iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 9 **Issue:** VIII **Month of publication:** August 2021

DOI: https://doi.org/10.22214/ijraset.2021.37262

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: 7.429
Volume 9 Issue VIII Aug 2021- Available at www.ijraset.com

Raspberry Pi based Object Detection and Text Reader using Voice Assistance

Sushmitha M[1],Geetha Nayana B[2], G. Sree Varshitha[3], Bhoomika HR[4], Keerthi KS[5]

[1]*Assistant Professor, Department of Computer Science and Engineering, S. J. C. Institute of Technology,Chikkaballapur, India*
[2, 3, 4, 5]*Student, Department of Computer Science and Engineering, S. J. C. Institute of Technology,Chikkaballapur, India*

Abstract: Communication is the basic requirement for humans to connect and it requires text and speech but visually impaired people cannot able to perform this. This project helps them to read the image. This project is an automatic document reader for visually impaired people, developed on the Raspberry Pi processor board. It controls the peripherals like a camera, a speaker which acts as an interface between the system and the user. Here, we use a raspberry pi camera which is used to capture the image and scan the image using Image Magick software. Then the output of the scanned image is given to OCR(optical character recognition) software to convert the image to text. It converts the typed or printed text to the machine code. Then we use Text to Speech (TTS), which is used to convert speech to text. The experimental result is very helpful to blind people as there was much analysis of the different objects.

Keywords: Raspberry Pi camera, raspberry pi, image magick, optical character recognition (OCR), text to speech (TTS) engine.

I. INTRODUCTION

Human Life without visualization is hard to imagine and is one of the most feared conditions. For those who are visually impaired need access to information in a format which can be accessible to visually impaired persons like braille, audio etc.,. Only few persons can learn and understand braille and others can't learn this. For those people they need other technology to communicate with people. Because communication plays an important role in human life. To overcome this problem for the visually impaired persons we have to design and implement a text to speech conversion and object detection in which theoutput is in the format of voice. So, the visually impaired persons can easily understand and communicate because it is in the format of sound.

The Tensor module: A module to manipulate symbolic objects. Also, the pi camera uses Python's Tensor for scanning the image of the object. The main idea is to create a way in which they can identify the objects in front of them without a hazel. Python is easy to work with and easy to use with Raspberry. It lets us convert a project into a real-time application but how it works is whenever there is any obstacle or an object in front of them, the Raspberry Pi helps them detect it .How exactly would be the big question. Initially, the raspberry pi cam captures the object in front of it. Scans the image using the OCR program, converts it into text which gets displayed on the screen. Meanwhile for them to hear what the camera see's, the text gets converted into speech using theTTS program. Finally, they hear it through the speaker, hence by warning them of the danger ahead or simply identifying a certain object. All this is done by using Python and its amazing libraries. Henceforth the raspberry pi combined with the power of python, has the capability to be their eyes, in case of emergencies .This is a portable device and it does not require an internet connection and can be used independently by people with low vision. It also has a user interface that allows them to interact with the device easily.

II. LITERATURE SURVEY

Smart reader for visually impaired

The blind people use braille language for reading the text, but those are limited to some since not every textbook contains this bailey language and it is difficult to construct so we use this technique of ocr in which each character is being recognized and it will be said in speakers wherein the visually impaired persons can hear it. They can read any text document that will be present in the paper. In this paper, when the image is sent it is savedusing MIL software which works in both RGB and gray and supports all the image files. Later on, the unwanted noise Should be removed by pre-processing, and background subtraction is done for the image. Then the image is converted RGB to gray. The filtering process is done to improve visual quality then it is converted gray to binary. There will be segmentation done by using MATLAB which is used for matching algorithms of OSTU and camy algorithms. Here we also use the Graphic Interface design that is being used which is used to capture an image from real-time and at one more place we have the picture of a database that is used to compare the images. The smart reader uses a simple efficient conversion

©IJRASET: All Rights are Reserved

which is used to convert image to text.[1]

Voice Assisted Text Reading System for Visually ImpairedPersons Using TTS Method Smart specs

The system proposed is Smart spec with an inbuilt camera for blind people to help them read any printed text in voice form. The main motive of this system is to allow visually impaired persons to touch the printed text and receive voice output.

This system is designed using technologies like Tesseract OCRfor text extraction and TTS to convert text to speech. Whatever The output is recognized that it will be viewed through Matlabsimulator with an audio output but the disadvantage was that nouser-friendly products are developed with high performance andaccuracy [2].

There are about 1.3 billion people in the world who have visual impairment issues. They would have to read printed material using Braille. However, there are limitations for these people when the material is not printed in Braille.

Thus, this paper proposes an easy way which is designed for thevisually impaired person. The raspberry pi captures the image ofthe object with a pi camera. The captured image is then converted to text by using image-to-text conversion in the Optical Character Recognition (OCR) framework. Finally, the text will be read out into speech format using text-to-speech conversion in the Text to Speech (TTS) framework. As a result, a person who has visual impairment can understand it by listening unlike feeling the printed text. It is a user-friendly device since the system has a sound guideline so they can always get to know the process of the application[3].

Text recognition and face detection aid for visually impairedpeople using raspberry pi. Speech and text is the main mode for communication. Any human would need the vision ability toaccess the information in the text. This paper proposes camera Based assistance to read the text so that visually impaired personscan read the text present on the captured image without a hazel.

The proposed idea involves text extraction from the scanned image using the tesseract optical character recognition(OCR) and converting the text to speech through speaker. It is carried out usingraspberry pi [4].

Efficient portable camera-based text to speech converter for a blind person.

A portable camera-based assistive text reader is a device designed that is helpful for a blind person. The proposed systeminvolves four modules like identification of the object, positioning of text, identification of text, and text to voice conversion. The system works efficiently with OCR. Convolutional a recurrent neural network is a proposed method used to train individual words separately. The disadvantage being that the proposed method CRNN failed to deliver output ifone word is above and another word is below in the image [5].

III. PROPOSED APPROACH

At first, image will be captured using Raspberry pi camera butthis image will not be available for the object detection so it will be sent to image processing module wherein it will be converted to jpg format. In the next step, objects are detected using JSON library where only objects with the high accuracy can be detectedrest all the images are being sent to image processing where in unwanted noise is removed. This will be removed using Imagemagick software. It is used for image sharpening and text cleaning. Sharpening increases contrast between bright and dark regions. Text cleaning helps the image to be readable for OCR. Then it converts into text using OCR software. Now it is in .txt file. Then it converts into text file using TTS to speech and the output is converted into sound. After all this, it will be hearableto visually impaired persons as an output.

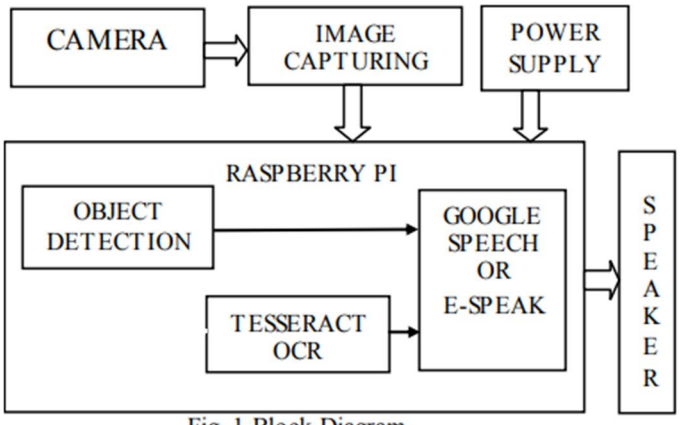


Fig. 1 Block Diagram

©IJRASET: All Rights are Reserved

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.429
Volume 9 Issue VIII Aug 2021- Available at www.ijraset.com

IV. IMPLEMENTATION AND RESULTS

1) Image capturing: in this step the image of the object is captured using the pi camera , then the resolution and framework is calculated .
2) Object detection: the ROI(region of interest) is detected using the supervised pre trained coco model which already hasabout 90 different collections of objects.
3) Text extraction: the const_config , a module in tesseractOCR converts the array of images into string and den initializes it to a text variable .
4) Text to speech conversion: the text is given to the videolan player which runs the audio and hence the text gets converted to speech .
5) The speech output will be heard from the speaker.

But how all This happens is a question that arises. All dis is done on the basis of supervised learning where no mathematical computation is required, only the input is to be given .
Firstly,

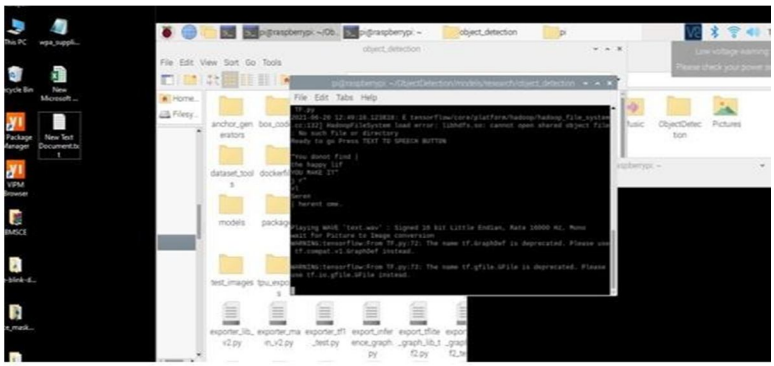
a) The labeling of the image takes place using supervisedlearning of the model.
b) Then the mapping of labels to images so that computationcan be done using image segmentation.

A. Hardware Implementation

The hardware used in this system is composed of raspberry p*i* camera and raspberry pi module. The 5- megapixel raspberrypi camera is used for capturing the image. For raspberry pi module we used Python language and Raspbian OS. The 8 GB SD card is used for storing the OS, code of the proposed system and images. Mobile hot spot and LAN Ethernet are used for internet connection which is taken to the raspberry pi module. A power bank is used to give input power supply of 5V to the raspberry pi module. The headset and speaker are used to hear the voice or audio output.



Mobile Phone



System output

©IJRASET: All Rights are Reserved

V. FUTURE WORK

The output of the device is voice that will be helpful for blindpeople and it will be helpful for them to identify the objects.It can also be help in blind schools and colleges, it also uses AI technique which helps further. It also less in size so it can be carried anywhere.

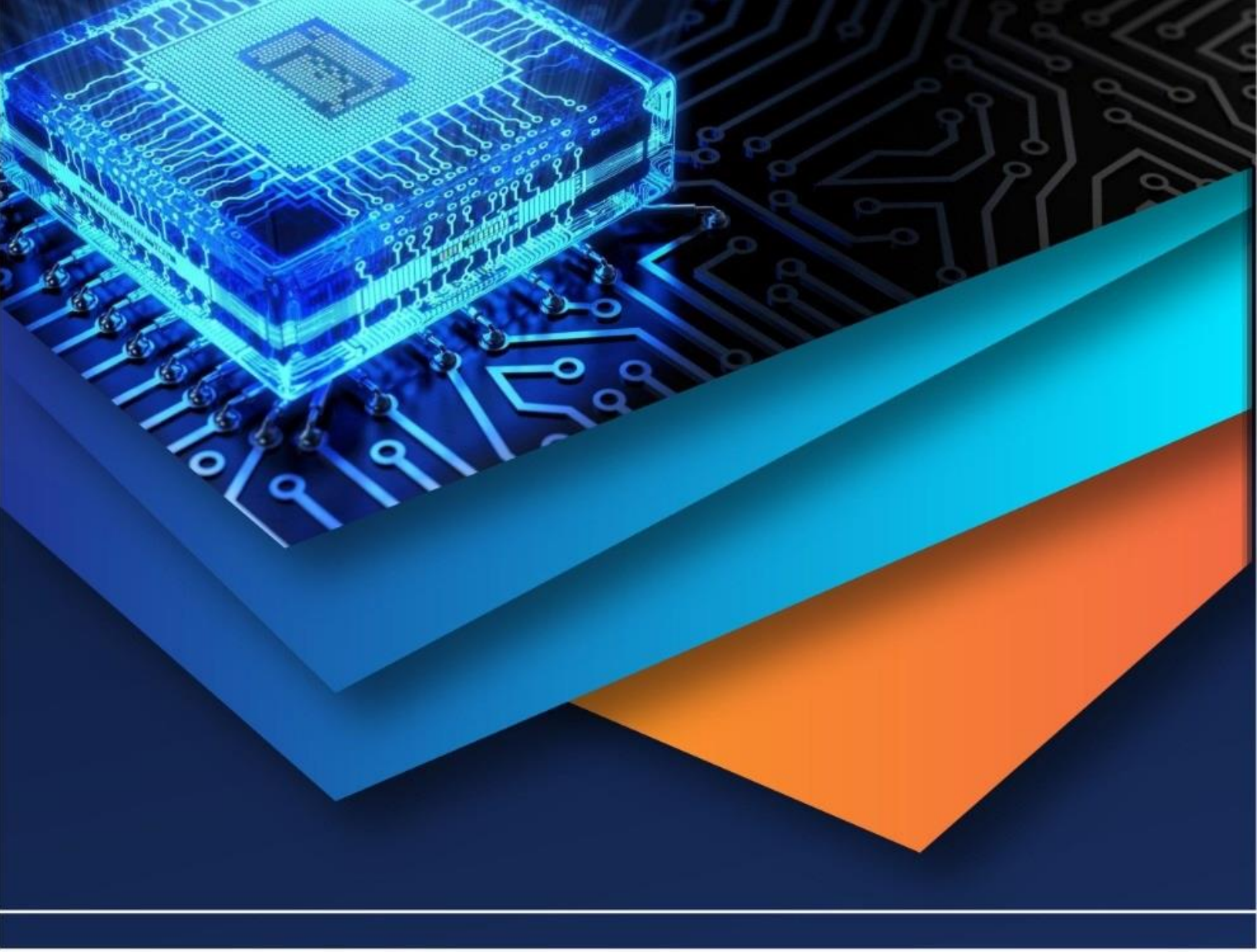
VI. CONCLUSION

We have implemented a voice assisted text reading system using raspberry pi. The experimental results are verified successfully and the output of hardware is also verified using different captured images. Our methodology processes the captured image and reads it out clearly. In simple words this canbe used as a powerful tool for the betterment of the people who are differently abled.

REFERENCES

[1] Musale, Sandeep & Ghiye, Vikram. (2018). Smart reader for visually impaired. 339-342. 10.1109/ICISC.2018.8399091..
[2] Monisha, M. & Nandhini, A.. (2015). Portable computer vision based assisting device for the visually impaired people. International Journal of Applied Engineering Research. 10. 14379- 14387.
[3] Liambas, Christos & Saratzidis, Miltiadis. (2016). Autonomous OCR dictating system for blind people. 172-179. 10.1109/GHTC.2016.7857276.
[4] Ali, Maghfirah & Tang, Tong Boon. (2016). Smart Glasses for theVisually Impaired People. 9759. 579-582 10.1007/978-3-319- 41267-2_82.
[5] Shah, Trupti & Parshionikar, Sangeeta. (2019). Efficient Portable Camera Based Text to Speech Converter for Blind Person. 353-358. 10.1109/ISS1.2019.8907995.
[6] A. Goel, A. Sehrawat, A. Patil, P. Chougule, and S. Khatavkar, “Raspberry pi-based reader for blind people,” 2018.

©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)