# HELLA- A SMART READING BOT

L. LATHA, V. GEETHANI, M. DIVYADHARSHINI
Computer Science and Engineering, Kumaraguru College of Technology
Saravanampatty, Coimbatore
INDIA
latha.l.cse@kct.ac.in

*Abstract:-* The main objective of this project is to present a smart reading bot for visually impaired people. According to World Health Organization (WHO), 81% are visually impaired who live in developing countries. Nowadays Human communication is mainly focused on text and speech. To read the text a human needs a vision. The existing method is still incomplete to meet their requirements, so current technology is implemented in the project to perform their day to day activities irrespective of their impairments. In order for the humans to be updated with the current trend, a smart reading bot is implemented. This system consists of a camera interfaced with Raspberry Pi for processing the text in the form of image capturing. The camera is used to capture the image of both handwritten and printed text. The raspberry pi makes use of python programming and libraries to perform image to text conversion and text to speech conversion. The bot not only reads the printed or handwritten text and gives the output in audio form by using Text to Speech synthesis. Also search for the word given by the user in the form of voice is recognized and particular word is searched in the line of text. Additionally, it helps to find the meaning of the specific word pronounced by the user, both searching and finding the meaning is done through speech recognition methodology, and all these outputs are converted into audio output (Speech). Searching the word gives the output as whether the word is present or not. The main advantage of this project is the captured image and audio will not get stored in the Raspberry Pi it will get directly uploaded in AWS server called S3 bucket which in returns give the public URL for accessing the image and audio for the use of text to speech conversion, image to text and all other features done by the Raspberry Pi. The future technology aims at providing a portable setup with battery backup which can be used anywhere and at any time. The application of this proposed system mainly focuses in the areas where the information and notices are to be read like in auditoriums, libraries and offices. The assistant is applicable for visually impaired people as well as for normal people in order to increase their level of comfort.

*KeyWords:* - Text to speech, Image to Text, Image Recognition, Raspberry Pi, Speech output, Speech synthesis, S3 bucket, AWS

## 1 Introduction

According to the World Health organization (WHO) survey, 285 million people are estimated to be visually impaired worldwide among which 90% live in developing countries. and forty-five million blind individuals world-wide. Though there are many existing solutions to the problem of assisting individuals who are blind to read, however none of them provide a reading experience that in any way parallels that of the sighted population. Our project aims in creating a Smart reader, smart way of finding the particular word and finding the meaning of the word using raspberry pi. Where the smart reader will be very useful for visually impaired people and also useful for normal categorized people.

Human communication today is mainly via speech and text. To access information in a text, a person needs to have vision. However, those who are deprived of vision can gather information using their hearing capability. Book reader is a camera based assistive text reading to help blind person in reading the text present on the text labels, printed notes and products. It involves Text Extraction from image and converting the text to speech converter, a process which makes blind persons to read the text.

## 2 Literature Review

[1] The text is extracted from the image, text to speech conversion is done for the extracted text, a process which makes blind persons to read the text. This is carried out by using Raspberry pi, where portability is the main aim which allows the user to

carry the device anywhere and can use any time. [2] The Image to Text conversion and then Text to Speech conversion is done. The OCR algorithm involves various stages like Scanning, Pre-processing, Feature Extraction, Classification and Recognition. Finally, E-Speak voice command software is used to convert the obtained text from OCR into speech command. [4] This methodology is only for the conversion of printed Tamil text. The device will be held like a image capture camera and captured over a printed page. The input is taken by a camera in the device and the output is given as speech through microphone using the hardware interface. [7] Pre-processing is the basic operation on input image like binarization, noise reduction. Segmentation stage for segment the given image into line by line. Future extraction calculates the characteristics of character. A classification contains the database and does the comparison. [12] The general principle of operation for smart glasses is by giving instructions via switches and listening to the output through an earpiece. For text recognition mode, the view is processed in real time to get the image sent to an OCR software for text extraction and subsequently forwarded to a text-to-speech synthesizer. The text is then read through the audio output port.

## 3 Scope

Visually impaired people report numerous difficulties with accessing printed text using existing technology, including problems with alignment, focus, accuracy, mobility and efficiency. We present a smart device that assists the visually impaired which effectively and efficiently reads paper-printed text. The proposed project uses the methodology of a camera based assistive device that can be used by people to read Text document, search for the specific word and can also find the meaning of that word. The framework is on implementing image capturing technique in an embedded system based on Raspberry Pi board. The speech output is given as audio form where the user can hear the voice either through speaker or headphones. The user voice is taken as input with the device called Raspberry Pi mic which converts the user voice to text and perform the necessary action like find and search.

## 4 Block Diagram

Figure 1 shows the block diagram. The written text is to be placed underneath the camera view by the visually impaired individual to make sure the image is of fine quality and fewer distortions. On startup, the application checks for the availability of all the devices and also for the connection. The Raspberry Pi is an embedded system which has integrated peripheral devices.

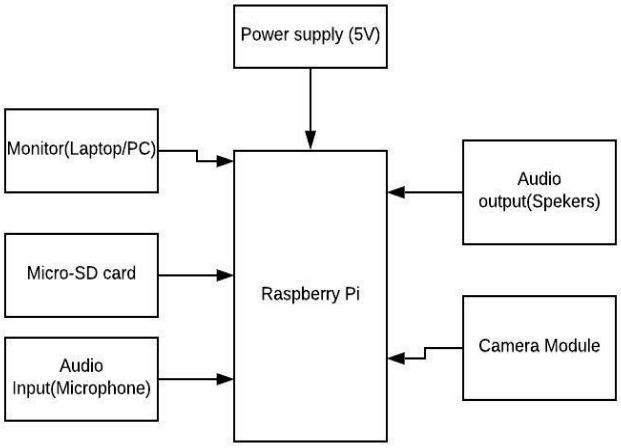


Fig. 1 Block diagram

## 5 Objective

The main objective of this project is to build a smart bot using raspberry pi which is not only used for reading the text but also it will be helpful in finding a word in large counts of text and also one can able to say a word and find out the exact meaning of the given word. Although a number of reading assistants have been designed specifically for the visually impaired, to our knowledge, no existing reading assistant can read text from the kinds of challenging patterns and backgrounds found on some books. Such text information can appear in multiple scales, fonts, colors, and orientations.

In order to help visually impaired people to read text from hand-held objects as well as books, we have developed a camera-based assistive text reading framework to track the interested area within the camera view and extract printed text information from the object or book. The proposed algorithm can be effectively used to handle different background patterns, and extract text information from any kind of hand-held objects or books. Nowadays people getting more tired of reading large number of pages and also a person is needed for visually impaired people to read a text loud and to explain them while we build this anyone can hear the text, find the word and meaning by their own.

Another main advantage of this project is this does not consume the storage space in the raspberry Pi through memory card or other extra storage devices for image and audio files. The captured images are directly stored into the Amazon web server (AWS) S3 bucket which can be accessed by the public URL it can either be uploaded or downloaded from and to the bucket. Thus, the image and audio files are stored in secured manner. The public URL is fetched by the python programming that get executes with the S3 bucket.

# 6 Methodology

The main objective is achieved by capturing the hard copy of documents with raspberry pi camera. The documents can be either handwritten or printed text and the captured image is processed and delivered it in the form of speech which is audio output.

**Added Features:**

- One can search for a specific word
- Capable of finding the meaning of a word
- Storage to the AWS server (S3 bucket)

The design is motivated by preliminary studies with visually impaired people, and it is small-scale and mobile, which enables a more manageable operation with little setup. The proposed fully integrated system has a camera as an input device to feed the printed text document for digitization and the scanned document is processed by a software module the OCR (optical character recognition engine). A methodology is implemented to recognition sequence of characters and the line of reading. As part of the software development the OpenCV (Open source Computer Vision) libraries is utilized to do image capture of text, to do the character recognition. Most of the access technology tools built for people with blindness and limited vision are built on the two basic building blocks of OCR software and Text-to-Speech (TTS) engines.

The proposed architecture of the smart reading bot is shown in fig 1, where the plastic sheet indicates the text to be read. Optical character recognition (OCR) is the translation of captured images of printed text into machine-encoded text. OCR is a process which associates a symbolic meaning with objects (letters, symbols a number) with the image of a character. It is defined as the process of converting scanned images of machine printed into a computer processable format. Optical Character recognition is also useful for visually impaired people who cannot read Text document, but need to access the content of the Text documents. Optical Character recognition is used to digitize and reproduce texts that have been produced with non-computerized system. Digitizing texts also helps reduce storage space. Editing and Reprinting of Text document that were printed on paper are time consuming and labor intensive. It is widely used to convert books and documents into electronic files for use in storage and document analysis.

OCR makes it possible to apply techniques such as machine translation, text-to-speech and text mining to the capture / scanned page and the final recognized text document is fed to the output devices depending on the choice of the user. The output device can be a headset connected to the raspberry pi.

# 7 Modules

**Capture Image with recognizing the Voice note given by User.**

This module is implemented with the feature of image recognizing. when the user gives the voice as **“Start capture”**, the raspberry pi camera is enabled and image starts captured what is present in front of the camera that can either be printed or handwritten image. When camera is ON the audio is given out as **“Image is Capturing”.** Thus, the given image is captured.

**Captured image is pushed into AWS S3 bucket.**

Finally, the given image gets captured and the captured image is named uniquely and get pushed into the EC2 instance with S3 bucket. The stored image into the unique S3 bucket returns the public url for accessing the image through web browser thus the url is used into the python code for further processing with the image. This makes space consumption in raspberry pi which is stored in the secured place. Thus, the access permission of the url can be set to private also when no one needs to view the image otherwise it is set to public URL.

**Image to text conversion.**

This module deals with the feature called image to text conversion. The public URL is used here to recognize the image by the python code and the text present in the image is given out as audio by recognizing each and every word in the image. The

text recognizing and the accuracy of the extracted text entirely depends upon the camera clarity. High end pixel camera gives more accuracy and clear text extraction from the image.

**Text to speech conversion.**

In this module the retrieved text from the module 3 is given out as audio output. While module integration the text is given out as an audio output whenever the capturing if the image is done. The audio file gets stored as a separate audio file in the device or it can also be uploaded to s3 bucket with (.mp3 format). The advantage of soring the audio file is that whenever the user wants to hear it again, they can just play the file without repeating the whole process.

**Handwritten image recognition (image to text for handwritten image).**

This module is same as module 1 only the method of processing the printed and handwritten image is different. However, the user voice is same as "Start Capture" for both printed and handwritten images only the processing gets different. Remaining all the output method and input method by the user is same.

**Speech to text recognition**

This module is developed in focus of recognizing the speech given by user and converting them to text. The converted text is passed in the form of string for the required modules so that the required feature gets executed like capturing the image, finding the string and finding the meaning of the given word all these are done with the help of speech to text module.

**Search for a word**

This module is implemented for searching a given word in the retrieved text. The word to be searched as given as speech as per the module 6 the speech is converted to text and the given word is searched after performing search operation the audio output is given as "String is present" and if not present the output is given as "String is not present" please search for some other word.

**Find meaning of the given word through voice.**

This module is same as module 7 the only thing is the feature of this module is to find the meaning of the given word from the oxford dictionary implemented by python code. It fetches the meaning of the word and gives as text which is then processed by text to speech algorithm and the audio file given. The user can hear the meaning of the word through headphones or speakers connected with the system.

The result and output of each and every module is attached and discussed in chapter 7 (Result and Discussion). Each module gives the accurate result and the audio output. The screenshots of the module give detailed explanation about each module what it exactly performs and how the output get displayed in the screen.

# 8 Flow of Process

Figure 2 illustrates the flow of process of the proposed method.

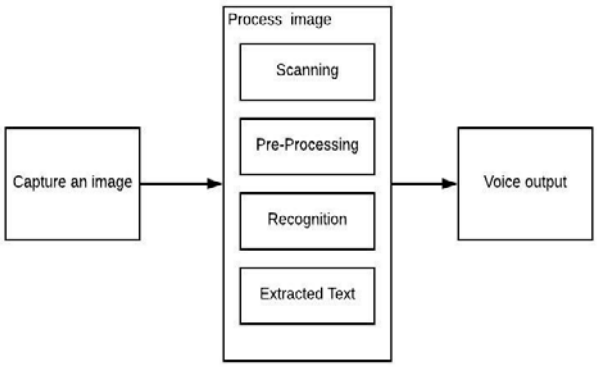


Fig- 2: Process Flow

# 9 Hardware Implementation

- Raspberry Pi 3B+
- HDMI cable
- Laptop or Desktop
- USB speakers or Headphones
- USB Mic
- Camera 5MP
- Ethernet Cable
- USB cable
- Connecting wires

Fig- 3: System Architecture

## 10 Software Implementation

- **IDLE-** Python 3.6 or Pycharm
- **Operating System** -Raspbian OS (jessie)
- Ubuntu OS
- Library files
- AWS account

The operating system which most of the systems or projects is built is Raspbian which is derived from the Debian operating system. Python language is a script language is used for writing the algorithm. OpenCV library is used for calling the functions in the algorithm. OpenCV is an open source computer vision library, which is written using C and C++ and runs in Linux, Windows and Mac OS. OpenCV was designed for machine potency and with a robust target on real-time applications. OpenCV is written in optimized C and can make the most of multi-core processors.

## 11 Result & Discussion

The proposed idea is not only helpful for visually challenged people it can be used by the normal people also. The whole setup will be designed in the form of product with battery backup so that people can take with them anywhere. In future the features can be added like different language output. Accuracy of text output depends upon the image clarity and the camera pixels which is used in the product. The setup cost will be normal like only raspberry Pi costs high other requirements like speaker, camera and mic depends upon the user wish. If they expect the product with high quality the speakers, camera and mic will be installed with high prices else it will be normal. As per the components installed into the product the prices vary.

The main advantage over this concept is battery storage space consumption. The only thing the user should concentrate in the voice the voice input should be clear and audible so that it recognizes the speech more clearly and perform any of the action driven by voice. So, whoever is using this device the voice should give in the proper format else the alert will be given so that the user has to repeat it once again When it gets recognized the output will be given as per the voice input given by the user. In future this can be extend with other features like different languages and with full battery backup and it can be a successful device.

Here we end up with the note that this can be served as

**"THIRD EYE FOR THE BLIND"**

*References*:

[1] Amal Jojie, Ashbin George, Dhanya Dhanalal Nayana J, Book Reader for Blind, IOSR Journal of Engineering (IOSRJEN).

[2] S. Aditi, SP. Annapoorani, A. Kanchana, Book Reader Using Raspberry Pi for Visually Impaired, International Research Journal of Engineering and Technology (IRJET), Volume 05, Issue 03, March 2018.

[3] K A. Aslam, Tanmoy Kumar Roy, Sridhar rajan, T. Vijayan, B. Kalai Selvi Abhinayathri, Smart Reading System for Visually Impaired People, International Journal of MC Square Scientific Research, Volume 09, Issue 02, 2017.

[4] V. Ajantha Devi, Dr. S Santhosh Baboo, Embedded optical character recognition on Tamil text image using Raspberry Pi, International Journal of Computer Science Trends and Technology (IJCST), Volume 02, Issue 04, Jul-Aug 2014.

[5] Mallapa D. Gurav, Shruti S. Salimath, Shruti B. Hatti, Vijayalaxmi I. Byakod, B-LIGHT: A Reading aid for the Blind People using OCR and OpenCV, International Journal of Scientific Research Engineering & Technology (IJSRET), Volume 06, Issue 05, May 2017.

[6] S. Rajakumar, Dr.V. Subbiah Bharathi, Century Identification and Recognition of Ancient Tamil Character Recognition, International Journal of Computer Applications, Volume 26, Issue 04, July 2011.

[7] Rahul R. Patil, Audumbar R. Misal, Ketan R. Nalawade, Survey paper on Text Recognition Using Image Processing, International Journal of Advanced Research in Electronics and Communication Engineering (IJARECE), Volume 04, Issue 03, March 2015.

[8] Praveen Choudhary, Dr. Vipin Kumar Jain, Text Extraction from an Image by using Digital Image Processing, International Research Journal of Computer Science (IRJCS), Volume 05, Issue 07, July 2018.

[9] Nagaraja L, Nagarjun R S, Nishanth M Anand, Nithin D, Veena S Murthy, Vision based Text Recognition using Raspberry Pi, International Journal of Computer Applications, National Conference on Power Systems & Industrial Automation (NCPSIA), 2015.

[10] Anush Goel, Akash Sehrawat, Ankush Patil, Prashant Chougule, Supriya Khatavkar,

Raspberry Pi Based Reader for Blind People, International Research Journal of Engineering and Technology (IRJET), Volume 05, Issue 06, June 2018.

[11] D. Velmurugan, M.S. Sonam, S. Umamaheswari, S. Parthasarathy, K.R. Arun, A Smart Reader for Visually Impaired People Using Raspberry PI, *International Journal of Engineering Science and Computing* (IJESC), Volume 06, Issue 03, 2016.

[12] Esra Ali Hassan, Esra Ali Hassan, Smart Glasses for the Visually Impaired People, Computers Helping People with Special Needs, 15th International Conference, ICCHP, July 2016.

[13] Aaron James S, Sanjana S, Monisha M, OCR based automatic book reader for the visually impaired using Raspberry PI, International Journal of Innovative Research in Computer and Communication Engineering, Volume 04, Issue 7, January 2016.

[14] Abhijith Shaji, Abhishek Aravindan, Nisham Rafeeque, Naveen K K, Reading assistant for visually impaired people, International Research Journal of Engineering and Technology (IRJET), Volume 05, Issue 04, April 2018.