SELF-ASSISTIVE DEVICE FOR THE BLIND, DEAF AND MUTE

A PROJECT REPORT

Submitted by

Diksha Garg (15BCE1148)

Anushka Sharma (15BCE1255)

Muktak Pandya (15BCE1263)

in the partial fulfilment for the award

of

Bachelor of Technology

in

Computer Science and Engineering

Vandalur - Kelambakkam Road, Chennai - 600 127

April - 2019

School of Computing Science and Engineering

DECLARATION

We hereby declare that the project entitled "**Self-Assistive device for the Blind, Deaf and Mute**" submitted by us to the School of Computing Science and Engineering, Vellore Institute of Technology, Chennai Campus, Chennai 600127 in partial fulfilment of the requirements for the award of the degree of **Bachelor of Technology in Computer Science and Engineering** is a bona-fied record of the work carried out by us under the supervision of **Dr. Karmel A**. We further declare that the work reported in this project has not been submitted and will not be submitted, either in part or in full, for the award of any other degree or diploma of this institute or of any other institute or university.

Place: Chennai

Signature of Candidate (s)

Diksha Garg (15BCE1148)

Date:

Anushka Sharma (15BCE1255)

Muktak Pandya (15BCE1263)

School of Computing Science and Engineering

CERTIFICATE

This is to certify that the report entitled "**Self-Assistive device for the Blind, Deaf and Mute**" is prepared and submitted by **Muktak Pandya (15BCE1263)**, **Anushka Sharma (15BCE1255)** and **Diksha Garg (15BCE1148)** to VIT Chennai, in partial fulfilment of the requirement for the award of the degree of B.Tech CSE programme is a bona-fide record carried out under my guidance. The project fulfils the requirements as per the regulations of this University and in my opinion meets the necessary standards for submission. The contents of this report have not been submitted and will not be submitted either in part or in full, for the award of any other degree or diploma and the same is certified.

Approved by the Guide and the B.Tech CSE Program Chair.

Name: **Dr. Karmel A.**

Date:

Name: **Dr. Rajesh Kanna B.**

Date:

Internal Examiner

Signature:
Name:
Date:

External Examiner

Signature:
Name:
Date:

ACKNOWLEDGEMENT

We wish to express our sincere thanks to a number of people without whom we could not have completed the thesis successfully.

We would like to place on record our deep sense of gratitude and thanks to our guide Dr. Karmel A., School of Computer Science and Engineering (SCSE), VIT, Chennai campus whose esteemed support and immense guidance encouraged us to complete the project successfully.

We would like to thank our Program Chair Dr. Rajesh Kanna B. & Co-Chair, Dr. C Sweetlin Hemalatha, B.Tech. Computer Science and Engineering and Project Co-ordinator Dr. B V A N S S Prabhakar Rao, VIT Chennai campus, for their valuable support and encouragement to take up and complete the thesis.

Special mention to our Dean to Dr.Vaidehi Vijayakumar , Associate Deans Dr. Vijayakumar V and Dr.Subbulakshmi T, School of Computing Science and Engineering (SCSE), VIT Chennai campus, for spending their valuable time and efforts in sharing their knowledge and for helping us in every minute aspect of software engineering.

We thank our management of VIT, Chennai campus for permitting us to use the library resources. We also thank all the faculty members for giving us the courage and the strength that we needed to complete our goal. This acknowledgment would be incomplete without expressing the whole hearted thanks to our family and friends who motivated us during the course of our work.

Muktak Pandya (15BCE1263)

Anushka Sharma (15BCE1255)

Diksha Garg (15BCE1148)

EXECUTIVE SUMMARY

Assistive device for the blind, deaf and dumb caters to the need of assistance for people with disabilities like blindness, deafness and muteness. It reads for the blind, displays text for the deaf and speaks for the mute. Since it is a small portable device it can be used at homes or on the go. It provides users with an embedded camera that takes images as input and converts them to text and ultimately into speech so that the visually impaired can hear and know which street they are headed towards or what is on the next page of a book. It uses image recognition and text detection to do so. It also reads handwritten text and converts it into speech for intercommunication between the visually impaired and audibly challenged. It also uses text to speech and speech to text techniques and displays them on a display screen with a keyboard for communication for the audibly challenged.

All in all it is a combined solution for three problems that will eradicate the need for buying separate devices for use and intercommunication of the differently abled community.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

Sr. No.	Abbreviations	Full Form	Description
1.	WHO	World Health Organisation	It is an organisation that strives to combat diseases – communicable diseases like influenza and HIV, and non-communicable diseases like cancer and heart disease.
2.	API	Application Programming Interface	It is a set of subroutine definitions, communication protocols, and tools for building software. In general terms, it is a set of clearly defined methods of communication among various components.
3.	TTS	Text-to-Speech	A Text-to-Speech (TTS) system converts normal language text into speech.
4.	MUD	Multi User Dungeon	It is a multiplayer real-time virtual world, usually text-based, MUDs combine elements of role-playing , hack and slash, player versus player, interactive fiction, and online chat.
5.	IP	Internet Protocol	It is the principal communications protocol in the Internet protocol suite for relaying datagrams across network boundaries.
6.	POP3	Post Office Protocol 3	It is the most recent version of a standard protocol for receiving e-mail. POP3 is a client/server protocol in which e-mail is received and held for you by your Internet server.
7.	GSM	Global System for Mobile communication	It is a digital mobile network that is widely used by mobile phone users in Europe and other parts of the world.
8.	OCR	Optical character recognition	It Is the mechanical or electronic conversion of images of typed, handwritten or printed

			text into machine-encoded text, whether from a scanned document, a photo of a document, a scene-photo or from subtitle text superimposed on an image.
9.	I2C	Inter-Integrated Circuit	I2C is a serial protocol for two-wire interface to connect low-speed devices like microcontrollers and other similar peripherals in embedded systems.
10.	HTTP	HyperText Transfer Protocol	It is the underlying protocol used by the World Wide Web and this protocol defines how messages are formatted and transmitted, and what actions Web servers and browsers should take in response to various commands.
11.	LCD	Liquid Crystal Display	It is the technology used for displays in notebook and other smaller computers.
12.	DSP	Digital Signal Processing	It is the use of digital processing, such as by computers or more specialized digital signal processors, to perform a wide variety of signal processing operations.
13.	GPRS	General Packet Radio Service	It is a packet oriented mobile data standard on the 2G and 3G cellular communication network's global system for mobile communications.
15.	GPU	Graphic Processing Unit	It is a specialized electronic circuit designed to rapidly manipulate and alter memory to accelerate the creation of images in a frame buffer intended for output to a display device.
16.	ARM	Advanced RISC Machine	An ARM processor is one of a family of CPUs based on the RISC architecture.
17.	USB	Universal Serial Bus	It is an industry standard that establishes specifications for cables, connectors and protocols for connection, communication

			and power supply between personal computers and their peripheral devices.
18.	GPIO	General Purpose Input Output	It is an uncommitted digital signal pin on an integrated circuit or electronic circuit board whose behavior—including whether it acts as input or output—is controllable by the user at run time.
19.	HDMI	High Definition Multimedia Interface	It is a proprietary audio/video interface for transmitting uncompressed video data and compressed or uncompressed digital audio data from an HDMI-compliant source device, such as a display controller, to a compatible computer monitor, video projector, digital television, or digital audio device.
20.	I2S	Inter Integrated Sound	It is a serial bus interface standard used for connecting digital audio devices together.
21.	GUI	Graphical user interface	It is a form of user interface that allows users to interact with electronic devices through graphical icons and visual indicators such as secondary notation, instead of text-based user interfaces, typed command labels or text navigation.
22.	SSH	Secure SHell	It is a cryptographic network protocol for operating network services securely over an unsecured network.
23.	JPEG	Joint Photographic Experts Group	It is a commonly used method of lossy compression for digital images, particularly for those images produced by digital photography.
24.	STT	Speech-to-Text	A Speech-to-Text (STT) system converts normal language text into speech.
25.	SDRAM	Synchronous Dynamic	It is a generic name for various kinds of

		Random Access Memory	dynamic random access memory that are synchronized with the clock speed that the microprocessor is optimized for.
26.	CPU	Central Processing Unit	It is the electronic circuitry within a computer that carries out the instructions of a computer program by performing the basic arithmetic, logic, controlling, and input/output operations specified by the instructions.
27.	MIPI	Mobile Industry Processor Interface	It is standard defines industry specifications for the design of mobile devices such as smartphones, tablets, laptops and hybrid devices.
28.	REST	Representational State Transfer	It is a software architectural style that defines a set of constraints to be used for creating Web services.
29.	SFTP	SSH File Transfer Protocol	It is a separate protocol packaged with SSH that works in a similar way over a secure connection.
30.	MIMO	Multiple Input and Multiple Output	It is a method for multiplying the capacity of a radio link using multiple transmission and receiving antennas to exploit multipath propagation.
31.	RISC	Reduced Instruction Set Computer	It is a microprocessor that is designed to perform a smaller number of types of computer instructions so that it can operate at a higher speed.

LIST OF ABBREVIATIONS FOR UNITS

Sr. No.	Abbreviation	Full Form
1.	MHz	Mega Hertz
2.	Fps	Frames Per Second
3.	Mbits/s	Megabits per second
4.	Mm	Millimetre
5.	G	Grams

ABSTRACT

Focusing and addressing the problems faced by the differently abled people such as visually, audibly and vocally challenged, through a single device is a tough job. A lot of research has been done on each problem and solutions have been proposed separately. But not all of them are addressed together. The aim of the project is to create a single device solution in such a way that is simple, fast, accurate and cost-effective. The main purpose of the device is to make the differently abled people, feel independent confident by seeing, hearing and talking for them. The device provides a Google API and Raspberry Pi based aid for the blind deaf and dumb people. The proposed device enables visually challenged people to read by taking an image. Further, Image to text conversion and speech synthesis is done, converting it into an audio format that reads out the extracted text translating documents, books and other available materials in daily life. For the audibly challenged, the input is in form of speech taken in by the microphone and recorded audio is then converted into text which is displayed in the form of a pop-up window for the user in the screen of the device. The vocally impaired are aided by taking the input by the user as text through the built-in customized on-screen keyboard where the text is identified, text into speech conversion is done and the speaker gives the speech output. This way the device speaks for the user.

1. INTRODUCTION

1.1 Objective

The objective for this project is to integrate the solutions of various problems into a simple modern solution. The aim of this project is to combine the solution of three different disabilities namely- seeing, hearing and speaking into one. This integration will help the differently abled by giving an all in one solution instead of buying different gadgets for different problems. Upon the successful implementation the project's aim is to make the disabled independent and able to use the device on their own. This will give them a sense of independence and will also allow secure aid without the help of an outside person or agent effectively increasing secure communication and preventing the leak of sensitive information. This project is to increase the confidence of the differently abled by helping them in communication and establishing a connection to the world. A problem like this specially in India, where other problems of higher weight age prevail, has not yet been taken up. The main objective of the project is to build a secure device that is cost effective and of high functionality to help all the disabled communities regardless of the area they live in or the monetary conditions that they can afford.

1.2 Motivation

The motivation behind this project was the lack of focus on the problems of the disabled community in India. This is a major concern because the blind, deaf and mute community should be able to do mundane tasks like reading and intercommunication without assistance from people. Specially in the case of India where the special schools for the disabled children or special work places harbour people from all of the above stated disabilities. In such situations it becomes hard for the newly disabled or even the previously disabled to communicate with each other. The cases of blindness, deafness and muteness are mainly of three kinds. The first case being disability from birth, the second case is if a person is losing their senses gradually as they grow up, which is the only case in which they get the time to learn and prepare the special languages that are derived for them. The third kind is the loss of sense in a sudden accident. In cases where a person is already grieving from an accident that

brought upon the misfortune of the loss of a sense it becomes even tougher to adjust to the condition and learn how communicate. The thought of helping people in such a situation was what drove us to find a compact and easy to use solution for the problem of intercommunication among the disabled community.

In India, where health and sanitation are bigger problems than problems of a minority community, not many people focus on solving such problems. This is what bothered us enough to start work in this field because a community that needs most assistance is barely getting enough work done for them and all the work that is pre existing has too many devices to wear and use. For them to communicate they have to take many steps before someone can even understand one sentence of what they are trying to say. This also creates a problem in communicating in case of emergencies. The motivation is to make a device that could do faster processing and required much less steps to be used before the translation of sentences.

Not only was the aim to make a simple solution but also to make it secure so in case of sensitive topics being discussed the disabled community will have a way for secure communication and not be overheard protecting their basic right of privacy. Although there were some similar products they always required an assistance or they were not a secure source of communication. Furthermore, one driving force was also the cost of the already existing products. The objective was to focus on the development of a product that could be bought by and used for all the three disabled communities which was almost comparable to the cost of the already existing products.

But the main motivation of this project was to help the disabled communities be independent and to be able to use electronics easily. The motivation was to help lives by making them self sufficient and increasing their confidence and making them realise that they can stand shoulder to shoulder with other non- disabled people.

1.3 Background

Approximately 1.3 billion people live with some sort of vision impairment out of which 188.5 million people have a mild vision impairment, 217 million have moderate to severe vision impairment, 36 million people are blind and the majority of people with vision impairment are over the age of 50 years. India is considered to home the largest number of blind people. Around 9.1 billion people are deaf and mute. According to WHO, around 5% of the world's population or 466 million people suffer from disabling hearing loss.

Technology is advancing day by day and during the last few decades, it has made our lives easier and convenient. But some- how the physically impaired part of our society has not been paid enough attention to. They are deprived of the advancements of science and still face plenty of problems in their day to day lives.

Communication is a major aspect of human lives. But there still exists a gap. Braille and sign language are the means of their communication but it is out of their comfort zone. They always have to learn these traditional modes of communication or they bank on support such as another person.

This project majorly focuses filling this gap by trying to make them feel independent and that they too can walk hand in hand with the other normal people. Raspberry Pi and Google API being the two pillars of this device, make it accurate, efficient and robust. The device consists of three major modules, each dedicated to the visually, audibly and vocally challenged. It uses Raspberry Pi supported by Google API as the main unit, and also consists of a camera, microphone, speaker and a screen. For the visually challenged, the inbuilt camera takes a picture of the writ- ten or printed document and this image is then converted into digital text using the Google Vision API. This text is then converted into audio using the TTS (Text to Speech) library and voice converted output according to the written document or book is obtained. The audibly challenged are aided by recording the speech or audio, converting it into text and displaying it on the screen for the user to read it. The device speaks for the vocally challenged as it provides the user with a customized keyboard on the screen where the user can type the message. This text is converted into speech using the TTS (Text to Speech) library and audio for the input given by the user is obtained in a synthesized voice.

1.4 Literature Survey

Development of user centered interfaces and technologies have become crucial in the process designing for the differently abled people. Adding an extra element is just not enough to assist the use of technology for the visually disabled . Many device-based hardware and software technologies exist to assist the visually disabled. They have functions like reading printed or writ- ten text, expanding characters on braille systems and machines Based on computer vision. Prototypes that function with cell phone, cameras, help in processing images to identify patterns of movement, are applied for musicians who are blind.

AudioMUD [4] is a multiuser virtual environment exclusively made for the blind people and is associated with spoken cues. The original MUDs (Multi User Dimension) are generally text based and do not contain any sort of graphical interfaces. Users generally use MUD (Multi User Dimension) style games to per- form a set of actions in a virtual environment with a navigable space in the presence of direction, orientation and restrictions. There is high potential for the description of spaces and interactions due to its possible types of interaction and text based interface between players and virtual environment in AudioMUD with collaborative aspects. Their project focuses on the development of a server and client from scratch where the state of the world is stored in the server in such a way that when the server connects to the client, the state of virtual game is received and players can enter or exit anytime. The game starts when the blind user enters the IP name and server in the client, the player comes inside the kingdom of the human body like the respiratory system in a random location with attributes and can explore the system.

In [2] a system is suggested for the visually disabled to enhance the quality of their life. The wearable system consists of facial recognition to recognize people's faces and can identify a person through prior system training using fisher faces algorithm, obstacles detection where the user wears the device which uses ultrasonic sensors to generate vibration signals that indicate an obstacle, email reader which accesses user's email using POP3 protocol and enables the user to listen to the email using headphones, medication reminder to remind the user about the medication prescribed, MP3 player as a source of entertainment enabling the user to listen to music.

In [3] a system is suggested using raspberry pi that uses image acquisition using interfacing a webcam, preprocessing of image to obtain the region of interest, template identification to detect characters and objects, converting image to text using OCR algorithm which scans

image and gives a corresponding text output, and save the text data in a text file, and convert text to speech using E-speak for the blind user to hear the text.

Sign language which principally uses manual communication including hand movements, facial expressions to express, connects with people and convey their messages. In [10] a wearable device is designed for the deaf-blind users called GlovePi to identify the person, number and position of people, and their facial expressions in front of the user. It mainly comprises of a gardener glove which is attached to capacitive touch sensor with raspberry Pi using a I2C interface. Using many to many architecture in order to include maximum amount of users into an account, the Glove enables the user to register on the server usually by sending a HTTP request and eventually the user is added on the server after which the server sends a updated list of all the connected users and thus uses peer to peer communication to send or receive messages.

In [12] a system has been designed which is a mouth gesture recognition system using the help of an infrared sensor that collects the data from the audibly impaired person's mouth and detects the state of the mouth. They have designed three states: OSCS (Open Slow Close Slow), OSCF (Open Slow Close Fast) and OFCS (Open Fast Close Slow). When the sensor reaches its threshold, the sensor indicates and records the signal. Using different combinations, 27 patterns have been achieved which generated 26 alphabetic letters. The output of this proposed system depends on the light reflected from the object that the sensor subjected on, where the intensity supposedly gets affected by the surface color, shape and distance, after which the circuit gets the appropriate output analog voltage range.

Systems that suffice all solutions for the blind, deaf and dumb users in one compact device are rare to find. In [1] an arrangement has been intoduced for the visually impaired can understand words using Tessaract which is an OCR (Optical Character Recognition) algorithm by python, vocally impaired can ex- press and communicate by text which is read through E-speak, and audibly impaired can hear by speech to text conversion using OpenCV. In [11] an ideology has been brought forward that consists of a Sharon bridge which is a wearable technology that makes communication between differently abled on the extent of their capabilities. The Sharon Bridge comprises of small units to form a complete circuit to enable them to convey messages among the differently abled and their different combinations. It comprises of a sensor glove that is made up of arduino circuit board, tactile and flex sensors, and accelerometer which is used to convert the American sign language to audio that is further changed to text which is displayed on the LCD(Liquid Crystal Display)

for the user, Arduino GSM(Global System for Mobile communications) shield to communicate over long distances using the internet and GPRS(General Packet Radio Service) wire- less network, Beagle bone that converts analog to digital and vice versa. It works in a way where the message to be sent is the input as text, audio or braille language which is converted to the respective forms for the disabled to hear, speak or see. For long distances, the input in converted and sent through wireless GSM network to the receiver but the user is supposed to possess a phone number. Sharon Bridge works for all combinations of the blind, deaf and dumb.

2. PROJECT DESCRIPTION AND GOALS

2.1 Project Description

This project provides assistance for the blind, deaf and mute community. There are three modules in out project. The first module is for the blind. It takes input as the image and converts it to text using the Google Cloud Vision API. The image text is identified and using gTTS the obtained text is converted to speech and given as a output through the speaker.

The second module is for the deaf and the module takes spoken sentences as the input through the microphone and converts it to text which is displayed through the LCD screen. The last module is for the mute in which a keyboard is displayed on the LCD through which text input can be taken. The input is converted to speech to give the text a voice and the output is through the speaker.

The modules are integrated into a small portable device that is controlled by a three way switch that can change the modules according to the user's usability. The device has a single reset button for the reuse of the same module and is simple to use and manipulate.

2.2 Goals

The goals of the project include:

- Providing comfortable capturing of clear images and then converting the image text to normal text.
- Providing fast and easy conversion of text obtained by the image to speech into language pre set by the user.
- Providing clear recognition of speech captured by the microphone and conversion of recorded speech into text.
- Providing clear display of the text converted by an image or by the speech to text module.
- Integration of all the above modules into a consolidated easy to use, button push device so that the user can easily change modes by the push of a button.

3. TECHNICAL SPECIFICATION

3.1 Hardware Specification

3.1.1 Raspberry pi:

Raspberry Pi is a low cost, credit card sized processor, which can easily perform all task we expect from a desktop. It is very easy to connect raspberry pi with computers and Televisions. It also provides GPIO pins to connect with other components. Because of this efficiency to intercommunicate with the cross-disciplinary domain, it has been used in a variety of projects. Raspberry pi operates in an open source environment such as Raspbian (Linux based operating system).

Technical Specifications:

- Broadcom Soc BCM2836 (CPU, GPU, DSP, SDRAM)
- 900 MHz quad-core ARM Cortex A7 CPU (ARMv7 instruction set)
- Broadcom VideoCore IV @ 250 MHz GPU
- 1 GB Memory (shared with GPU)
- 4 USB ports
- 17 GPIO Peripherals plus specific functions
- 15-pin MIPI camera interface video input connector
- HDMI video outputs, composite video via 3.5 mm jack
- I²S audio input
- Analog audio output via 3.5 mm jack; digital via HDMI and I²S
- MicroSD for storage
- 10/100Mbps Ethernet speed
- 800 mA power rating (4.0 W)
- 5 V power source via MicroUSB or GPIO header
- 85.60mm × 56.5mm
- Weighs 45g (1.6 oz)

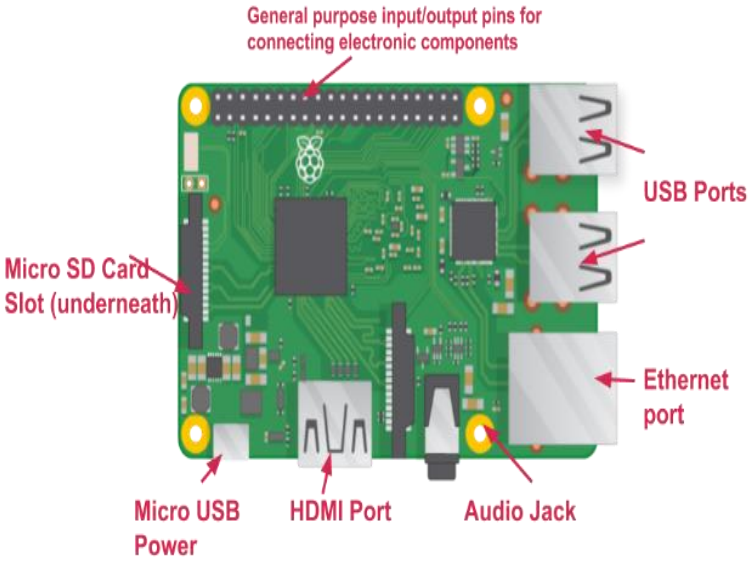

Figure 1: Raspberry pi 2 model

Figure 1 shows the Raspberry pi 2 model. The Input-Output pins are used to connect other components to the pi. The USB ports are for devices that take input or output through the USB ports, in our case the Microphone. The Ethernet port is to connect the pi to Internet. The Audio jack outputs audio in our case through speaker or earphones, according to the usage of the user. The HDMI port is to connect a monitor or a display screen to the Raspberry pi. The Micro USB is to give the pi a power supply and the Micro SD card slot is to give the pi a pre stored data like in case the user is using an Operating System it can be loaded up in the SD card and given to the pi.

3.1.2 Logitech C310HD Webcam:

The camera used in the project is a C310HD Logitech webcam shown in Figure 2 with a resolution of 720p/30fps. The images taken are crisp and contrasted. This camera fits perfectly in the project as it adjusts to the lighting conditions to produce brighter contrasted images. It uses a universal clip to attach itself firmly to the device. It is small, adjustable and agile and is therefore handy in the project.

Technical Specifications:

- Max Resolution: 720p/30fps
- Lens technology: standard
- Focus type: fixed focus
- Field of View: 60°
- Built-in mic: mono

- Cable length:1.5 m
- Universal clip fits laptops, LCD or monitors

Figure 2: Logitech C310HD webcam

3.1.3 Wave share 5 inch display:

The project consists of a 5inch resistive touch shown in Figure 3, with a high hardware resolution and HDMI Interface specially designed for the Raspberry Pi. It has a resistive touch control It is compatible and has a direct connects with any revision of the existing Raspberry Pi. It provides drivers and the backlight can be turned on or off for the lower power consumption. According to the requirements of the project, a keyboard has been hardwired in this 5-inch display for the vocally challenged to type their text in the screen.

Technical Specifications:

- Drivers provided (works with your own Raspbian/Ubuntu/Kali/Retropie)
- HDMI interface for displaying, no I/Os required (however, the touch panel still needs I/Os)
- High-quality immersion gold surface plating

Figure 3: Wave share 5 inch display

3.1.4 USB microphone:

A mini portable high quality USB microphone (Figure 4) is used in the project. It is a noise cancelling microphone which filters out unwanted background noise. It comes as a brownie point for the project as it is portable, compact and easy to use. It can be made more efficient according to the user or background by increasing the gain control or capture for better accuracy.

Technical Specification:

- 4.5V Working voltage
- Weight 99.8 g
- 2cm x 2cm x 0.5cm in size

Figure 4: USB microphone

3.2 Software Specification

3.2.1 Google cloud vision API:

The Google Cloud Vision API encapsulates powerful machine learning models in an easy to use REST API and enables developers and users to apprehend the content of an image. It is used for classification of images into thousands of categories, detecting individual objects and faces within images, and reading printed words contained within images. Optical Character Recognition (OCR) is used to enable the user to detect text within images, along with automatic language identification. Vision API supports a huge and broad set of languages. Initially Conventional Neural Network based model is used to detect localized lines of text and generates a set of bounding boxes. Script identification is done by identifying script per bounding box and there is one script per box. Text recognition is the core part of the OCR which recognizes text from image. The process is shown in Figure 5.

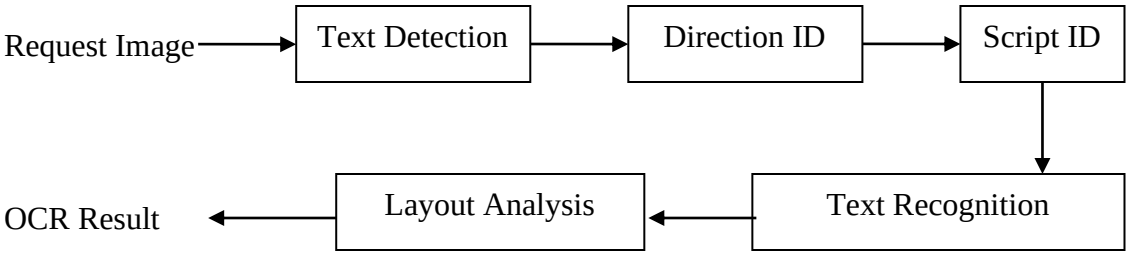

Figure 5: Goggle Vision API converting image to text

3.2.2 Tkinter:

Various options for the development of graphical user interfaces are provided by python. Tkinter is the standard GUI provided as a library for python. GUI applications can be created in a faster and easier way using Tkinter, and it also provides a prevailing object-oriented interface to the Tk GUI toolkit.

3.2.3 Google Speech to Text API:

Google cloud Speech to text aides the developers in the conversion of audio into text as it applies robust neural network models in a convenient API. It enables voice

command and control and transcribes audio. It is capable of processing real-time streaming or pre-recorded audio using Google's ML technology. The accuracy is unparalleled as the most advanced deep learning neural network algorithms are applied by Google. It streams text results, returning text as it is recognized from audio stored in a file and is capable of long-form audio.

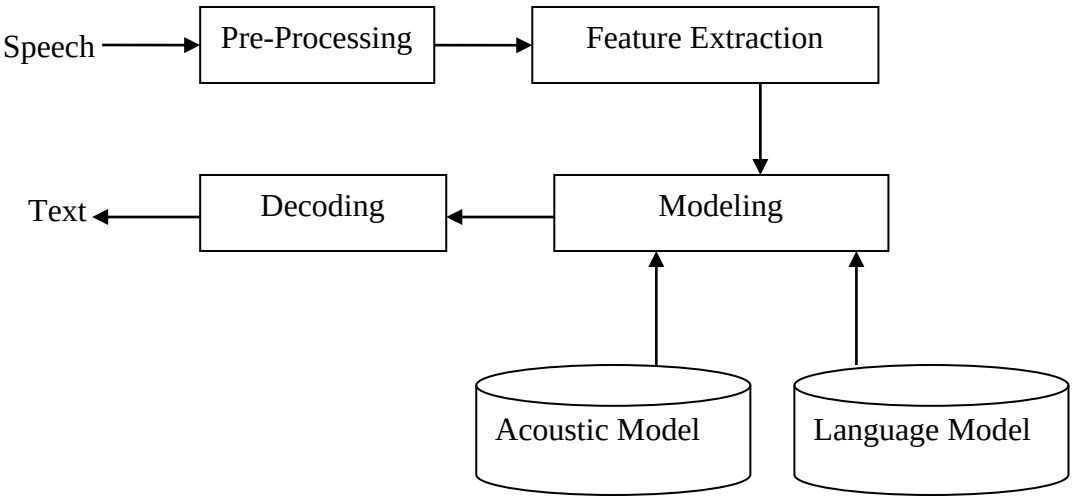

Figure 6: Google speech API converting speech to text

3.2.4 Google Speech to Text API:

Google Text to Speech API is one of the several APIs available in python to convert text to speech. It is commonly known as the gTTS API. It is an easy and efficient tool which converts entered text, into audio that can be saved as an mp3 file.

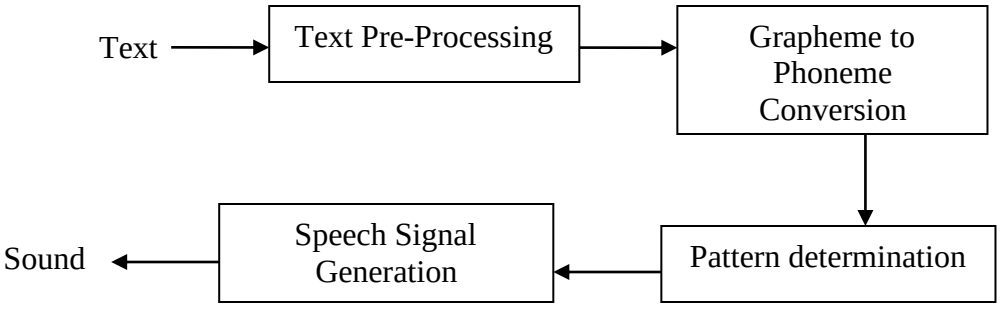

Figure 7: gTTS converting text to speech

3.2.5 Bitwise SSH:

Bitwise ssh is one of the advanced and flexible SFTP protocol. The bitwise ssh helps us to securely connect with raspberry pi and access all the resources of raspberry pi. In addition, the user can transfer the files from local host to raspberry pi; compile the programs; and provides a secure link for further connection.

4. DESIGN PLAN

4.1 Design approach

The figure 8 shows the outline of the model that is used in the project. The raspberry Pi is the support system of the device which connects the camera, microphone, speaker and LCD display. The device works for the visually impaired as the camera clicks a picture of the document and the output is in audio format through the speaker, audibly impaired as the microphone takes the spoken words as input and displays it as text on the LCD display, and for the vocally impaired as the user types the message in the LCD and the speakers gives the output as an audio.

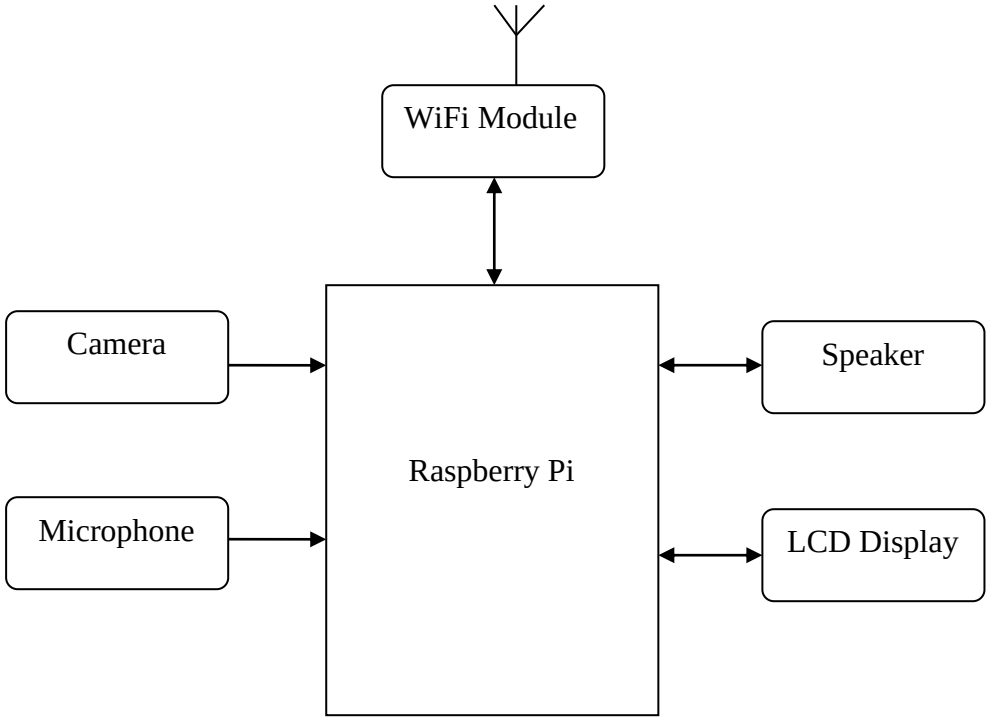

Figure 8: The system architecture with all main modules

The device has been created by formulating a unique design for assisting the differently abled people. It has been divided into three modules for enhancing the experience of the user with the device. The device consists of three modes and a three-way slider to change mode. Each mode is separately dedicated for the blind, deaf and dumb respectively in the device. The device is designed to make the user feel individualistic, self-reliant and self-sufficient. The

detailed design of the device is in the figure. The main component of the device is the raspberry PI. Figure 9 shows the flowchart and the flow procedure of the modules that have been explained below.

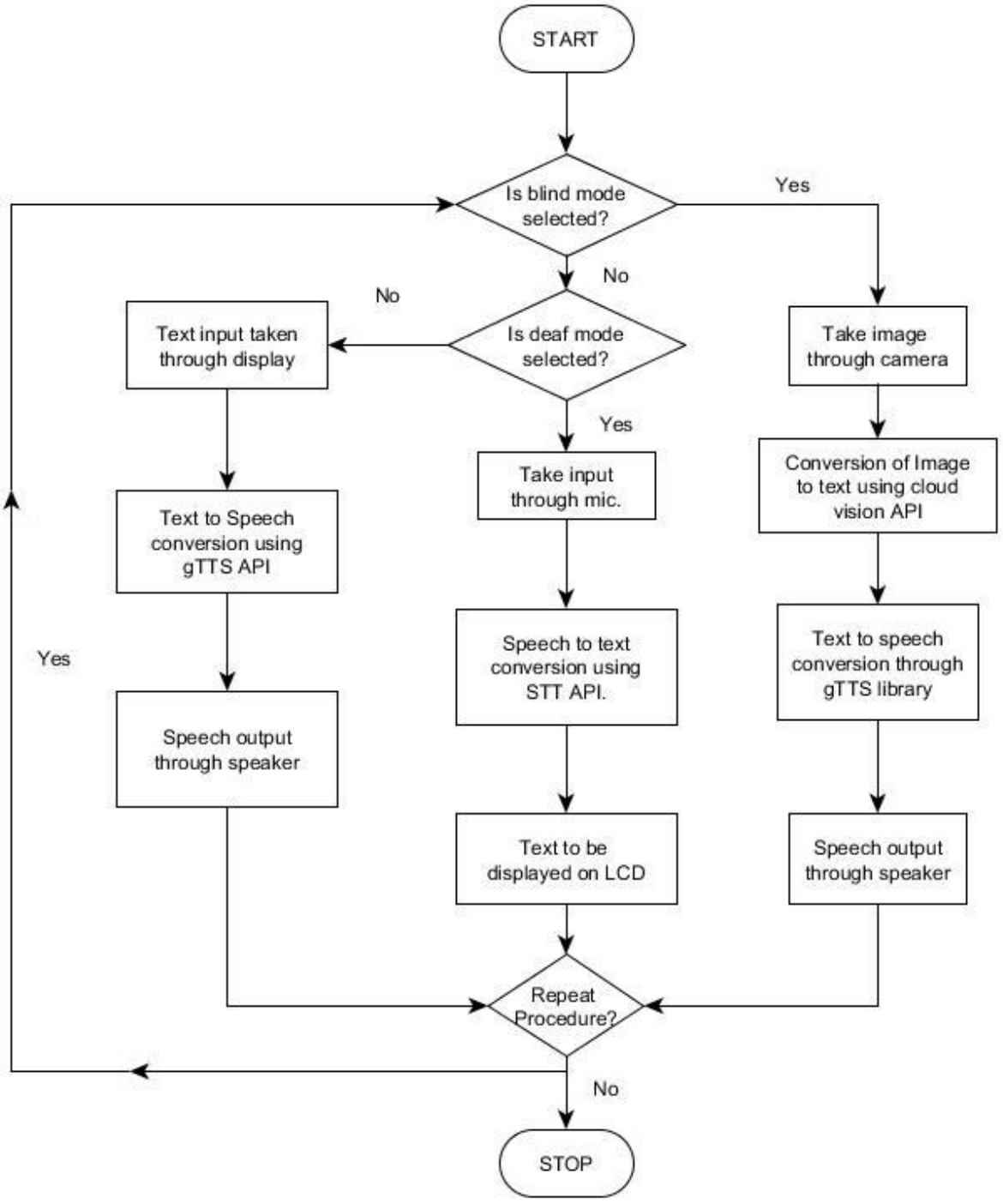

Figure 9: The flowchart of the modules

4.1.1 Blind Module:

The figure methodology of this module which consists of three steps.

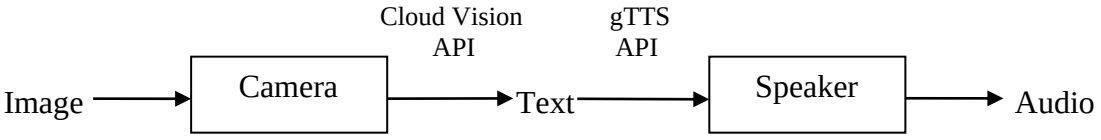

Figure 10: Working of Blind Module

1. For the module to work, the three-way slider is set to the blind mode. The camera connected to the raspberry Pi of the device takes a picture of the written document or book placed on the holder of the device.
2. The picture is saved in JPEG format and is passed to the Google Cloud Vision API to be converted to text where the API extracts the text to be converted.
3. The extracted text then gets converted into speech using the gTTS API and the required text is thus converted to the audio format.
4. This audio is given as an output by the high quality speaker connected to the Raspberry Pi and thus the device enables the visually impaired person to understand the written document or book through the audio.

4.1.2 Deaf module:

The audibly impaired can virtually hear using this device as it enables them to read, what is being spoken. The figure describes the respective procedure.

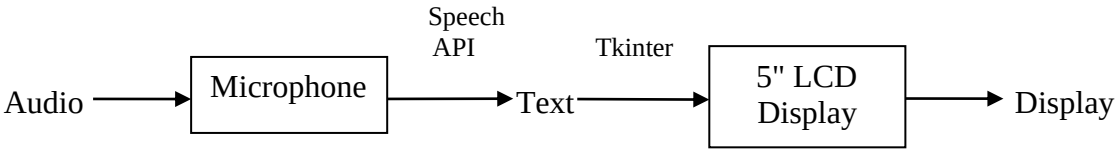

Figure 11: Working of Deaf Module

1. The three-way slider is set to the deaf mode. The audio or the words being spoken to the user, who in this case might be a deaf person, are recorded as input by the USB

Microphone connected to the Raspberry Pi of the device and is saved as a file in mp3 format.

2. This audio file is passed to the Google Speech API which converts the audio into text for the user to understand.
3. The converted text is then displayed on the 5 inch HDMI LCD screen available in the device, as a pop up window exclusively created using python tkinter for this module. This way the user understands everything that is being spoken to him quickly and efficiently. To change modes, the slider can be set accordingly.

4.1.3 Mute module:

This module makes the device handy for the vocally disabled as it enables them to vocalise words by typing it on the screen. The figure explains the methodology.

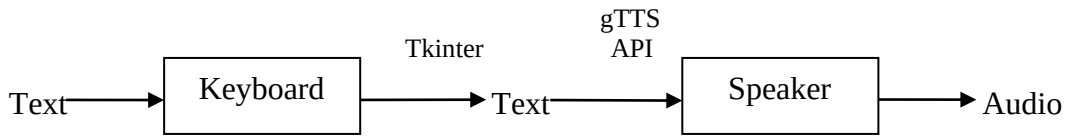

Figure 12: Working of Mute Module

1. When the three-way slider is used to set the device on the dumb mode, a pop up is displayed along with a customized keyboard which has been created using python tkinter, in it on the HDMI screen connected to the Raspberry Pi.
2. The user who possibly is vocally impaired can type whatever he wants to convey using the keyboard in the screen as text
3. The typed text is converted into audio format using the gTTS API and the audio file of the required text is obtained.
4. The high quality speaker connected to the Raspberry Pi in the device plays this audio file thus vocalising the message given by the impaired person.
5. Modes can be changed in the device according to the convenience of the user.

4.2 Codes and standards

4.2.1 IEEE 802.11

802.11n is a wireless-networking standard that uses multiple antennas to increase data rates. Sometimes referred to as MIMO, which stands for "multiple input and multiple outputs", it is an amendment to the IEEE 802.11 wireless-networking standard. Its purpose is to improve network throughput over the two previous standards—802.11a and 802.11g—with a significant increase in the maximum net data rate from 54 Mbit/s to 600 Mbit/s.

4.2.2 USB

USB, short for Universal Serial Bus, is an industry standard that was developed to define cables, connectors and protocols for connection, communication, and power supply between personal computers and their peripheral devices.

USB was designed to standardize the connection of computer peripherals to personal computers, both to communicate and to supply electric power. It has largely replaced a variety of earlier interfaces, such as serial ports and parallel ports, as well as separate power chargers for portable devices – and has become commonplace on a wide range of devices.

4.2.3 SSH

The SSH protocol is a method for secure remote login from one computer to another. It provides several alternative options for strong authentication, and it protects the communications security and integrity with strong encryption. It is a secure alternative to the non-protected login protocols and insecure file transfer methods.

The protocol works in the client-server model, which means that the connection is established by the SSH client connecting to the SSH server. The SSH client drives the connection setup process and uses public key cryptography to verify the identity of the SSH server. After the setup phase the SSH protocol uses strong symmetric encryption and hashing algorithms to ensure the privacy and integrity of the data that is exchanged between the client and server.

4.3 Constraints, Alternatives and Tradeoffs

Some of the constraints of this project are:

- Delay in the recognition of image.
  Solution: The use of faster and better processors like ARM-based processors.
- Delay in the switching of modules.
  Solution: The use of faster processors.

5. SCHEDULES, TASKS AND MILESTONES

The project was divided into modules and the milestones were set accordingly. The following are the lists of modules for the project that had to be completed:

1. TTS module
2. Image recognition module
3. STT module

Following are the Timeline and Gantt chart that display the important milestones according to the date they were completed:

5.1 Timeline

AT RISK	TASK NAME	SUB TASK NAME	STATUS	START DATE	END DATE	DURATION in days
☑	Meeting with guide		Completed	05-Dec-18	06-Dec-18	1
☑	Finalizing Title		Completed	06-Dec-18	18-Dec-18	12
☑		Literature Review	Completed	08-Dec-18	18-Dec-18	10
☑	Review 1		Completed	18-Dec-18	25-Jan-19	38
☑		Module 1 - for blind people	Completed	22-Dec-18	24-Jan-19	33
☑	Review 2		Completed	25-Jan-19	25-Feb-19	31
☑		Module 2- for deaf people	Completed	25-Jan-19	14-Feb-19	20
☑		Module 3 - for dumb people	Completed	11-Feb-19	10-Mar-19	27
		Paper/patent Process	Completed	20-Feb-19	15-Mar-19	23
☐	Final Review	Complete integration	Completed	10-Mar-19	01-Apr-19	22

Figure 13: Timeline of the project

5.2 Gantt chart

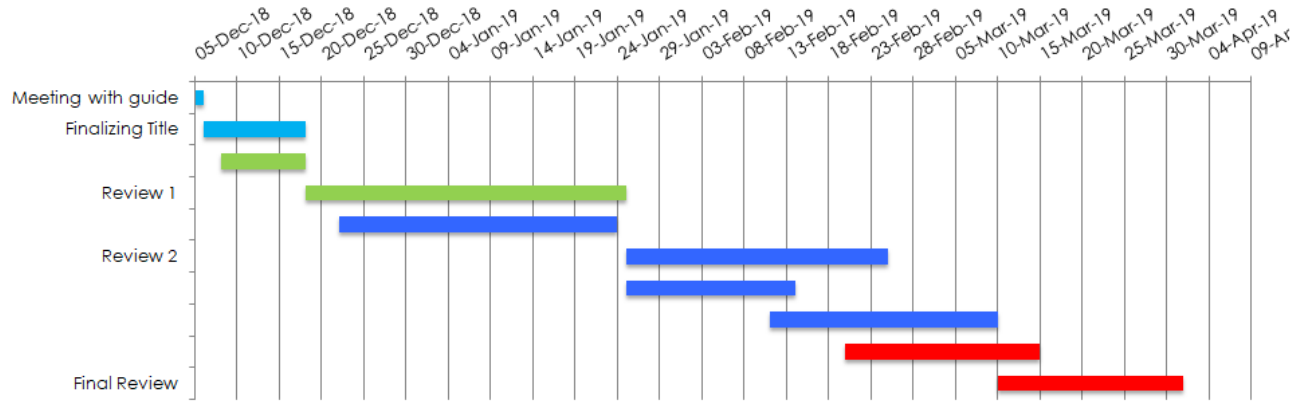

Figure 14: Gantt Chart of the project

6. PROJECT DEMONSTRATION

6.1 Project Views

A few screen captures of the project to provide a broader understanding of the project and a holistic view of the functionalities have been given below.

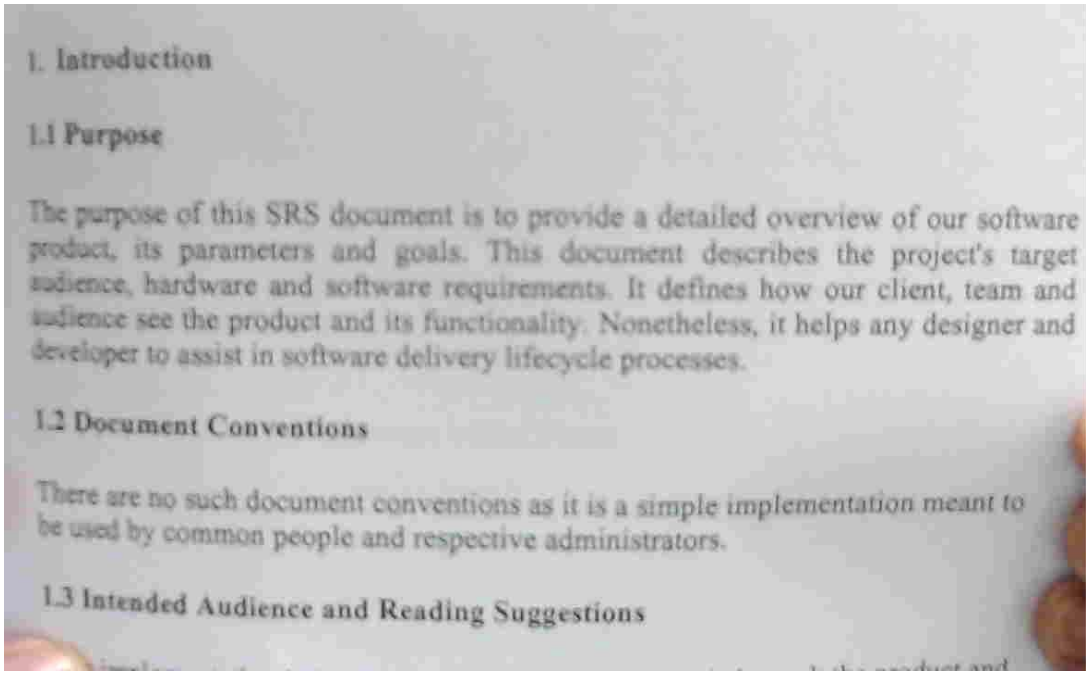

Figure 15: Image captured by the camera

Figure 15 shows the text that was taken as an image for the cloud vision API that was used in the testing for Module 1.

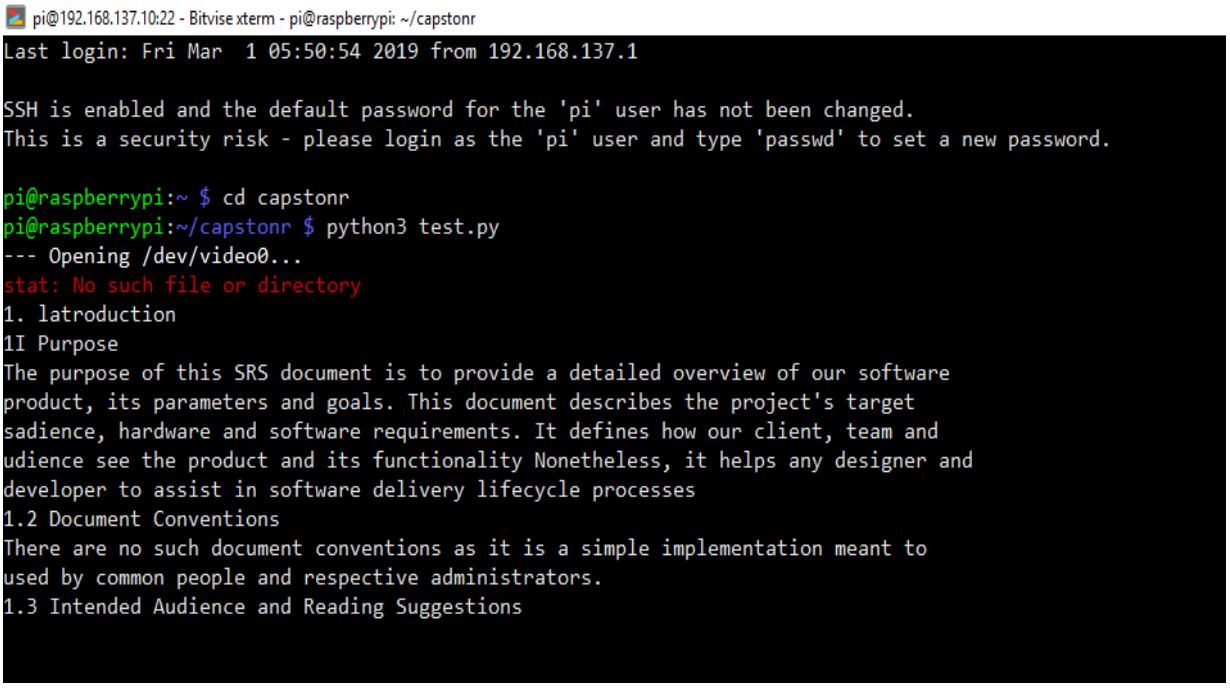

Figure 16: Resultant image to text conversion by the Google cloud vision API

Figure 16 shows the output of after Figure 15, the image that was taken for testing was converted to text by the cloud vision API. and Figure 17 shows the completion of and the output given through the speaker.

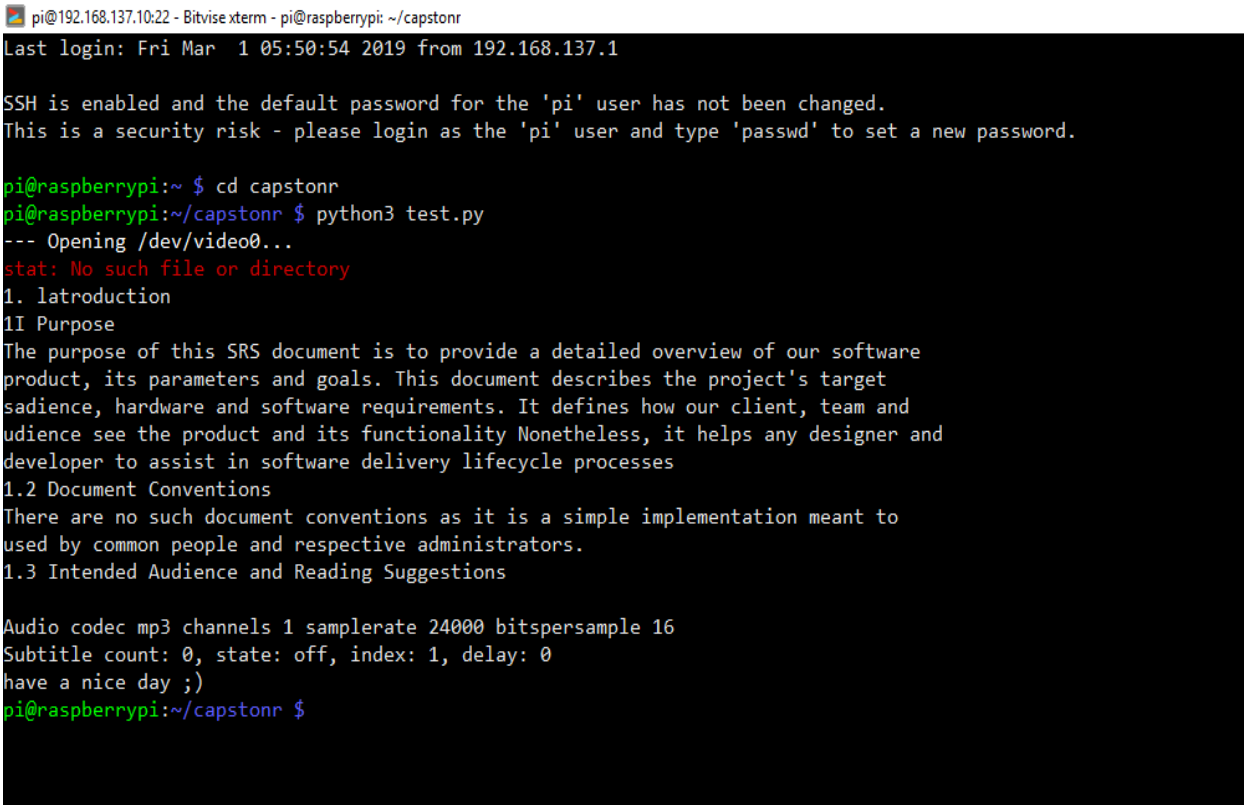

Figure 17: Conversion of gained text from image to audio

Figure 18 shows how the audibly impaired can read as the audio or spoken words “ Hello this is Muktak doing testing for module 2” are identified as and converted to text.

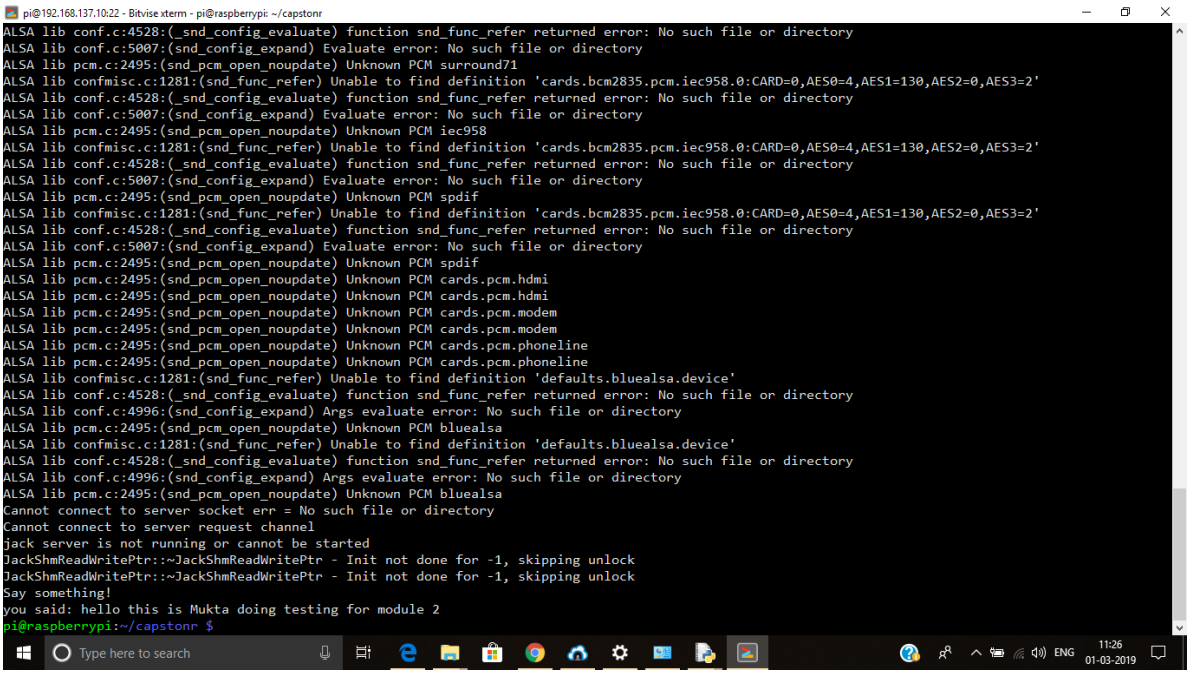

Figure 18: Conversion of Audio to text

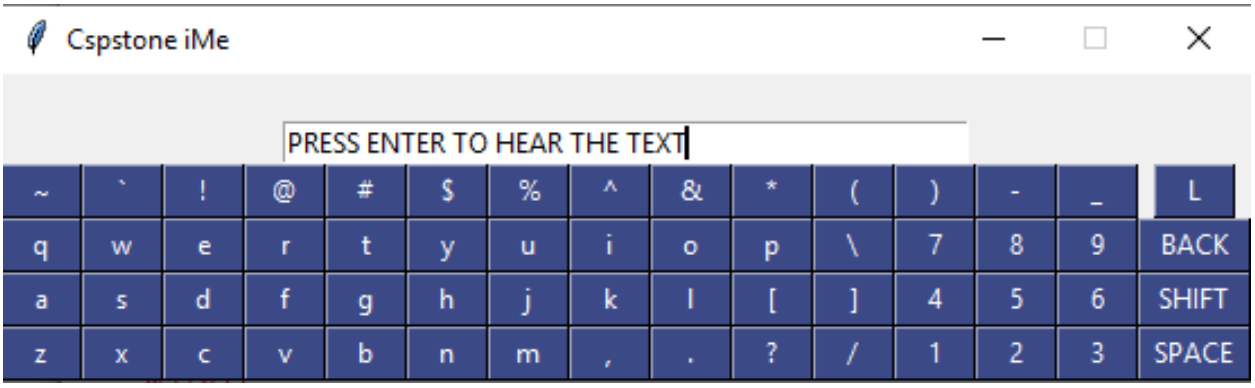

Figure 19: On-screen Keyboard (text to speech)

Figure 19 shows the on screen keyboard that takes the text input for the usage of the audibly impaired.

7. MARKETING AND COST ANALYSIS

7.1 Market Analysis

The market analysis for this project was done and it was observed that there existed separate solutions for these problems but we did not find one single product for the same. The analysis showed that the products were either for a single problem or if they were targeting two disabilities then they were full of constraints. Moreover most of the products were research based prototypes and not actual marketable products . We are using Cloud Vision for Image to text, Speech to text and text to speech Google APIs which are reliable and have a faster processing speed. Since these are Google products they are secure and reliable too. They are also bound to increase customer satisfaction and trust. We are also using Raspberry pi 3 which has a higher processing speed itself. All other components except for camera also have a very less delay. Since we are using hardware like Camera, microphone and LCD, which not only don’t harm the environment. The components used are also recyclable so the ecological impact of the project is minimal. The main component of the project is a raspberry pi 3 which is a processor that can be used with versatility. With just the change of a memory card the whole OS of the processor can be changed making this product highly recyclable.

7.2 Cost Analysis

	Parts	Cost
1.	Raspberry pi	Rs 2600.00
2.	Logitech C310HD camera	Rs 1425.00
3.	Wave share 5 inch display	Rs 1600.00
4.	USB microphone	Rs 700.00
6.	Speaker	Rs 700.00

Total Parts Rs 7025.00

Total Cost per Unit Rs 7025.00

Table 1: Cost Analysis

8. SUMMARY

Through this project, an unprecedented prototype has been created to aid the visually, vocally and audibly disabled. This project not just focuses on empowering and facilitating the differently abled, it is also compact and a resource saver. The overall cost has been cut down by eliminating Braille books and the energy spent in understanding them. It is a less costly solution, as all the components used in the device are cost effective and efficient. The latest and most trending technology makes this device portable, adaptable and convenient. The device proposed in this paper can be a major help in solving a few of the many challenges faced by the differently abled.

To further extend the project, the device can be made more compact and wearable to make it easy for the user to use. One more future work can be to extend the project by adding cloud storage and automatically storing the image and text to audio files in the cloud storage to be retrieved by the push of a button. Upon doing this the person will have easier access to usually spoken sentences and would not have to type again and again to be converted by the gTTS API.

9. REFERENCES

[1] N. K., S. P. and S. K., Assistive Device for Blind, Deaf and Dumb People using Raspberry-pi, vol. 3, issue-6. Imperial Journal of Interdisciplinary Research (IJIR), 2017 [Online]. Available: https://www.onlinejournal.in/IJIRV3I6/048.pdf.

[2] L. González-Delgado, L. Serpa-Andrade, K. Calle-Urgiléz, A. Guzhñay-Lucero, V. Robles-Bykbaev and M. Mena-Salcedo, "A low-cost wearable support system for visually disabled people," 2016 IEEE International Autumn Meeting on Power, Electronics and Computing (ROPEC), Ixtapa, 2016, pp. 1-5.

doi: 10.1109/ROPEC.2016.7830606

[3] Anusha Bhargava, Karthik V. Nath, Pritish Sachdeva & Monil Samel (2015), International Journal of Current Engineering and Technology, E-ISSN 2277– 4106, P-ISSN 2347– 5161

[4] J. Sanchez and T. Hassler, "AudioMUD: A Multiuser Virtual Environment for Blind People," in IEEE Transactions on Neural Systems and Rehabilitation Engineering, vol. 15, no. 1, pp. 16-22, March 2007.

doi: 10.1109/TNSRE.2007.891404

[5] M. Lumbreras and J. Sánchez, “Interactive 3-D sound hyperstories for blind children,” in Proc. ACM-CHI ’99, Pittsburgh, PA, 1999, pp. 318–325.

[6] R. McCrindle and D. Symons, “Audio space invaders,” in Proc. ICDVRAT 2000, Alghero, Sadinia, Italy, Sep. 23–25, 2000, pp. 59–65.

[7] T. Westin, “Game accessibility case study: Terraformers-Real-time 3-D graphic game,” in Proc. ICDVRAT 2004, Oxford, UK, 2004, pp. 120–128.

[8] Y. H. Lee and G. Medioni, “Rgb-d camera based wearable navigation system for the visually impaired,” Computer Vision and Image Understanding, vol. 149, pp. 3–20, 2016

[9] J. Bajo, M. A. Sanchez, V. Alonso, R. Berj ´ on, J. A. Fraile, and J. M. ´ Corchado, “A distributed architecture for facilitating the integration of blind musicians in symphonic orchestras,” Expert Systems with Applications, vol. 37, no. 12, pp. 8508–8515, 2010.

[10] L. Monti and G. Delnevo, "On improving GlovePi: Towards a many-to-many communication among deaf-blind users," 2018 15th IEEE Annual Consumer

Communications & Networking Conference (CCNC), Las Vegas, NV, 2018, pp. 1-5.doi: 10.1109/CCNC.2018.8319236

[11] R. Rastogi, S. Mittal and S. Agarwal, "A novel approach for communication among Blind, Deaf and Dumb people," 2015 2nd International Conference on Computing for Sustainable Global Development (INDIACom), New Delhi, 2015, pp. 605-610.

[12] A. M. Hassan, A. H. Bushra, O. A. Hamed and L. M. Ahmed, "Designing a verbal deaf talker system using mouth gestures," 2018 International Conference on Computer, Control, Electrical, and Electronics Engineering (ICCCEEE), Khartoum, 2018, pp. 1-4.

doi: 10.1109/ICCCEEE.2018.8515838