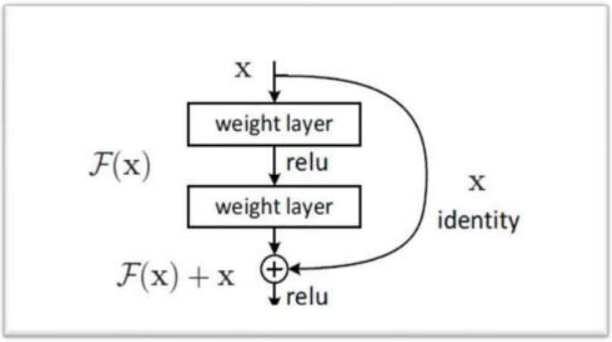


Fig 2. Residual CNNs for Image Classification Tasks

grouping, and synthesis of images. Although going deep can seem cool, neural networks have a problem known as deterioration, therefore doing so is useless. Here, precision is crucial. Decreasing Gradient Drop is another issue that results from this. As a result, the weights are unable to update correctly during the backpropagation stage. The chain rule is used in the stage of backpropagation. Each layer's derivatives are compounded as we move lower down the network. When deeper networks start to converge, though, degradation issues are noticed. Accuracy reaches a saturation point as network depth grows.

D. IMPLEMENTATION

1. **Data Pre-Processing And Data Augmentation:**
   Our handwriting recognizer was trained mostly using the IAM Handwriting Dataset. There are more than 1500 forms in this dataset that have handwritten information. It is a paper with over 600 lines of text, over 5500 sentences, and over 11500 words written by over 600 authors. The corresponding XML files offer all relevant form label metadata. After segmenting and meticulously validating the words, the source text was taken from the Lancaster-Oslo/Bergen (LOB) corpus, which contains texts of whole English phrases with a word count of over 1 million. Also included in the database are 1,066 forms created by around 400 different authors.
2. **Grayscale Conversion:**
   It is a pre-processing technique in which the images which need to be converted are changed into the black-and-white format, we do this because basically as we know text in images can be read clearly when they are in black-and-white, and generally, it can reduce the complexity making it easier for the model.
3. **Padding images:**
   Before using our dataset to train our models, we used several pre-processing and data augmentation techniques to make it more compatible with the models and more flexible to real-world scenarios. Because words can be different heights and lengths, there are many sizes of images. Due to the length of the phrases, the word "error" creates a smaller image than the word "congratulations,". Like every other convolutional neural network architecture, ours is based on the assumption that the input images would all be the same size.

4. **Rotating images**

Several words showed a small tilt, even though each word's image was separately recorded for our dataset. This occurred because participants in the dataset were told to write on blank paper without any lines, which caused some of the words to be written more awkwardly. Whether lines are there on paper or not, this condition happens regularly in real life. To increase the resistance of our training data to this issue, we decided to include a photo in our training set that had been randomly rotated to the right by a very small angle. We were able to enhance our model against some relatively common but minor features that could occur by using this data augmentation technique.

5. **Page Segmentation**

This is the first step of the process where our model detects the handwritten area from the given image file, this method is mainly used in detecting the texts of old documents of history like manuscripts, such documents consist of decoration, backgrounds, and margins, as we know that convolution is a type of encoder-decoder model so that it can recognize strong features of such documents automatically, this features which are detected are sent to a support vector machine to classify and learn the features to detect the handwritten area in documents using segmentation.

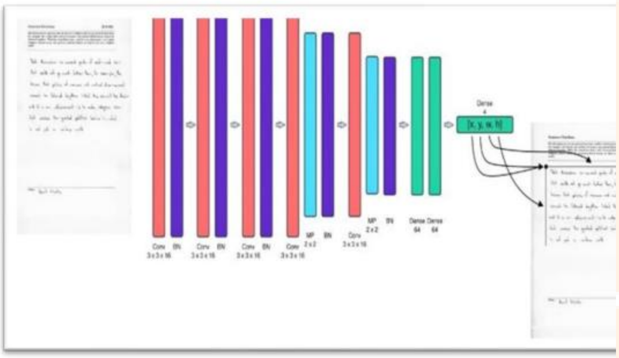
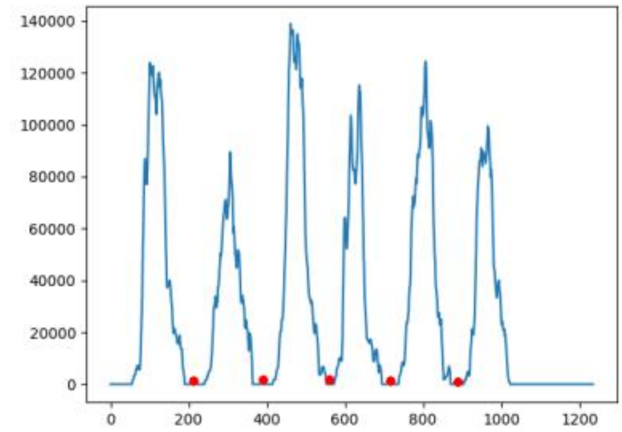


Fig 3.CNN layers in page segmentation

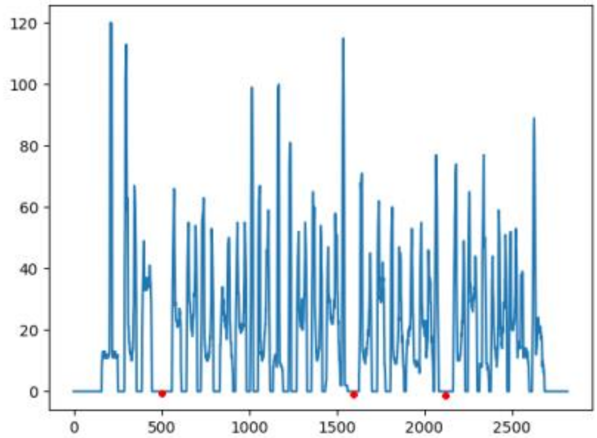
6. **Line Segmentation**

From the previous step the DCNN model recognizes the handwritten area and proceeds the output to the next step here in this method the ha- add written text area which was detected is segmented into lines, concisely this step divides each line of the paragraph and recognizes each of the handwritten lines which can be used by further steps in the process, this model takes the images as input which consists of handwritten text we can get this input from the previous page segmentation output which consists of borders.



(a)

Fig. 4: Horizontal Histogram of Image



(a)

Fig. 5: Vertical Histogram of Image

IV. EXPERIMENTAL RESULT

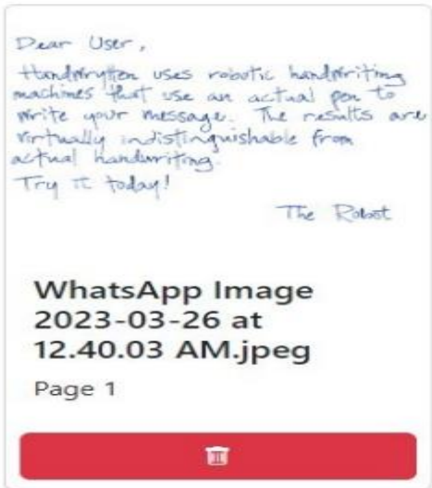


Fig 6. Selected image

The image that the user selected and they have their choice The chosen 22 image will be displayed in this instance. Now, if you want to make any modifications, such as deleting the image, we can do that too. The subsequent loading of our photograph into PDF text format will take a little while. Finally, a pdf text version of our handwritten image is downloaded.

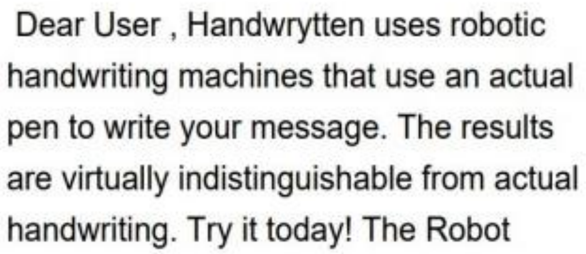


Fig 7.Final image.

V. CONCLUSION AND FUTURE SCOPE

There are various typefaces for numerous regional languages all around the world. by the HCR system utilizing the proper methods and techniques. We are learning how to recognize the alphabet. Because there are many unusual characters and certain characters resemble one another in shape, some people have trouble reading handwritten characters. Characters are separated by individual letters once the scanned image has been cleaned up and pre-processed. Prior processing is done to do normalization and filtering employing procedures that result in output that is clear and noise-free. The Control of Evolutionary Algorithms An effective training program would evaluate many stepwise methods to provide a more efficient system output. Better letter recognition for English letters is achieved by neural networks using some statistical and geometric aspects. This development will make it easier for researchers to use different scripts. The research on character segmentation in other languages has since been added. faxes and newspapers can be converted to text format with this tool. You can utilize numerous ANNs for classification and use them at the post office to detect postal codes and identify words and sentences in paragraphs.

REFERENCES

[1] Alex Graves, Santiago Fernandez, Faustino Gomez, Jrgen Schmidhuber, Connectionist temporal classification: labeling unsegmented sequence data with recurrent neural networks, Proceedings of the 23rd international conference on Machine learning. 2006

[2] Oivind Due Trier, Anil K. Jain, Torfinn Taxt. Feature Extraction Methods for Character Recognition–A Survey. Pattern Recognition. 1996

[3] Build a Handwritten Text Recognition System using TensorFlow - Harold Schiedl.https://towardsdatascience.com/build-a-handwritten-text-recognition-systemusin TensorFlow-2326a3487cd5

[4] S. Johansson, G.N. Leech, and H. Goodluck. Manual of Information to accompany the Lancaster Oslo/Bergen Corpus of British English, for use with digital Computers. Department of English, University of Oslo, Norway, 1978.

[5] Elie Krevat, Elliot Cuzzillo. Improving Off-line Handwritten Character Recognition with Hidden Markov Models.

[6] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner: Gradient-Based Learning Applied to Document Recognition, Intelligent Signal Processing, 306-351, IEEE Press, 2001

[7] M. Liwicki, H. Bunke, J. A. Pittman, and S. Knerr. Combining diverse systems for handwritten text line recognition. Machine Vision and Applications