IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.301
Volume 3, Issue 5, May 2023

Text to Speech Conversion using Raspberry PI

S. Sreenivasulu [1], P. Satya Priyanka [2], P. Jaahnavi[3], S. Harshini[4]
Asst. Professor, G. Narayanamma Institute of Technology and Science, Hyderabad[1]
Student, Electronics and Communication Engineering
G. Narayanamma Institute of Technology and Science, Hyderabad[2], [3],[4]

Abstract: *In present situation, communication plays a vital part in the world. Transferring on information, to the correct person, and in the proper manner is veritably important on a particular and professional position. The world is moving towards digitization, so are the means of communication are Phone calls, emails, textbook dispatches etc. have come a major part of communication vehicle in this digital world. In order to serve the purpose of effective communication between two parties without any detention, numerous operations have come to actuality, which acts as a middleman and help in effectively carrying dispatches in from textbook to the speech signals over long hauls of networks. The main purpose of this design is to overcome the problems facing by the eyeless people and illiterates. Because the eyeless people and illiterates can be fluently manipulated, this leads to abuse. To overcome this problem we're proposing a device which helps in conversion of hard dupe of textbook which is fitted into the device will be converted to speech. utmost of these operations find the use of functions similar as articulators, conversion from textbook to synthetic speech signals, language restatement amongst colorful others. In this design, we'll be executing different ways and algorithms that are applied to achieve the conception of Text to Speech(TTS).*

Keywords: Tesseract, Text to speech (TTS), Optimal Character Recognition(OCR), Raspberrry Pi.

I. INTRODUCTION

From the once many times, Mobile Phones have come a main source of communication for this digitalized society. One can make calls and textbook dispatches from a source to a destination fluently. It's known that verbal communication is the most applicable modem of passing on and conceiving the correct information. This design is substantially designed for eyeless people which offers computer generated voice affair to engage with original or remote services. In this design, textbook to speech conversion process using Optical Character Recognition is shown. Optical Character Recognition can be used extensively in healthcare operations to prop eyeless people. This design can be used to help eyeless people to read books, understand different textbook on banners, film land and large announcement boards. OCR includes substantially three factors the camera to capture the images, the programmable system to convert the captured camera into whatever format we want and eventually the affair system to show the affair of OCR. Raspberry pi, a compact size and low weight system for creating an OCR system is used in this design. The capturing of image is done through raspberry pi camera or through android camera which can be connected with Ethernet, Bluetooth or Wi- Fi. The webcam or raspberry pi camera prisoner the images large fountain size image and store it into raspberry pi system. By using tesseract OCR library handed by python to convert the captured image to the textbook, the display of Raspberry pi is shown on the computer display with the help of putty software and Ethernet string. The camera is connected to the raspberry pi system which works as an OCR tool for this project. The camera generally takes 5- 7 twinkles to capture and reuse the images. With the help of tesseract OCR library of python, the captured image is converted to the textbook format in the raspberry pi hard fragment position. The converted textbook is handed to TTS system which converts the textbook to the voice format.

DOI: 10.48175/IJARSCT-10059

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.301
Volume 3, Issue 5, May 2023

II. BLOCK DIAGRAM

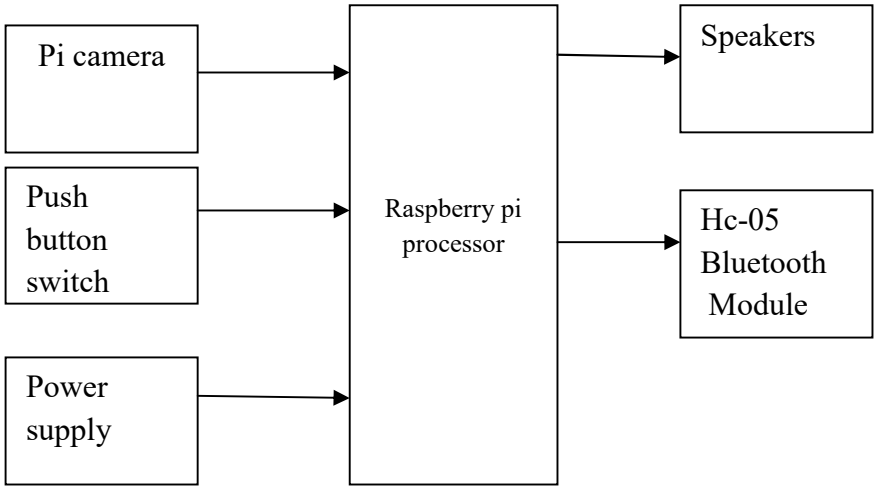

Fig.1. Block Diagram of proposed system

The above figure shows the connection between the different blocks involved in the proposed system. The image of textbook on a published paper is captured with pi camera using drive button switch. The captured image is converted to the textbook and saved at same position of image. It takes roughly 7- 8 sec to convert the textbook. The converted textbook is reused in TTS. The speech is attained as an affair in Headphones or speaker. It recognizes figures as well as letter in English. The range of reading distance is 30 cm.

III. HARDWARE REQUIREMENTS

- Raspberry Pi Processor 3: It is a series of single board computers. It is equipped with a quad-core 64-bit Broadcom BCM2837 ARM Cortex-A53 SoC processor running at 1.2 Ghz.
- HC-05 Bluetooth Module: This module is used to connect small bias like mobile phones using a short- range wireless connection to change lines.
- Power bank: 5V power force is used to give power force to the raspberry pi processor.
- Push button switch: It is used to capture the image of a paper placed in front of the pi camera.
- Speakers: Speakers offer audio affair of the image to the listeners.

IV. SOFTWARE REQUIREMENTS

Raspbian OS

Raspbian OS is a free operating system, grounded on Debian, optimized for the Raspberry Pi tackle. Raspbian Jessie is used as the interpretation is RPi's main operating system in our design. Our law is written in Python language(interpretation2.7.13) and the functions are called from OpenCV. OpenCV, which stands for Open Source Computer Vision, is a library of functions that are used for real- time operations like image processing, and numerous others(14). As of now, OpenCV supports plenty programming languages like C, Python, Java etc. and can be operated on different platforms including Windows, Linux, OS X etc. The interpretation used for our design is opencv-3.0.0. OpenCV's operation areas include Facial recognition system, Gesture recognition, mortal – computer commerce(HCI), Mobile robotics, stir understanding, Object identification, Segmentation and recognition, stir shadowing, stoked reality and numerous further. For OCR and TTS operations to be performed ,Tesseract OCR must be installed Tesseract is known as an opensource Optical Character Recognition(OCR) Engine. It can be used directly, or for programmers it can be used as an API to rize compartmented, handwritten or published textbook from images. It supports a large number of languages. It is generally called‘ tesseract ’ or ‘ tesseract- ocr ’.

Image Processing

Books and papers have letters. Our end is to prize these letters and convert them into digital form and also recite it consequently. Image processing is used to gain the letters. Image processing is principally a set of functions that's used

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.301
Volume 3, Issue 5, May 2023

upon an image format to conclude some information from it. The input is an image while the affair can be an image or set of parameters attained from the image. Once the image is loaded, the loaded image can be converted into gray scale image. The image which is given as output now is in the form of pixels within a specified range. This specified range is used to determine the letters in the image. In gray scale, the image will be having either white or black content; the white will be the distance between words or blank space.

Feature Extraction

In this stage we gather the essential features of the image called point charts. One similar system is to descry the edges in the image, as they will contain the needed textbook. For this we can use colorful axes detecting ways like Sobel, Kirsch, Canny, Prewitt etc. The most accurate in chancing the four directional axes vertical, perpendicular, right slant and left slant is the Kirsch sensor. This fashion uses the eight point neighborhood of each pixel.

Optical Character Recognition

Optical character recognition, generally shortened to OCR, is the mechanical or electronic conversion of scrutinized images of handwritten, typewritten or published textbook into machine decoded textbook. It's extensively used as a form of data entry from some kind of original paper data source, whether documents, deals bills, correspondence, or any number of published records. It's pivotal to the robotization of published textbooks so that they can be electronically searched, stored more curtly, displayed on- line and used in machine processes similar as machine restatement, text to speech and textbook mining. OCR is known as field of exploration in pattern recognition, artificial intelligence and computer vision.

Tesseract

Tesseract is a free software optical character recognition machine for colorful operating systems. Tesseract is considered as one of the most accurate free software OCR machines presently available. It can be installed on Linux, Windows and Mac OS. An image with the textbook is given as input to the Tesseract machine that's command grounded tool. Also it's reused by Tesseract command. Tesseract command takes two arguments First argument is image train name that contains textbook and alternate argument is affair textbook train in which, uprooted textbook is stored. The affair train extension is given as. txt by Tesseract, so no need to specify the train extension while specifying the affair train name as a alternate argument in Tesseract command. After processing is completed, the content of the affair is present in .txt train. In simple images with or without color(gray scale), Tesseract provides results with 100 delicacy. But in the case of some complex images Tesseract provides better delicacy results if the images are in the gray scale mode as compared to color images. Although Tesseract is command- grounded tool but as it's open source and it's available in the form of Dynamic Link Library, it can be fluently made available in plates mode**.**

Setting up the Raspberry Pi and camera interface:

To get started, we will need a raspberry Pi board formerly loaded with Raspbian OS and a raspberry Pi camera board module.

Step 1. Connect the camera module with the Raspberry Pi shown on the computer window.

Step 2. Enable the camera module which is shown on the window

Open the terminal and run “sudo raspi-config”.pi@raspberry:~$ sudo raspi-config

This will open a window as shown in fig.

Click on the the “Interfacing Options” on the window and select “Camera”.

The camera module is enabled.

Now the camera interfacing is done.

To check this, open the terminal and execute “raspistill -o image.jpg”. Camera module is activated by this, displays a preview of the image, and then captures the image and saves it in the current working directory as “image.jpg”.

pi@raspberry:~$ mkdir test

pi@raspberry:~$ cd test/

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.301
Volume 3, Issue 5, May 2023

```
pi@raspberry:~/test$ raspistill -o image.jpg
pi@raspberry:~/test$ ls
image.jpg
```

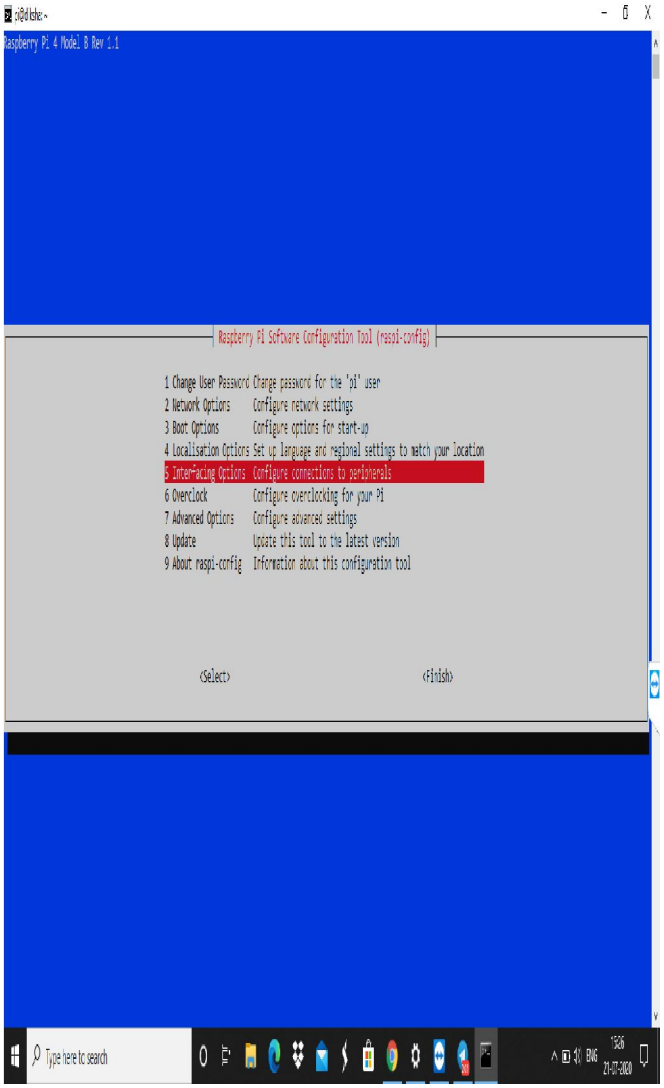

Fig 2(a) Software configuration

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.301
Volume 3, Issue 5, May 2023

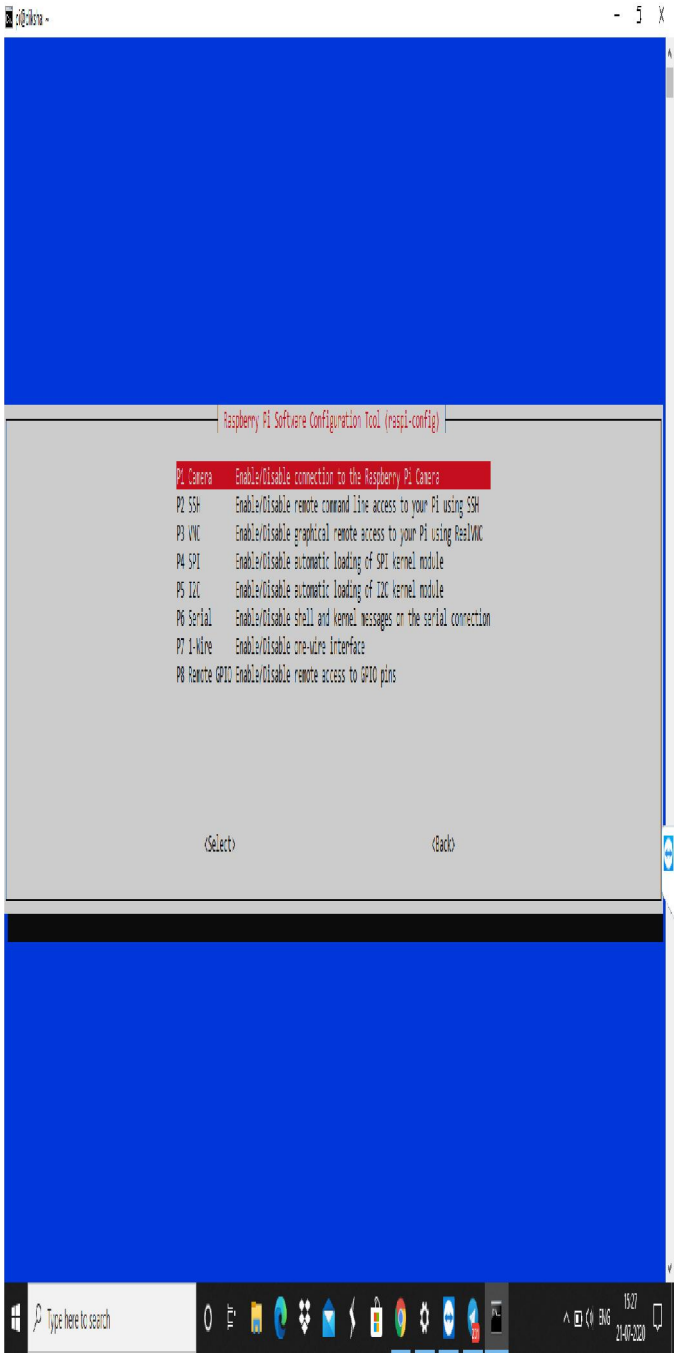

Fig 2(b) Software configuration

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.301
Volume 3, Issue 5, May 2023

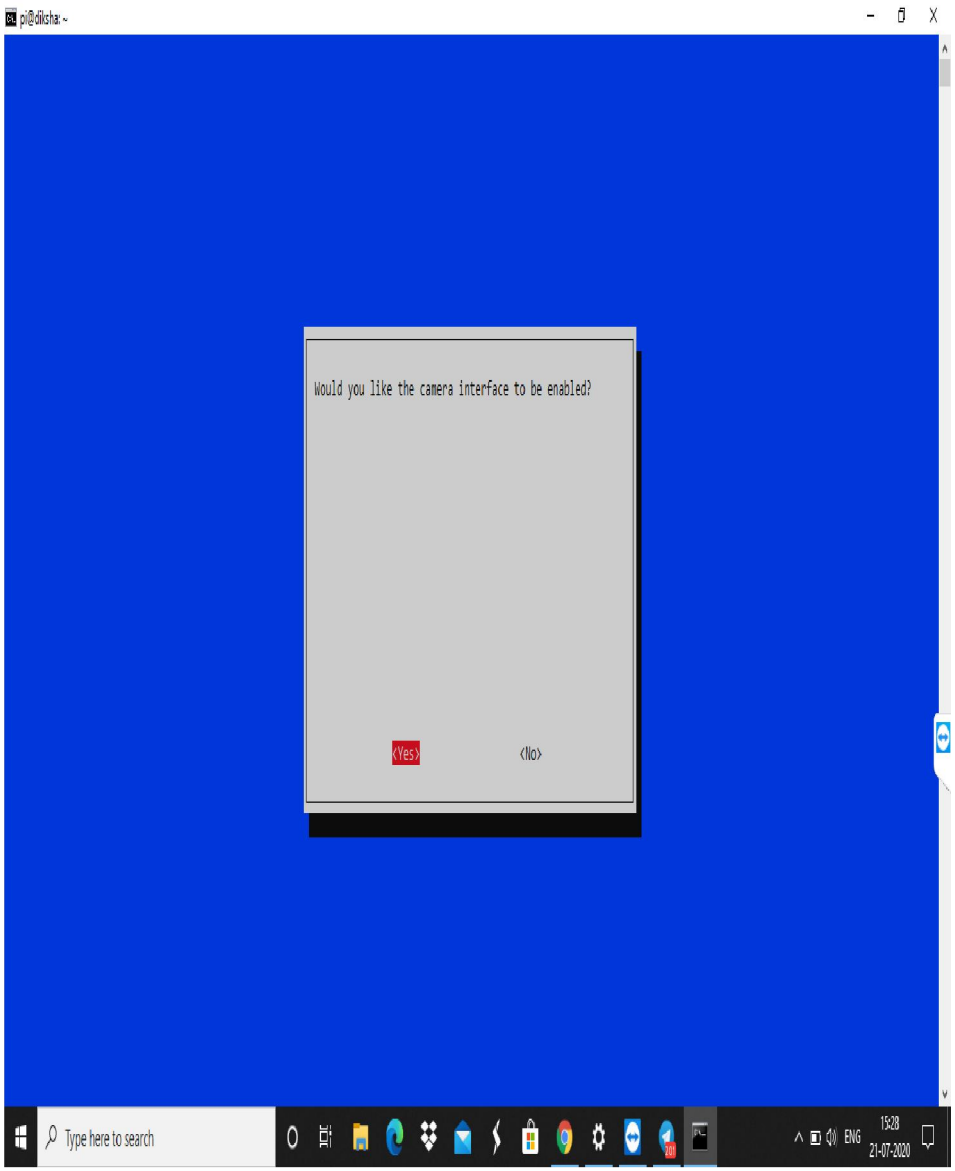

Fig 2(c) Software configuration

V. RESULTS

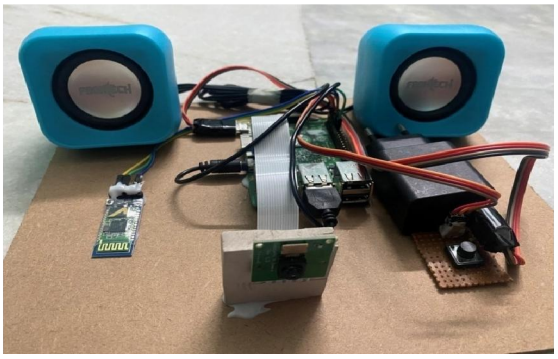

Fig 5 Proposed system prototype.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.301
Volume 3, Issue 5, May 2023

The portable system which captures the text on the paper which is placed in front of the pi Camera. These details were verified using Raspberry Pi Processor for extraction of text from image. The Raspberry Pi processor will sending the same text to the user mobile through HC-05 Bluetooth module in order to verify and monitor text wireless.

For this project tesseract-ocr is to be installed which is an optical character recognition engine so that pytesseract can use this to extract text from the image.The captured image is converted to the text and saved at same location of image. The converted text is processed in TTS. The speech is obtained as an output in the speakers.

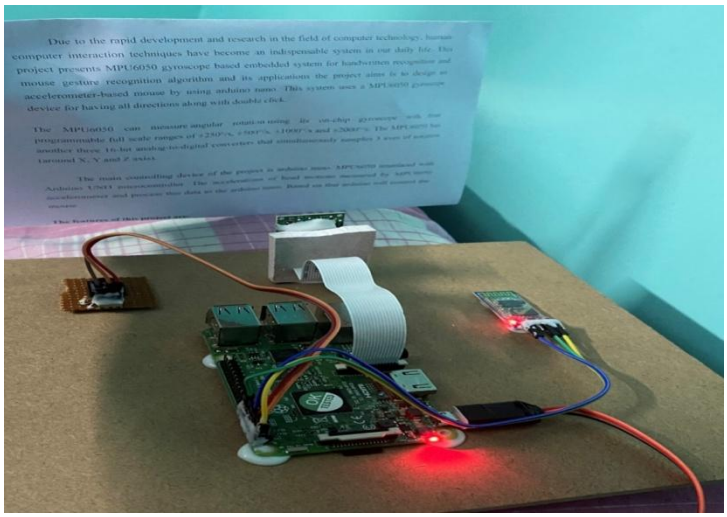

Fig 4 System capturing the text

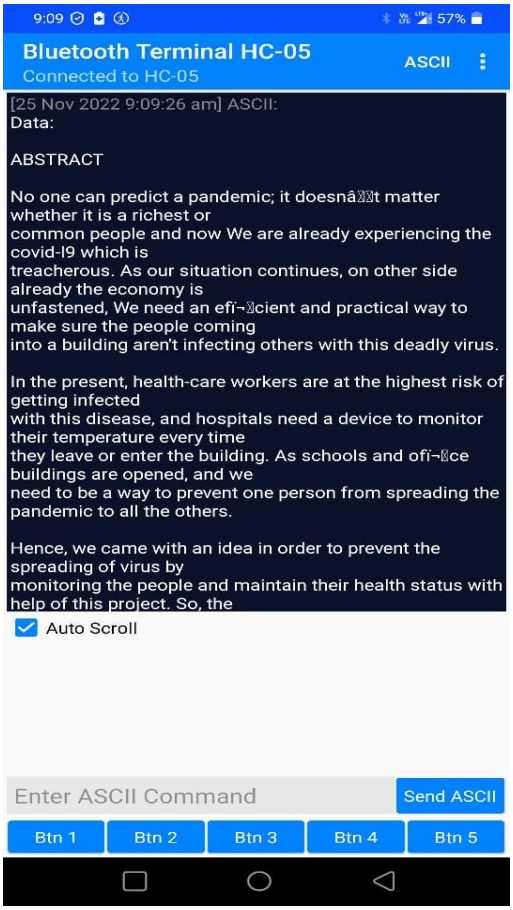

Fig 5 Bluetooth module output

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.301
Volume 3, Issue 5, May 2023

VI. CONCLUSION

The project was designed as a system which captures the text from printed paper which is placed in front of pi camera and display on Bluetooth mobile application to check whether the captured text is correct or not .This system is able to read the text using pytesseract. Speech is obtained as output.

VII. FUTURE SCOPE

1. By connecting GSM module this system can send the captured image of object details to the authorities in the form of SMS messages to the predefined numbers.
2. By connecting GPS module this system can track the location of the blind person.

VIII. ACKNOWLEGMENTS

We're thankful to tutoring and non-teaching staff members of ECE Department, GNITS for the precious information handed by them in their separate fields. We're thankful for their cooperation during colorful stages of the period of our assignment. Numerous thanks to our parents who have been a great source of strength all through this work and to our musketeers who support was veritably precious in completion of the work.

REFERENCES

[1]. J.Liang,D.DeMenthon,andD.Doermann,“Geometricrectificationof camera-captured document images,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 30, no. 4, pp. 591–605, 2008.
[2]. J. Banerjee, A. M. Namboodiri, and C. V. Jawahar, “Contextual restoration of severely degraded document images,” in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., 2009, pp. 517–524
[3]. S. Lu and C. L. Tan, “Retrieval of machine-printed Latin documents through Word Shape Coding,” Pattern Recognit., vol. 41, no. 5, pp. 1799–1809, May 2008.