iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 11 **Issue:** IV **Month of publication:** April 2023

DOI: https://doi.org/10.22214/ijraset.2023.50917

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 11 Issue IV Apr 2023- Available at www.ijraset.com

Implementing Image-to-Speech Recognition by Capturing Image Frames for Visually Impaired

Prof. Jayant V. Kulkarni[1], Vedant Bhoir[2], Prathamesh Deshpande[3], Sakshi Deshpande[4]
[1]*faculty & guide,* [2, 3, 4]*Student, Department of Instrumentation and Control Engineering, Vishwakarma Institute of Technology, Bibwewadi Pune*

Abstract: ***The difficulties that the visually impaired people go through is often been not considered. They have difficulty reading written materials on a daily basis such as newspaper or posters or many other things used frequently. Their everyday life is very difficult as communication is totally based on speech and text .Based on the literature survey, the research work until now is limited to the accessibility of using software also, the products are not cost effective. Hence, the main objective of the paper is to make a software that is easily accessible and cost effective. This paper is a prototype software that helps the visually impaired people as well as the people with low vision to read and understand the text from the image they are scanning. The prototype is based on the OCR tool and a google text to speech function. The software is tested on numerous photos after completion.The software correctly identified the text. This will make easier for them to read more efficiently.***
Keywords: Optical Character Recognition, Raspberry Pi, Raspberry Pi Camera, Tesseract , Visually Impaired .

I. INTRODUCTION

Visual impairment is one of the biggest limitations for humanity, especially in these days information is communicated a lot by text messages rather than voice. Our aim is to help people with visual impairment. In this project we develop a device that converts image text to speech.

The main research work includes the OCR applications in various flied**.** .Dr. Yusuf Parvej,The paper presents introduction, major research work and applications of Optical Character Recognition in various fields [1].The main drawbacks of the existing system is limitation to the languages. The proposed system is a prototype software that can be used to detect any language and output the scanned image in speech.

Recently, Miss. Tejashree U. Bagayatkar proposed a system that contents 4 modules they are: camera module, image processing module, optical character recognition module and text-to-speech module. The device output is in the form voice so, it can be easily hear by visually impaired people [5]. The main disadvantage is that the device is not cost effective. Hence, main of aim of this paper is to build a prototype that will help visually aid and low visioned people to read efficiently without any help. Also to make it cost effective so as it is available easily.

II. RELATED WORK

Dr. Yusuf Parvej,[6]The paper presents introduction, major research work and applications of Optical Character Recognition in various fields. At the first introduction of OCR will be discussed and then some points will be stressed on the major research works that have made a great impact in character recognition. And finally the most important applications of OCR will be covered and then concluded.

Karez Hamad,[7]Optical character recognition is an active research area that attempts to develop a computer system with the ability to extract and process text from images automatically. The objective of OCR is to achieve modification or conversion of any form of text or text-containing documents such as handwritten text, printed or scanned text images, into an editable digital format for deeper and further processing. Therefore, OCR enables a machine to automatically recognize text in such documents

Saharsa Vanria Thota [8] Visual impairment or vision loss is defined as the decreased ability to see clearly and cannot be fixed using glasses. Blindness is the term used for complete vision loss. The common causes of vision loss are uncorrected refractive errors, cataracts and glaucoma. People with visual impairment face a number of difficulties in normal daily activities like walking, driving and reading.

Bhalaji Natarajan[9] This project is implemented using a handheld page or document scanner, an external Bluetooth module when the scanner does not have an inbuilt Bluetooth module, an Android application to perform OCR and speech synthesis and an Android mobile phone.

©IJRASET: All Rights are Reserved

The cost involved in developing the system is significantly low and the system provides a friendly user interface for the visually impaired people.

K Karthick [10] In this paper, the author has discussed handwritten OCR systems classification and the steps involved in OCR which is one of the automatic identification techniques and also given information regarding recent applications used.

G.A.E.Satish Kumar [11] In this paper the author presented a computer vision technology to extract text from scene images and electronic aid is used to convert text to speech. Design and Development of Smart Assistive Device for Visually Impaired People presented in May 2021, 2016. In this paper the author introduced a smart stick with an RFID technology and IR sensors to give direction and to detect obstacles respectively.

Jamshed Menon, [12] The offline system is a static system in which input data is in the form of scanned images while in online systems the nature of input is more dynamic and is based on the movement of the pen tip having a certain velocity, projection angle, position and locus point. Therefore, an online system is considered more complex and advanced0, as it resolves the overlapping problem of input data that is present in the offline system.

III. METHODOLOGY/EXPERIMENTAL

Algorithm:

1) Step 1: Installation :

The most precise free OCR engine is Tesseract. In order to identify the text-containing area, it first converts the user's image to a grayscale image. It enters all the text from the.png file and transforms it to a.txt file

2) Step 2 :Image Capturing Process:

It gained access to the system's camera by using OpenCV and other tools. The image is saved as a file as soon as it is obtained. This document is a.png file.

3) Step 3 : Text to Speech :

The **text** from the text image is added using Pytesseract to the.txt file that is generated. The text file is then accessible using gTTTs, a feature that converts text to speech.

4) Step 4: Audio File :

It stores that file as an mp3 file. The operating system then processes that file, producing sounds of text being read from an image.

IV. FLOWCHART

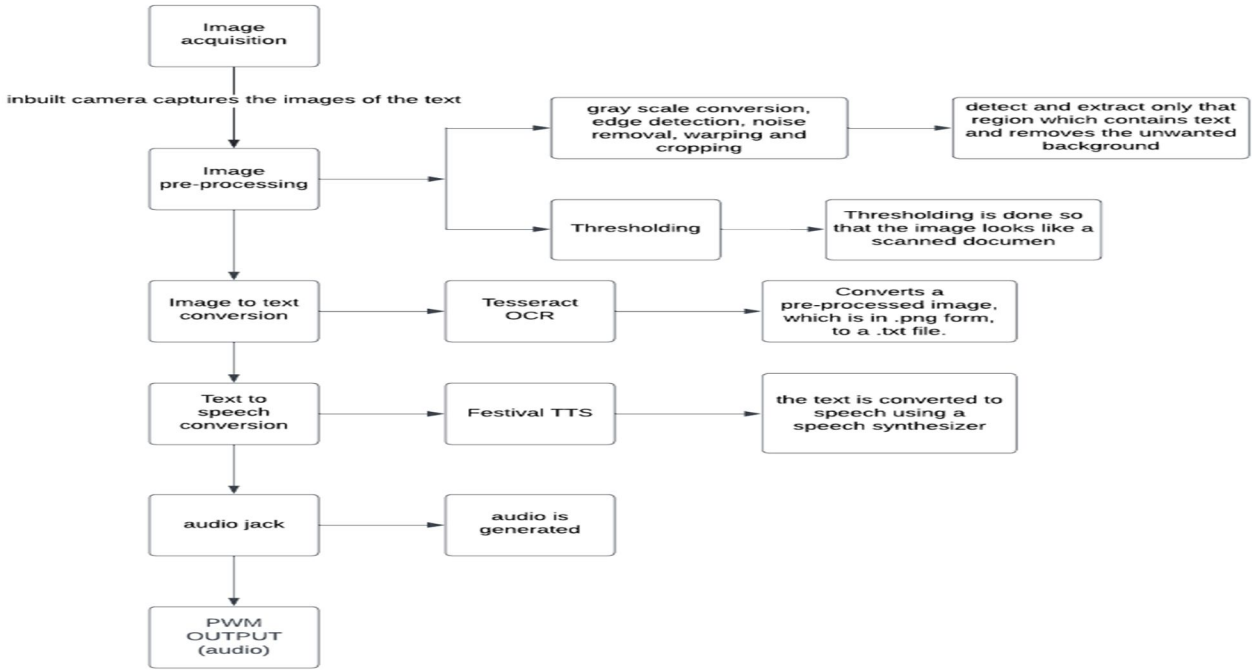


Fig: 1 Flow chart

1) We take a picture using the laptop's camera that includes text and characters.
2) After taking the picture, we prepare it for image pre-processing.
3) We also examine settings like cropping, noise removal, edge detection, grayscale conversion, and threshold.
4) It recognises and isolates the text-containing area of an image, turning it into a scanned document.
5) We installed the tesseract to convert images to text.
6) The best free OCR engine is this one. In order to identify the text-containing area, it first converts the user's image to a grayscale image. The.png file is changed into a.txt file.
7) The image is saved as a file as soon as it is obtained.
8) The text was then read from that file and turned to a picture using Pytesseract.
9) The text file is then retrieved via gTTTs, a feature that turns text into speech.
10) Once we obtain the audio file, we will be able to play it back using the laptop's audio jack.

V. RESULTS

This system has the accuracy to understand both words and numbers. After placing a particular image in front of the camera, OpenCV library had successful captured the image and pytesseract library converted the texts. We easily got the speech. After training the model to read numbers, the accuracy of the system increased, and hence the coding part of the system is completed .After completion the software is tested on various images. The total number of images count up to 500.The software was able to detect the text accurately.

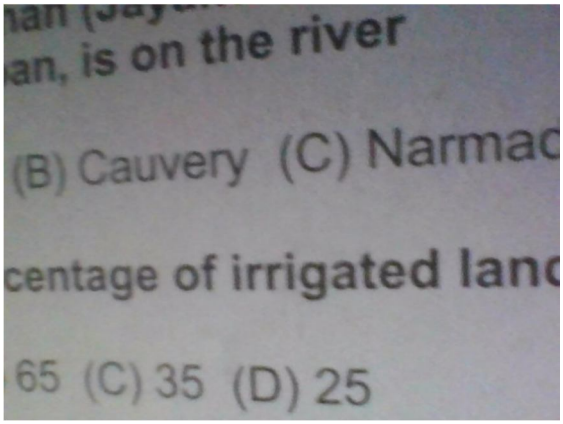


Fig: 2 Image captured

The above image is the screenshot of the text which has appeared on the console screen along with a speech reciting the above content.

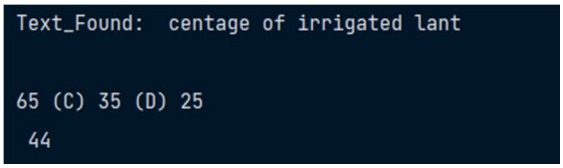


Fig: 3 The captured image is the record of image captured into a text.

We easily got the speech. After training the model to read numbers, the accuracy of the system increased, and hence the coding part of the system is completed.
Text file converted to audio file shows the text file produced by the tesseract software and it produced audio file.

VI. CONCLUSION

Visually handicapped people can easily listen to the document's text thanks to the installation of this system. One can change the text to the desired language using translation software. He or she can translate the text into another language and then again into audio by using the Google speech recognition tool. We can also widen the field of view for long-range capturing. work on creating an effective portable capture device could also be emphasized.

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

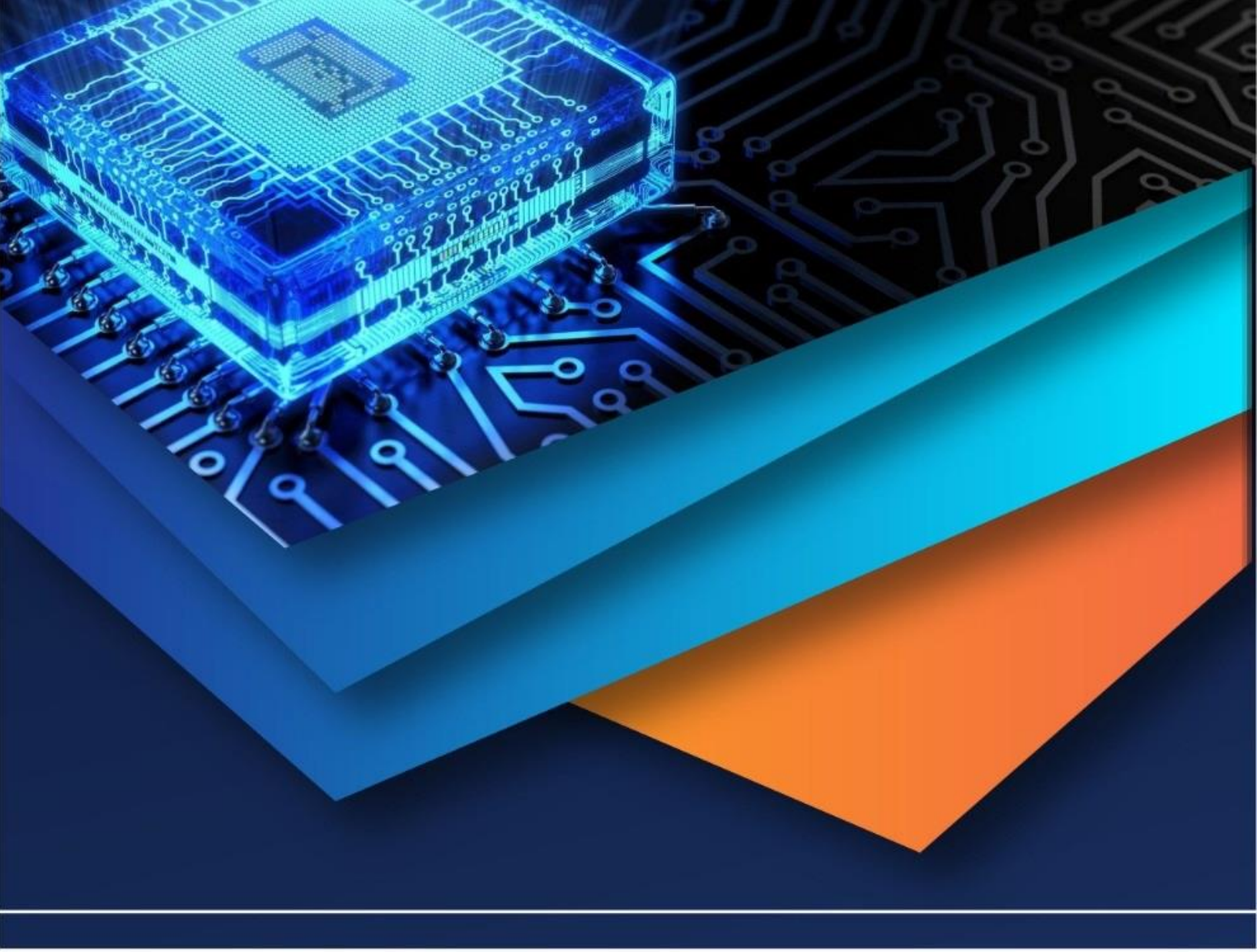
VII. ACKNOWLEDGEMENT

It is our great pleasure to acknowledge sense of gratitude to all, who have made it possible for us to complete this EDI project with success. It gives us great pleasure to express our deep gratitude to our guide Jayant Kulkarni for their support and help from time to time during project work. It is our pleasure to acknowledge the sense of gratitude to Prof. Rajesh Jalnekar our director for their great support and encouragement in project work. Finally we would like to thank all faculty members and all our colleagues for their valuable suggestions and support.

REFERENCES

[1] Prof. Vaibhav V. Mainkar,et al:Raspberry pi based Intelligent Reader for Visually Impaired Persons.IEEE Xplore,June 2020
[2] Dr. I S Akila,et al:A Text Reader for the Visually Impaired using Raspberry Pi.IEEE Xplore,2018.
[3] Mr.Rajesh M.,et al:TEXT RECOGNITION AND FACE DETECTION AID FOR VISUALLY IMPAIRED PERSON USING
[4] RASPBERRY PI.2017 International Conference on circuits Power and Computing Technologies [ICCPCT],2017.
[5] Rithika.H,et al:Image Text To Speech Conversion In The Desired Language By Translating With Raspberry Pi.IEEE Xplore,2016.
[6] http://elinux.org/RPi_Text_to_Speech_(Speech_Synthesis)
[7] https://en.wikipedia.org/wiki/Visual_impairment
[8] https://en.wikipedia.org/wiki/Braille
[9] https://www.classycyborgs.org/braille-literacy-statistics-india/ www.raspberrypi.org
[10] http://www.zdnet.com/article/raspberry-pi-11-reasons-why-its-the-perfect-small-server/
[11] http://aishack.in/tutorials/opencv/
[12] http://opencv-python-tutroals.readthedocs.io/en/latest/py_tutorials/py_setup/py_intro/py_i
[13] ntro.html http://hackaday.com/2016/02/28/introducing-the-raspberry-pi-3/

©IJRASET: All Rights are Reserved



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)