System to provide Reading Aid to Visually Impaired People

Saloni Chaturvedi
EXTC Department
SVKM's NMIMS, MPSTME
Mumbai, India.
saloni.chaturvedi07@nmims.edu.in

Milind Chechani
EXTC Department
SVKM's NMIMS, MPSTME
Mumbai, India
milind.chechani08@nmims.edu.in

Krisha Gautami
EXTC Department
SVKM's NMIMS, MPSTME
Mumbai, India
krishagautami.pandey37@nmims.edu.in

Kashyap Joshi
EXTC Department
SVKM's NMIMS, MPSTME
Mumbai, India
kashyap.joshi@nmims.edu

Abstract— **Artificial intelligence and Machine learning is fast evolving in today's world running towards digitization and automation, ranging from cancer detection using scans in Healthcare to driverless automatic electric cars. The system to provide reading aid is an approach to help said differently abled people to gain emotional independency in life, when it comes to reading text visuals, be it in the form of a book, newspaper or notice board. It will help increase scope of education for visually impaired which was formerly relatively less possible due to unavailability of books in Braille. For this technology, OCR (Optical Character Recognition) and Signal Processing (Raw images) is use, coupled with a camera module attached to the Raspberry Pi.**

Keywords— *Optical Character Recognition (OCR), OpenCV, Raspberry Pi, Raw Image Processing, TTS, Visual*

I. INTRODUCTION

This project was selected as this contributes to the healthcare sector of the country and upliftment of the differently abled visually impaired people, trying to increase their quality of life by just a tad bit, getting them to be able to live a life of normalcy just like any other individual. There is quite a big extent of unavailability of all the books in the language Braille. Even if books are made available in said language, they are quite population of the world. Therefore, the unavailability of all the books in the language combined with the expenses, becomes a very inconvenient purchase for the differently abled blind people. If people try to go about their lives with an alternate, that is either having a expensive thus rendering themselves unaffordable and inaccessible by majority professional or a family member helping them read, it adds to the cost of the individual's time for reading out books to the visually impaired person. Moreover, the emotional factor of dependency always has a higher weighted average on itself when it comes to talking about the mental health of the said people. To eradicate such problems from the world is the objective of this system, for which the idea was fixated on a venture to make a device which provides as an aid to the visually impaired people in order to read any kind of physical text present in front of the camera which will be mounted on the wearable device given to these people.

A. Detailed Objective

The device aimed to be built, a wearable camera, mounted on the glasses that a visually impaired person can wear, revolves around the idea of being able to capture any evidence of physical text that is present in front of the camera. This text can be a piece of paper from a book, an article of the newspaper, a notice on the noticeboard of the residing building, an LED screen of the lift or a road sign. It aims to increase the visually impaired person's quality of life by reducing the need of dependency on any external factor by the visually impaired individuals for any kind of daily chore that is otherwise easily available to the co-existing individuals. The camera, after having clicked the picture, processes the raw image and brings it up to a level of standardized processed image, which is then given to a pre written code instilled in the said device, mounted on the glass which will in turn be connected to an earpiece, giving out the output in the form of an audio signal via either an earpiece or speaker, pertaining to the choice of the user.

B. Statistics for Motivation

India houses approximately 62 million visually impaired people, of which 54 million have low vision, and 8 million people are completely blind. From the graph and statistics attached below, it can be seen that India and China or one of the most densely populated with the people have been slow or completely blind and visually impaired problem, with a very and significant difference between themselves.
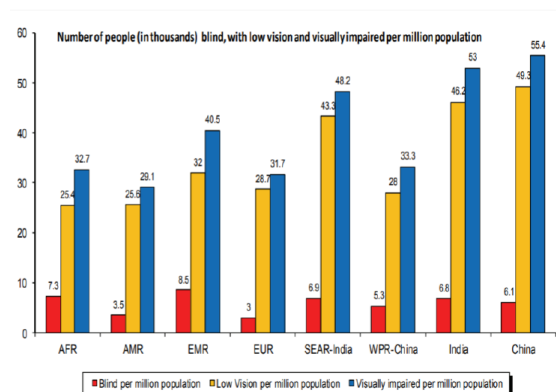

Fig. 1. Comparison of Visually Impaired Numbers in World (WHO) [1]

According to WHO survey and research, India accounts for about 20% of the entire world visually impaired people, which is an alarmingly high number, but is expected along with the population that India is a habitat to, which is also at the second rank of the entire world, thus cancelling out with the proportionality equivalence. The motivation to do this project, was to help this 20% of the world's population as first step, with this wearable and easily accessible, expensive in nature device, thus giving instant boost in the quality of life of the user.

Knowing India's economic conditions and government procedures, it is very difficult for users to apply for medical loan to get their impairments corrected (which deems to be the case for major population accounted in the said 20%) or get any kind of medical assistance, be it for fatal & lethal diseases like cancer, let alone an issue which is considered trivial by users and government themselves. System aims to provide an aid and contribute to this social and health cause prevalent since a long time, at prices affordable for one and all, residing in the country as first priority.

II. SYSTEM ANALYSIS

A. Hardware System Description

The heart of the system is raspberry pi unit which process all the input data and gives appropriate output. Firstly, an image is captured with the help of the text that encapsulates the text. This output raw image is provided to the Image processing section as input. The output from processing section, i.e., the processed image output is given as input to the filter for the noise signals reduction to minimise noise. The filtered image is provided as input to the edge reduction unit. Output image is then fed to the background separation unit. In the end, output is provided to the OCR (Optical Character Recognition). This fully processed image is further given to the raspberry pi that detects the image content and provides an output signal in the form of audio signal.

Step 1: Capturing the Image – Raspberry camera module

Step 2: Processing the Image – Python code in Raspberry Pi

Step 3: Filtering – Python code in Raspberry Pi

Step 4: Edge Detection – Python code in Raspberry Pi

Step 5: OCR conversion – Python code in Raspberry Pi

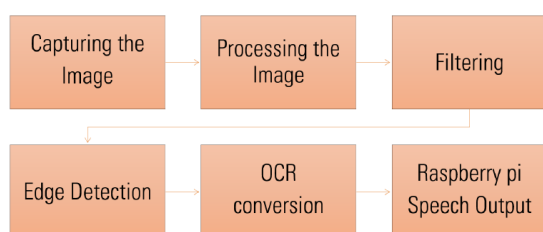*Step 6:* Raspberry pi Speech Output – Speakers

Fig. 2. Flowchart of Working

The following components are used in the design of
1. A 1.2GHz 64-bit quad-center ARMv8 CPU
2. 802.11n Wireless LAN
3. Bluetooth 4.1
4. Bluetooth Low Energy (BLE)
5. 4 USB ports
6. 40 GPIO pins
7. Full HDMI port
8. Ethernet port
9. Consolidated 3.5mm sound jack and composite video
10. Camera interface (CSI)
11. Display Interface (DSI)
12. Micro SD card space
13. Video core IV 3D Illustrations

A wooden box is used for keeping Raspberry pi, speakers, wires, power bank etc. inside the box. Using a wooden box helps to reduce the humming sound i.e., the vibrations when the audio(output) is heard from the top of the box. This box also helps to keep all the delicate equipment like Raspberry Pi safe from the outside world. The camera will be mounted on the top of the box with the help of a stand attached to one side of the box. The text to be read is put on the top.

Fig. 3. Wooden box for Hardware Integration

Fig. 4. Top view of Wooden Box

Step 1: A wooden box of size 16x8x4 inch is used.

Step 2: Magnetic lock is connected between the lid and the box so that it is easy to open and close.

Step 3: Hole(s) are drilled for the speaker on the top of the box.

Step 4: A piece of aluminum strip is used for the camera mount. It is easily adjustable for focusing the camera. The focus and size of the document decides the height of the camera. The adjustable stand helps to go high to read larger area document. *Step 5:* A 1"x1/16" slit id cut in the box top for the camera cable to pass through. No components are permanently cut unless a certainty of stability is achieved.

Step 6: A 15 cm long camera cable is connected to the camera.

Step 7: The camera is mounted in such a way that it is facing downward from the end of the 5" wood strip
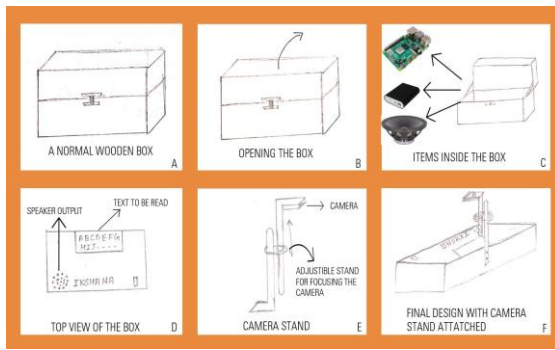
Fig. 5.   Sketch representation for visuals on Hardware Integration

Step 8:   The camera cable is slides along the slit then further attached to the Raspberry Pi. The Raspberry is turned off at this time.

Step 9:   The power supply given to the system by connecting to the power bank and plug into the Pi.

Step 10:   For the audio, a USB powered mono speaker is used and speaker is connected to the PI using an audio plug and USB cable connected Pi USB. The speaker is mounted under the lid of the box, and holes are there in the shape of the speaker where audio can be heard.

Step 11:   There is no on/Off switch, as it is assumed that the Pi should be running all the time, so it is assumed that it is ready to read something immediately. PI can run easily using mere power supply and can run 24/7 without issues. Though, it is possible for it to become corrupted if unplugged or power failure, it is rare.

B. *Software System Description*

Raspberry Pi OS was chosen for the project against options like Windows IOT core because Raspberry Pi OS is a full-blown OS available for Raspberry Pi while others are heavily trimmed down in terms of features. It provides more than a pure OS, the variety of packages and pre-compiled nature of software made for easy installation on the Raspberry Pi.

The code for the system is written using Python 3.7. The coding flexibility & dynamic nature of python made it the ideal choice to create an intelligent device within the limitations of the hardware. The libraries used in this code are:

1. OpenCV - The OpenCV functions used in the project are imread function to read the image input from the raspberry pi, image processing functions like resize, cvtColor, dilate, medianBlur, merge, erode, threshold and imwrite function to write the processed output image to provide as input to the OCR system.
2. Tesseract - Tesseract was chosen due to the accuracy it provides when compared with alternatives like Google Vision API. The processed image from OpenCV was provided to Tesseract to perform OCR on and extract the text from the physical image of the text. The final output received from this engine was text from the image which was fed to the text to speech engine.
3. PyTTS - Since it is based on various Google translate engine, PyTTS can be used to speak out various languages, even translate from one language to

another. This proved to be a major advantage as it could be useful in future modifications to the project.
4. Secondary Libraries - Numpy is used to write an optimised OpenCV code and yield the best results. OS library is used to provide functions to interact with the Operating System used in the code. Time library in Python was used to provide functions for working with times, and for converting between representations. Threading module was used to construct higher-level threading interfaces and provide optimization against threading in Python.
5. Pycharm is preferred in this case as opposed to an online environment like jupyter notebook because it has the ability to generate complete projects which can include files of various formats apart from just the .py files.

III.   METHODLOGY

Firstly, any physical evidence of text, like a piece of article from the newspaper, a page from any book, a noticeboard, an LED screen of the lift or a sign on the road, is captured by the camera.

This captured image is then sent for image pre-processing. There are various steps involved in the pre-processing of the image to convert the raw picture into a noiseless output which can be given as input to the raspberry pi for further programming. The first step that is taken in the pre-processing of the image is resizing where in the image is cropped or enlarged according to the requirements of the code and as per the availability of the captured image.
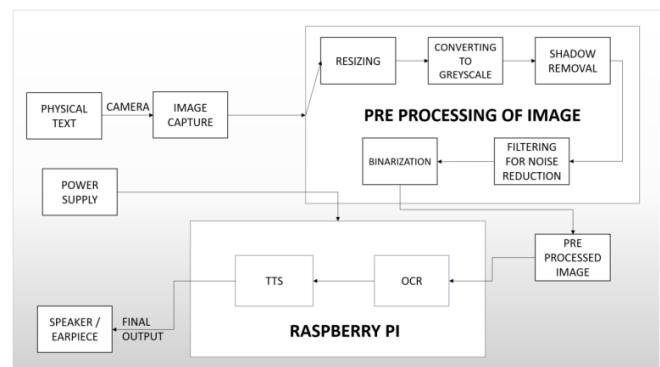

Fig. 6.   Block Diagram of the implemented system

The next involved step is the conversion of picture matrix to greyscale where in the color or black-and-white picture is converted to an array of greyscale that is easily readable and recognizable by the pre fed python code in the raspberry pi.

Following this, the removal of shadows is conducted by the preprocessing code in the captured image.

The next step is filtering of the image for noise reduction for gaining an output which has minimum possible noise for easy recognition by the machine.

The final step of reprocessing of the image binarization, which is conversion of the image into machine readable text.

With binarization, the entire process of pre-processing of the image is completed. After this, the, output of the preprocessing code is fed as input to the raspberry pie in which OCR, that is the optical character recognition, takes
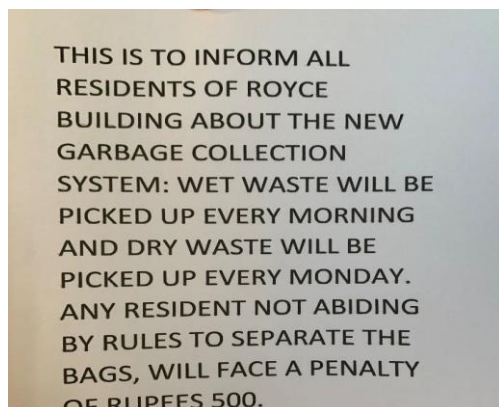
place via the code that is already fed into the device. After optical character recognition, text to speech conversion takes place of this particular input that is given to the TTS algorithm.

The final output is received in the form of audio signal which is given by the speaker or earpiece that is also attached to the raspberry pie, pertaining to the choice of the user.

In order to analyse the working of device 3 test images were considered to gauge the capacity of the system to extract the relevant text and convert it to a string which is then spoken by the system.

The table 1 illustrates the results for 3 different input images that were given to the system. In this section, results are analysed and reasoning presented for the obtained results and conclusion.

IV. RESULTS AND ANALYSIS

TABLE I. RESULT ILLUSTRATION FOR 3 DIFFERENT IMAGES

	Image 1	Image 2	Image 3
Nature	Image of a bill from a restaurant.	Image of a page from a novel.	Image of an official notice from a society to its residents.
Characteristic	Clear characters, no italics or special font, poorly lit and some characters close together due to a small area of paper.	Noisy image of a page from a novel where the font is italic, almost cursive. The book was folded in half hence the image is curved. Half the page was well lit and half of it was dark	Clear image of a notice with distinct characters and some shadows and noise.
Result of pre-processing	The pre-processing worked exceedingly well it removed the shadows and noise from the image.	Even after preprocessing there was a significant difference in the shadows on the half that was towards the light and the half that was away from it.	The pre-processing worked well removing the noise and shadows almost entirely,
Relevant text extracted	Yes	No	Yes
Does speech output convey the intended message	Partially	No	Yes

Fig. 7. Example- Input Vs Output Image

From analysing the result for 3 different inputs, it was concluded that even the cleanest of pictures will have shadows and the pre-processing block in place to remove shadows works well enough to remove the shadows drastically from even poorly lit photos. Another observation was that the system is not equipped with removing cursive or compacted text, this variable is completely dependent on the image taken by the user. This issue could be resolved by creating custom resizing settings for such text.

III. CONCLUSION

For the sole purpose of improving upon the problems faced by the visually impaired people in this world, of which India accounts for 20% of, be it chronic or acute combined with daily problems, this medical aid for the said users was worked upon. Here in, an image of any kind of physical evidence of text is clicked by the camera which is mounted on the wearable device intended for the user, which is then pre-processed to convert into a machine-readable text.

After OCR and TTS, the output is given in the form of audio waveform of the written text in the image captured by the camera via the earpiece or the speaker, attached to the device itself, pertaining to the choice of the user. This device aims on increasing the future scopes of the visually impaired, with advancements in technology, this device can help increase efficient mobility with almost no dependency, which in turn also increases the emotional factor of independency thus resulting in good mental health of the individual. Education and educational institutes can be revolutionized using this equipment, expanding career

options for the visually impaired. Keeping the poverty line of India in consideration, the aim is to build this device at a cost which is affordable to all, and help in contributing to this social cause of the country.

This device can be ventured for various future prospects like schools and colleges for blinds, text from videos (movies), vertical reading, mathematical equation solver and facial recognition.

REFERENCES

[1] Aravind Kumar, *"Text to Speech Conversion System using OCR"*, ISSN 2250-2459, ISO 9001:2008 Certified Journal, Volume 5, Issue 1, January 2015

[2] KN Kumar, P Surendranath, K Shekar, *"Assistive Device for Blind, Deaf and Dumb People using Raspberry-pi,"* Published 2017, Engineering Imperial journal of interdisciplinary research

[3] Urmila Shrawankar, Shruti Gedam, *"Finger Spelling in Air System for Deaf and Dumb"*, January 2018 International Journal of Technology Diffusion 9(1):28-44 DOI: 10.4018/IJTD.2018010103

[4] Tanay Choudhary, Saurabh Kulkarni, P Harshavardhan Reddy, *"A Braille-based mobile communication and translation glove for deaf-blind people"*, Published 2015 Computer Science 2015 International Conference on Pervasive Computing (ICPC)

[5] A Karmel, Anushka Sharma, Muktak Pandya, Diksha Garg, *"IoT based Assistive Device for Deaf, Dumb and Blind People"*, Volume 165, 2019, Pages 259-269

[6] Shalini Sonth, Jagadish S. Kallimani, *"OCR based facilitator for Visually Impaired"*, 2017 International Conference on Electrical, Electronics, Communication, Computer, and Optimization Techniques (ICEECCOT)

[7] Roberto Neto, Nuno Fonseca, *"Camera Reading for Blind People"*, Polytechnic Institute of Leiria, Leiria 2411-901 Leiria, PORTUGAL

[8] Rizqi Andry Ardiansyah, *"Design of an Electronic Narrator on Assistant Robot for Blind people"*, Research Center of Electrical Power and Mechatronics, Indonesian Institute of Sciences, Cisitu 21/154 D, Bandung 40135, Indonesia

[9] Nagaraja L, Nagarjun R S, Nishanth Anand, Nithin D, Veena S Murthy, *"Vision based Text Recognition using Raspberry Pi"*, International Journal of Computer Applications (0975)

[10] Yahvin Gali, *"Rethinking Braille with Pi reader: Raspberry Pi Based Optical Reader for the Blind and Visually Impaired using OCR & TTS"*, California Science Engineering Fair