ISSN No. 0976-5697

VOLUME 12 SPECIAL ISSUE 2, JUNE 2021

International Journal of Advanced Research in Computer Science

RESEARCH PAPER

Available Online at www.ijarcs.info

Recognizing Handwritten Characters Using OCR & Converting into TTS

Yerrabolu Sailendra Chakravarthy Reddy
School of C&IT Reva University
Bangalore, India
r15cs449@cit.reva.edu.in

Rohit Singh
School of C & IT Reva University
Bangalore, India
r14cs274@cit.reva.edu.in

Manju More E
School of C & IT Reva University
Bangalore, India
manjumore.e@reva.edu.in

Abstract: The aim of the project is "To make Neural Networks aware of handwritten characters", that is, to create a platform that converts handwriting into digital text using Neural Network & Optical Character Recognition. This paper provides an in-depth study of text acquisition, tracking and image recognition with three major contributions. First, it is proposed that a standard framework for the release of image text that equally describes the discovery, tracking, recognition and their relationships and interactions. Second, within this framework, the various methods, systems, and procedures for visualizing the text of an image are summarized, compared, analyzed and the extracted text is converted and extracted by voice. Thirdly, related applications, outstanding challenges, and future directions for image editing are also well discussed.

Keywords: Handwritten Characters, Neural Network, Optical Character Recognition, OCR, TTS, OpenCV

I. INTRODUCTION *(HEADING 1)*

The aim of the project is to "Train Neural Network to Identify Handwritten Character". According to this Artificial Intelligence is one of the fastest growing branches of computer science in which the neural network plays its part. By using this program, one could translate handwritten English words, numbers and special characters into digital text and save them. Neural Network is a branch of artificial intelligence inspired from the human system which is very complex and very fast. Using this concept, neurons are replaced by nodes and dendroids. Neural Network is unencrypted and unstructured but training. The more he trains, the stronger he becomes. Neural Network helps identify patterns of characters and extracts them by pointing. The project also uses image processing techniques that help to improve the image and convert it into a gray scale and then converted into a binary image of recognition. Before using the neural network, it must be trained using the MNIST database of all others because it has a large amount of data. With the training of this data, neural networks try to understand a given image and visualize it clearly. In our project, we wanted the device to be able to retrieve text from any complex background and read it carefully. Promoted the method used by applications such as "Cam Scanner", assuming that in any complex background, text will be boxed e.g. we think this is a required region that contains text. This is done using warping and cutting. A new image is found on the edge and a border is drawn above the letters. This gives it more meaning. The image is then processed by Optical Character Recognition and Text to Speech to provide audio output.

According to the references we have gathered Aisha Sharaf [8] aims at improving the existing handwritten text recognition using machine learning. Machine learning to say is an interesting yet little complicated branch of the computer science. In this paper they used convolution neural network. The neural network is capable of handling complicated data with ease. In general, the data which is handwritten text image is passed through different hidden layers between input and output layer forming a complex network which is correctly connected to the last layer that is output layer. But when coming to image recognition we could use convolution networks instead of general network in a view of computers. Convolution neural networks well in compared to general neural network. Because each pixel increasing in image, increases the parameters exponentially and J. Pradeep [9] mostly discuss about the method of the neural network recognition. The method is multi-layered Feed Forward method. This is a simple neural network method, which is unidirectional that follows a forward direction from input to output. This paper also discusses about a new method for feature extraction that is diagonal feature extraction method. In this paper they used fifty datasets written 26 English alphabet characters. These English alphabets are almost written in 570 different styles.

After the project is implemented, it can be used in many other sectors like banking, financing & digitalizing the hand written documents. We can be able to preserve the old documents even in this digital era.

II. LITERATURE REVIEW

Yingying Zhu [1] suggested that the document, as one of mankind's greatest inventions, played an important role in human life, dating back to antiquity. The rich and accurate information embedded in the text is very useful for many types

of applications that support the idea, so text acquisition and recognition in natural forums has become important and effective research topics in computer recognition and document analysis. Especially in recent years, the public has seen an explosion of research efforts and great progress in these fields, although various challenges (e.g. noise, blurring, distortion, prevention and diversity) remain. The objectives of this study are threefold: 1) to introduce timely tasks, 2) to identify high-level skills, and 3) to predict potential future research indicators. In addition, this paper provides comprehensive links to publicly available services, including benchmark data sets, source codes, and online demons. In summary, this review of the literature can serve as a valuable reference for investigators in the field of event record and identity.

Taraggy M. Ghanim [6] proposed a multidisciplinary approach to working in the field of offline Arabic letter recognition. His paper discusses the Hierarchical Agglomerative Clustering (HAC) process of dividing the database into slightly related clusters. The relationship between the support of the built-in database representatives as a great example of a search engine and helps to find the reduced difficulty of comparing each test image with the collection.

Aisha Sharaf [8] aims at improving the existing handwritten text recognition using machine learning. Machine learning to say is an interesting yet little complicated branch of the computer science. In this paper they used convolution neural network. The neural network is capable of handling complicated data with ease. In general, the data which is handwritten text image is passed through different hidden layers between input and output layer forming a complex network which is correctly connected to the last layer that is output layer. But when coming to image recognition we could use convolution networks instead of general network in a view of computers. Convolution neural networks well in compared to general neural network. Because each pixel increasing in image increases the parameters exponentially. So, this paper uses convolution neural network.

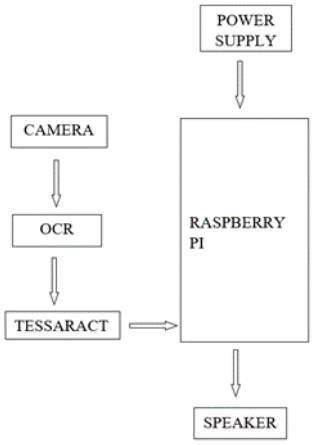
J. Pradeep [9] mostly discuss about the method of the neural network recognition. The method is multilayered Feed Forward method. This is a simple neural network method, which is unidirectional that follows a forward direction from input to output. This paper also discusses about a new method for feature extraction that is diagonal feature extraction method. In this paper they used fifty datasets written 26 English alphabet characters. These English alphabets are almost written in 570 different styles.

Megha Agarwal [10] discusses about feature extraction method like diagonal and direction techniques in the way that it generates high accuracy results. This paper proposes the horizontal and vertical methods of feature extraction in a traditional manner. This paper also proposes In neural network the model called as feed forward is trained using the back-propagation algorithm.

III. OBJECTIVES

The aim of this project is to create a model that will be able to identify and determine handwritten characters from its image using Neural Network & Optical Character Recognition concepts.

IV. METHODOLOGY



SYSTEM DIAGRAM

Visual impairment is one of the major limitations in humanity, especially in this age when information is transmitted more by text than by voice. The device we have developed is intended to help people with visual impairments. In this project, we have developed a device that converts image text into speech. A basic framework is an embedded system that captures an image, removes only the region of interest (eg image region containing text) and converts that text into speech. Made with Raspberry Pi 3B + and Raspberry Pi camera. Captured image makes a series of pre-processing steps to find only that part of the image that contains the text and remove the background. Two tools are used to convert a new image (containing text only) into speech. They are OCR (Optical Character Recognition) engines and TTS (Text-to-Speech) engines. Sound output is heard with a raspberry pi audio jack using speakers or earphones.

The device contains Raspberry Pi, speaker or earphones, Raspberry pi camera, power supply (230V AC) and power switch mode (SMPS). The camera must be manually displayed in this text and the image must be taken. This image is then processed by the Raspberry Pi and the audio output is audible.

A. OCR ENGINE

Image transcription is done using visual acuity (OCR) recognition. OCR is a field of research in pattern recognition, artificial intelligence and computer vision. It is the conversion of images to typed, handwritten or printed text into digital text or computer format text. Previous versions of the OCR had to be trained for each character in a particular text and font.

Today, there are advanced OCRs with a high degree of accuracy, supporting a wide variety of images, languages and fonts. For our project, we have used Tesseract OCR. It is the most accurate open source engine and is powered by google. Can be used on Linux platforms, mac and windows. A new version of Tesseract, 3.4 supports 100 languages. However, images should form multiple pre-processing categories such as audio output, rating etc.

B. TTS

The process of converting text into computer speech is called speech synthesis. Speech transcript (TTS) is used to perform speech comparisons. The TTS is made up of two parts: the rear end and the rear end. The previous ending converts text into symbols, for example, a number. Each punctuation mark is given a phonetic. Background and convert phonetics to sound. For our project, we have used Festival TTS. The festival is the

TTS of the most widely used open source. It has a variety of words and supports English, Spanish and English. We used the English language.

IMPLEMENTATION

Raspbian is a free, Debian-based app, used on Raspberry Pi hardware. Our code is written in Python language and the functions are called OpenCV. To perform OCR and TTS activities we include Tesseract OCR and Festival software.

The Raspberry pi 3B + is a device that contains several key functions on a single chip. It is a system on chip (SoC). The Raspberry pi camera is 5MP and has a resolution of 2592x1944. The Raspberry Pi has a 3.5mm audio hole so earphones or a speaker can be easily connected to it to hear sound.

Raspberry Pi 3B+

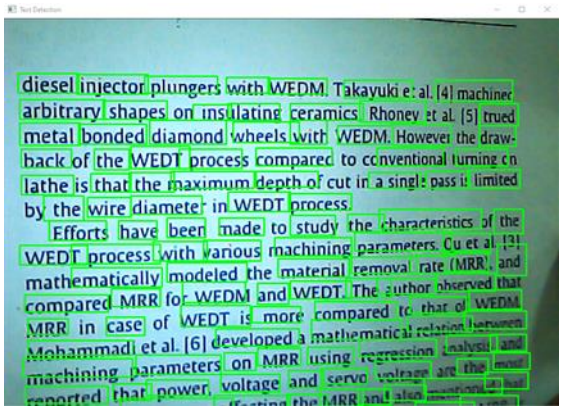
Raspberry Pi Camera

A. IMAGE ACQUISITION

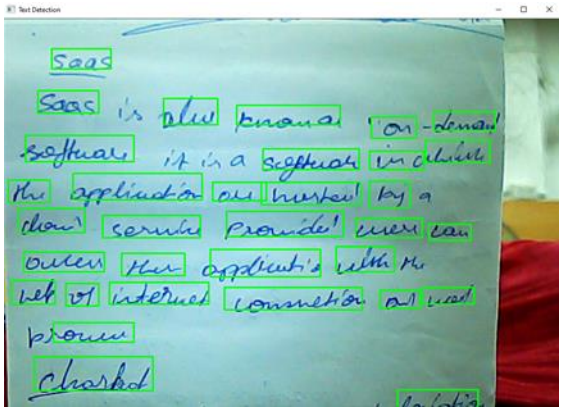
At this point, the Raspberry Pi camera captures text images. The quality of the photo taken depends on the camera used.

B. IMAGE PRE-PROCESSING

In this step the image is converted to a gray scale as most OpenCV functions require an input parameter as a gray scale image. Audio deletion is done using a two-state filter. The discovery of the canvas of Canny is made on a gray scale image to better determine the boundaries. Folding and cutting of the image is done according to the parameters. This enables us to locate and extract only the region that contains the text and to remove the unwanted domain. Finally, Thresholding is done to make the image look like scanned text. This is done to allow the OCR to properly convert the image into text.



Text detection using OCR on a text document



Text detection using OCR on a hand written document

C. IMAGE TO TEXT CONVERSION

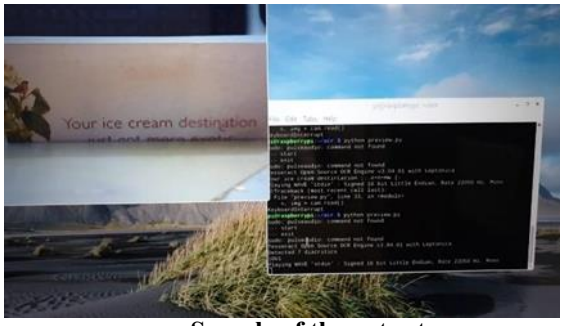
In this form, It converts the pre-processed image, which is in .png form, to a .txt file. We are using the Tesseract OCR.

D. TEXT TO SPEECH CONVERSION

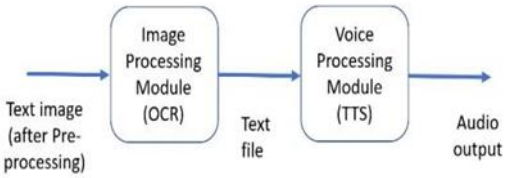
The second block is a voice processing module. Converts .txt file to audio output. Here, the text is translated into speech using a speech synthesizer called Festival TTS. The Raspberry Pi has a jack on the sound board, the sound of the board is made with PWM output.

V. RESULT

This system makes the visually impaired feel less stressed when it comes to reading unwritten text in braille. The image processing component allows the extraction of the required text region from a complex domain and provides good quality input to the OCR. The text, which is an OCR output, is sent to the TTS engine that produces the speech effect. To allow the portability of the device, the battery can be used to power the system. Future work can improve devices that detect object and extract text from videos instead of still images.



Sample of the output



Flow diagram of the project

VI. CONCLUSION

This project has still a long way to develop in the coming years. If this project is correctly implemented, then it can be used for blind people in reading out the texts or books and it could be helpful in many other sectors such as banking, financing and digitalizing the handwritten documents.

VII. FURTHER WORK

The use of commercial frames is something that needs to be tested, as it has produced better results than free frames when used without filtering, suggesting that the use of these frames in filters to improve visibility will lead to more satisfying results. The database used for this project can be upgraded, by requiring a wider set and representing the use of real-world scenarios. Other tests that can be done to check system quality include comparing the visual audio line with the text itself, asking users to write down what they think they have heard and finally make a comparison between the text and the original text. Another issue that should be addressed in development is real-time recognition.

At this stage, the project was designed and developed for the English language, but in the future, it should allow use in other languages, thus extending the number of people who can benefit from the advantages of the application. It is also reserved for future work the increment of the menu options, allowing the user to change other application options such as the language or the setting of sounds and orientations

VIII. REFERENCES

[1] Y. Zhu, C. Yao, and X. Bai, "Scene text detection and recognition: Recent advances and future trends," Frontiers in Computer Science, vol. 10, no. 1, pp. 19-36,2019.

[2] Oussama Zayene, Jean Hennebert, Sameh Masmoudi Touj, Rolf Ingold, and Najoua Essoukri Ben Amara. "A dataset for Arabic text detection, tracking and recognition in news videos-AcTiV." In 2015 13th International Conference on Document Analysis and Recognition (ICDAR), pp. 996-1000. IEEE, 2015.

[3] Ruxandra Tapu, Bogdan Mocanu, and Titus Zaharia. "DEEP-SEE: Joint object detection, tracking and recognition with application to visually impaired navigational assistance." Sensors 17, no. 11 (2017): 2473.

[4] Stanislav Mukhametshin, Alisa Makhmutova, and Igor Anikin. "Sensor tag detection, tracking and recognition for AR application." In 2019 International Conference on Industrial Engineering, Applications and Manufacturing (ICIEAM), pp. 1-5. IEEE, 2019.

[5] Xu-Cheng Yin, Ze-Yu Zuo, Shu Tian, and Cheng-Lin Liu. "Text detection, tracking and recognition in video: a comprehensive survey." IEEE Transactions on Image Processing 25, no. 6 (2016): 2752-2773.

[6] Taraggy M. Ghanim, Mahmoud I. Khalil & Hazem M. Abbas. Comparative Study on Deep Convolution Neural Networks DCNN-Based Offline Arabic Handwriting Recognition. IEEE, 10.1109/ACCESS.2020.2994290.

[7] Ahmed Mahdi Obaid., IIHazem M. El Bakry., IIIM.A. Eldosuky., IVA.I. Shehab. Handwritten Text Recognition System Based on Neural Network. IJARCST, 4(1), 2347 - 8446, 2016.

[8] Aisha Sharaf1., Bhagya Viswanath2., Kavya Chandran3., Nishana Salim4., Anju S Oommen5. Handwritten Text Recognition and Digitization System. IJIRSET, 8(5), 2319-8753, 2019.

[9] J.Pradeep1 ., E.Srinivasan2., S.Himavathi3. Diagonal based Feature Extraction for Handwritten Alphabets Recognition System using Neural Network. IJCSIT, 3(1), 2011.

[10] Megha Agarwal., Shalika., Vinam Tomar., Priyanka Gupta. Handwritten Character Recognition using Neural Network and Tensor Flow. IJITEE, 8(6S4), 2278-3075, 2019.

[11]Ashwinkumar.U.M and Dr. Anandakumar K.R, "Predicting Early Detection of cardiac and Diabetes symptoms using Data mining techniques", International conference on computer Design and Engineering, vol.49, 2012

[12] Tariq Rashid "Make Your Own Neural Network: A Gentle Journey Through the Mathematics of Neural Networks, and Making Your Own Using the Python Computer Language" CreateSpace Independent Publishing Platform, 2016

