

Handwritten Text Recognition using Machine Learning Techniques in Application of NLP



Polaiah Bojja , Naga Sai Satya Teja Velpuri, Gautham Kumar Pandala, S D Lalitha Rao Sharma
Polavarapu Pamula Raja Kumari

Abstract; Handwriting Detection is a technique or ability of a Computer to receive and interpret intelligible handwritten input from source such as paper documents, touch screen, photo graphs etc. Handwritten Text recognition is one of area pattern recognition. The purpose of pattern recognition is to categorizing or classification data or object of one of the classes or categories. Handwriting recognition is defined as the task of transforming a language represented in its spatial form of graphical marks into its symbolic representation. Each script has a set of icons, which are known as characters or letters, which have certain basic shapes. The goal of handwriting is to identify input characters or image correctly then analyzed to many automated process systems. This system will be applied to detect the writings of different format. The development of handwriting is more sophisticated, which is found various kinds of handwritten character such as digit, numeral, cursive script, symbols, and scripts including English and other languages. The automatic recognition of handwritten text can be extremely useful in many applications where it is necessary to process large volumes of handwritten data, such as recognition of addresses and postcodes on envelopes, interpretation of amounts on bank checks, document analysis, and verification of signatures. Therefore, computer is needed to be able to read document or data for ease of document processing.

Key Terms— E.G – For Example, NLP -Natural Language Processing, CNN – Convolutional Neural Network, OCR - Optical Character Recognition.

I. INTRODUCTION

The Handwritten text system is commonly used system in various applications, and it is a technology that is a mandatory need in this world as of now. Before the correct implementation of this technology we have dependent on writing texts with our own hands that result in errors. It's difficult to store, access physical data and process the data in efficient manner.

Revised Manuscript Received on December 30, 2019.

* Correspondence Author

Professor.Polaiah Bojja*, Department of Electronics and Communication Engineering, Koneru Lakshmaiah Education Foundation,

Polaiah Bojja, Department of Electronics and Communication Engineering, Koneru Lakshmaiah Education Foundation, Professor.

Naga Sai Satya Teja Velpuri, Department of Electronics and Communication Engineering, Koneru Lakshmaiah Educational Foundation.

Gautham Kumar Pandala, Department of Electronics and Communication Engineering, Koneru Lakshmaiah Educational Foundation.

S D Lalitha Rao Sharma Polavarapu, Department of Electronics and Communication Engineering, Koneru Lakshmaiah Educational Foundation.

© The Authors. Published by Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP). This is an [open access](https://creativecommons.org/licenses/by-nc-nd/4.0/) article under the CC-BY-NC-ND license <http://creativecommons.org/licenses/by-nc-nd/4.0/>

Manually it is needed to update, and labor is required in order to maintain proper organization of the data. Since for long time we have encountered a severe loss of data because of the traditional method of storing data.

Modern day technology is boon, and this technology is making people to store the data over machines, where the storage, organization and accessing of data is easier.

Adoption of the Handwritten Text Recognition software is a practical idea and, it is easier to store and access data that was traditionally stored. Furthermore, it provides more security to the data. One such example of Handwritten text Recognition software is the Google Lens and example for hardware is OCR scanners. The aim of our project is to make an model for handwritten text recognition and convert them into speech for application in healthcare and personal care that can recognize the handwriting using concepts of deep learning. We approached our problem using Tensor Flow and OpenCV as they contain Pre-trained Models that are directly used to provide results accurately compared to other methods over such task. The model that developed in this paper is used to convert the text in different forms ie, mainly Text document and Voice files. These files are stored in respective folder and information can be extracted. We mainly use Opensource models for the development of the project. The architecture of models that we have used are based on NLP. This NLP architecture has some basic components of data acquisition, processing and Query and visualization. This text is further converted into the voice

II. PROGRAMMING INTERFACE

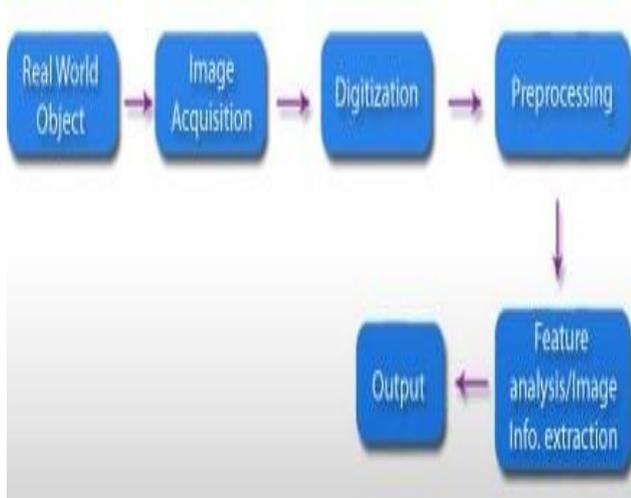
In the development of this project we have used PYTHON version 3 as language to develop the software in order to meet the project requirements. For the initial development and exploration, we have used SPYDER as an ide and then later we have shifted to PYCHARM another ide which supports Python and makes the task completion ease.

Various other libraries and pre-trained models we have used include Pytesseract, OS, Gtts (Google Text To Speech) etc.

These Programing algorithms have well defined models that are trained with some samples. We use this algorithms and models to increase the accuracy of the result. The result will be dependent on the amount of training given to the model to get better output.

III. PROPOSED METHODOLOGY

A. Block Diagram

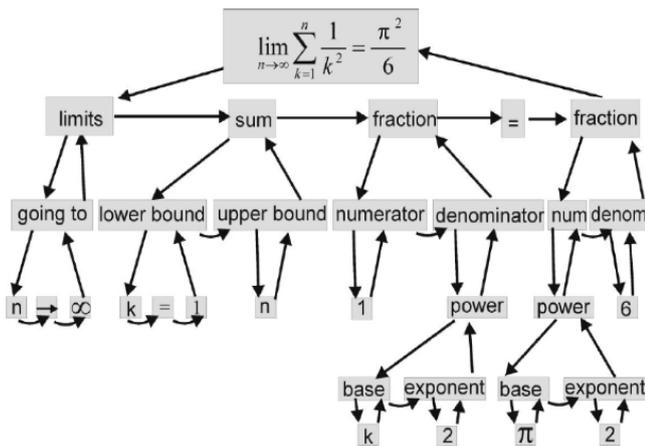


B. Abbreviations and Acronyms

Some of the abbreviations and acronyms used in this work are:

1. .py: Extension used for python scripts.
2. GTTS: Google Text To Speech.
3. NLP: Natural Language Processing.
4. OCR: Optical Character Recognition.
5. PIL: Python Imaging Library

C. Math behind recognition



IV. RESEARCH METHODOLOGY

- A. Recognition strategies heavily depends on the nature of the data to be recognized.
- B. In the cursive case, the problem is made complex by the fact that the writing is fundamentally ambiguous
- C. As the letters in the word are generally linked together, poorly written and may even be missing.
- D. On the contrary, hand printed word recognition is more related to printed word recognition, the individual letters composing the word being usually much easier to isolate and to identify.

The Algorithm that we have used in mainly based on the Pretrained Models, these models helps us to get the results required within less time.

These pretrained model of Tesseract and GTTS are used in this research. This research mainly concentrates on the working procedure of pretrained models and their application to solve problems. These Model intern uses Neural Network, Decision Trees to solve and segregate the Characters and words.

a. Decision Tree:

These trees flow the approach of trail and error method, solves every possible way and conclude the decision. These models improve their accuracy for every time they used to detect the outputs.

A typical example of decision tree:



Equations:

Gini Index :

$$I_G = 1 - \sum_{j=1}^c p_j^2$$

Entropy :

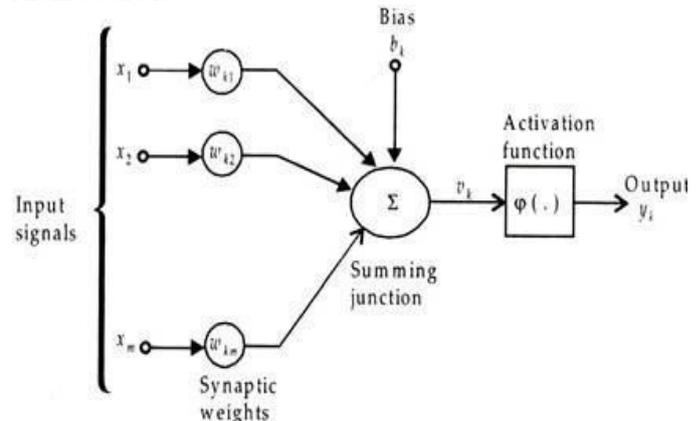
$$I_H = - \sum_{j=1}^c p_j \log_2(p_j)$$

Where p is the proportional of samples belongs to a class c, for a particular node. Based on these Character selection is done.

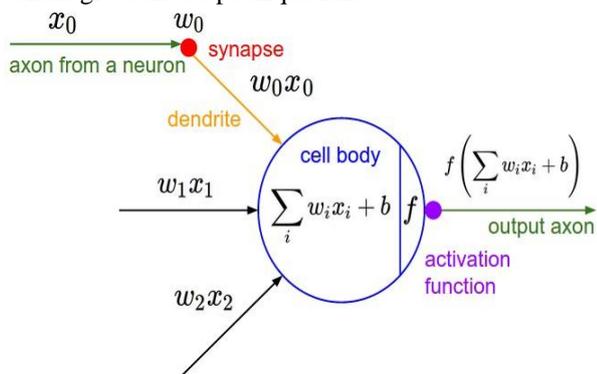
b. Neural Network Model

Neural Networks are artificial Networks that are working same the Human Neurons, but these carry information in form of weights. Different equations hold for the different situations or types to use the Threshold Equation

Architecture is:



These are the Weighted sum of the Neurons passing to threshold gives the output. Equations:



By using these Types of Formulae and we find the Threshold value and feature extraction matching. This is useful to conclude the handwritten text into characters of different types by taking the features of text and then the constructing the real time system.

The Characters recognise and word recognise were there and constructed them with help of neural network algorithm but the data is segregated or accepted by combination algorithms in pretrained models. These Time saving techniques are helpful for industrial in order to solve the problem in very efficient way.

Layer wise Dividing and comparing to database original characters and determine the identity of character and these a combination of the words and then convert file to text format and mp3 format.

V. RECOGNITION STRATEGIES

Recognition heavily depends upon Python Text to Speech In python, there a score of API's available to convert text to speech such as gTTS (Google Text to Speech API), eSpeakNG, pyttsx3 et. The gTTS convert the input text in to audio and saved as mp3 file for speaking purpose. It works only when the internet is connected and it has a very natural sound. The cross platform and offline TTS library is pyttsx3 which works for both python versions. It supports different speech engines like espeak, sapi5, nsss. The eSpeakNG is an open source, small size and multiple language speech synthesizer supporting python.

VI. FUNCTIONS

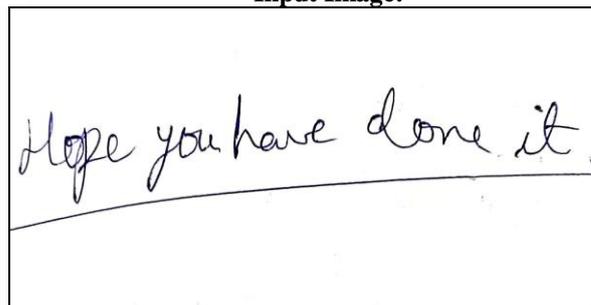
Some of the important functions used in this project are:

- **get_tesseract_version** Returns the Tesseract version installed in the system.
- **image_to_string** Returns the result of a Tesseract OCR run on the image to string
- **image_to_boxes** Returns result containing recognized characters and their box boundaries
- **image_to_data** Returns result containing box boundaries, confidences, and other information. Requires Tesseract 3.05+.
- **image_to_osd** Returns result containing information about orientation and script detection.
- **run_and_get_output** Returns the raw output from Tesseract OCR. Gives a bit more control over the parameters that are sent to tesseract.

VII. RESULTS

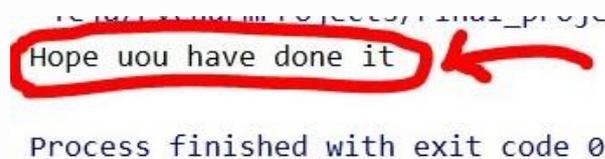
For the Given inputs in various forms we have developed the model for the research and the input output is noted as the.

Input Image:



These input images which we have produced has the 2 types of the outputs out of which the first one being output as computerized text based output and the other one being speech signal of the recognized text using the speakers of the computer or the attached ones to the system.

Output-1



This is the Text Format Output

Output-2:



This is second output as a speech signal form

VII. .DISCUSSION AND ANALYSIS

In the input image out of 17 characters, our model has recognized and processed 17 characters out of which 16 were correctly recognized and 1 was faulty. Hence with an accuracy of 94% for this input and an overall accuracy of 92.7%. These Results are accurate and hypothetical results that have good accuracy and clear voice file which helps us to hear after the Subject.

VIII. CONCLUSION

Handwritten character recognition is a complex problem, which is not easily solvable. The necessity is around dataset and database. This model is built to analyze the text we have written and convert it in Computer Text and Voice formats. This application is applicable in many sectors of healthcare and Consumer sector. This type of model used in health application can save understanding perspectives of people and store each and every record digitally.

ACKNOWLEDGMENT

We thank our research organization for supporting us during the course of this work. We also thank our guide Dr. Bojja Polaiiah for his remarkable guidance and support throughout.

We also acknowledge our parents for their financial and moral support throughout the period of this research. Finally we like to acknowledge our efforts in accompanying each other at work.

REFERENCES

1. G. Cohen, S. Afshar, J. Tapson, and A. van Schaik, "Emnist: an extension of mnist to handwritten letters," arXiv preprint arXiv:1702.05373, 2017.
2. L. Eikvil, "Optical character recognition," citeseer. ist. psu. edu/142042. html, 1993.
3. Singh, Sameer, Mark Hewitt, "Cursive Digit And Character Recognition on Cedar Database", Pattern Recognition, 2000.Proceedings. 15th international conference on. Vol. 2. IEEE 2000.
4. Anita Pal and Davashankar Singh, "Handwritten English Character Recognition Using Neural Network",
5. International Journal of Computer Science and Communication, pp: 141- 144, 2011.
6. 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
7. Polaiiah Bojja, etal; "Evaluation and Development of Advance Control Strategies Which Uses Pattern Recognition Technique for Nonlinear System", International Journal of Computer Theory and Engineering, Vol.3, No.1, February, 2011, pp: 1793-8201.
8. V. SAMPATH KUMAR REDDY, M. SRIMURALI and B. POLAIAH; "Prediction Of Pm10 And So2 Concentrations In Ambient Air Using Artificial Neural Networks For Hyderabad", International Journal of Civil Engineering (IJCE), Vol. 5, Issue 3, 2278-9995 Apr - May 2016; pp: 1-10.
9. Kumari, P.R., Polaiiah Bojja, "SO2 concentrations in the ambient air at vijayawada station using artificial neural networks computing", International Journal of Recent Technology and Engineering, 2019
10. Vijaya Lakshmi A., Gopitilak V., Muzammil Parvez M., Subhani S.K., Ghali V.S., "Artificial neural networks based quantitative evaluation of subsurface anomalies in quadratic frequency modulated thermal wave imaging", Infrared Physics and Technology, ISSN: 13504495, Vol No:97, 2019, pp:108-115, DOI: 10.1016/j.infrared.2018.12.013
11. Aparna P., Kishore P.V.V., "Biometric-based efficient medical image watermarking in E-healthcare application", IET Image Processing, ISSN:17519659, Vol No:13, Issue No:3, 2019, pp:421 - 428, DOI: 10.1049/iet-ipr.2018.5288
12. Gajula S., Rajesh V., "Multiple medical image watermarking scheme with DWT, SVD transforms", Journal of International Pharmaceutical Research, ISSN: 16740440, Vol No:46, 2019, pp: 274 – 277.
13. Kumar D.A., Sastry A.S.C.S., Kishore P.V.V., Kumar E.K., Kumar M.T.K., "S3DRGF:Spatial 3-D relational geometric features for 3-D sign language representation and recognition", IEEE Signal Processing Letters, ISSN:10709908, Vol No:26, Issue No:1, 2019, pp: 169 - 173, DOI: 10.1109/LSP.2018.2883864
14. Karunakaran S., Gattim N.K., "VLSI implementation of folded FIR filter structures using high speed multipliers", Journal of Engineering and Applied Sciences, ISSN: 1816949X, Vol No: 14, Issue No:4, 2019, pp:1070 -1077, DOI: 10.3923/jeasci.2019.1070.1077
15. More A.R., Prasad M.S.G., Wankhede Vishal A., "Optimized throughput and minimized energy consumption through clustering techniques for cognitive radio network", International Journal of Innovative Technology and Exploring Engineering, ISSN: 22783075, Vol No:8, Issue No:4S2, 2019, pp:474 - 479.
16. Ramesh C., Venkat Rao D., Murthy K.S.N., "Short note on the application of compressive sensing in image restoration", Smart

AUTHOR PROFILE



First Author Professor. Polaiiah Bojja, Department of Electronics and Communication Engineering, Koneru Lakshmaiah Education Foundation,



Second Author Naga Sai Satya Teja Velpuri, Department of Electronics and Communication Engineering, Koneru Lakshmaiah Educational Foundation.



Third Author Gautham Kumar Pandala, Department of Electronics and Communication Engineering, Koneru Lakshmaiah Educational Foundation.



Fourth Author S D Lalitha Rao Sharma Polavarapu, Department of Electronics and Communication Engineering, Koneru Lakshmaiah Educational Foundation.



Fifth Author Pamula Raja Kumari, Department of Mathematics, Koneru Lakshmaiah Educational Foundation.