Journal of Physics: Conference Series

PAPER • **OPEN ACCESS**

Smart Glass with Multi-Functionalities for Assisting Visually Impaired People

Smart Glass with Multi-Functionalities for Assisting Visually Impaired People

G Sudharshan, S Sowdeshwar and M Jagannath

School of Electronics Engineering, Vellore Institute of Technology (VIT) Chennai, Tamil Nadu 600127, India

*E-mail: jagan.faith@gmail.com

Abstract. Smart glass is a device that is obliging for visually impaired people. It becomes difficult for impaired people to recognize objects in front of them, study, etc. The objective of the work is to design and develop a device that is obliging for visually impaired people. Though, there are many approaches to help them, the major disadvantage of some of the products that are currently present in the market is that they are very costly and bulky. Also, some of the proposed systems either perform only object recognition or only help in reading the text. Almost all proposed systems did not help the visually impaired people to write or to have a hardcopy in the form of braille code which then helps them to read offline. To overcome the disadvantages mentioned above the paper discusses various techniques- Seeing mode, which can recognize 550 classes of objects along with object's position (Left, Right or Centre). Moreover, in the reading mode, it recognizes the text in front of the person which is then read out loud to the user along with creating a copy of the braille script. Further, in the writing mode, speech is converted to a file containing braille codes.
Keywords: Braille Code, Computer Vision, Optical Character Recognition, Smart Glass.

1. Introduction

According to Investigative Ophthalmology & Visual Science 2020 [1], around the world, in 2020, there would be 49.1 million blind people, 221.4 million with moderate vision impairment, and 33.6 million with severe visual impairment. While there is no cure for blindness, there must be a way to assist them. The high proportion of vision-impaired people are unable to complete their studies. Special schools for disabled children are not broadly accessible, and the majority of them are expensive and exclusive.

As a result, their only option was to study at home and learn the basics from their family. This education does not help with learning concepts, and as a result, it is unable to compete with others. There are various levels of need, and not all of them necessitate the use of special facilities and schools. People with vision problems, for example, can learn with other students provided they are in the right atmosphere. To address this problem, computer vision technology can be employed to create specific aids that allow visually impaired persons to live as comfortably as possible [2].

The basic idea is to help visually impaired people to read the text that is present in front of them, recognize the objects that are present in front of them, and to help them write in braille code. The main benefit of the product is that visually impaired people can use the product to learn about their surroundings better. This reduces the need for other people to help them do these basic works and they can independent for these activities. Another benefit of the proposed product is that it also mentions the position of the object in the frame of view. It tells the person whether the object is present in the left, center, or right. This helps them to be spatially aware and helps them if they want to grab that particular object. The key benefit of this proposed idea is writing and saving every text that was detected in front of the person as braille code and saving that file in the cloud. This is also used in writing mode. The person can write in braille code and that file is then sent to the cloud. A braille file (Grade 1) is generated and stored in the Raspberry pi. Which can then be copied and passed to a braille embosser for hard copy of the file. This helps them to have a hard copy of the book or even a page. The text is converted to braille as well as the same text is converted to speech signals. This helps the impaired person in 2 ways: (i) Reading out aloud the book in real time to make them aware of what they are studying currently, and (ii) Converting the text to braille, which will help them in having a backup when studying without the glasses. This also helps in faster conversion of a book to a book containing braille code which helps visually impaired students in college or schools to read and study.

The proposed device uses a Raspberry pi Zero W board. It is one of the smallest Raspberry pi boards with minimal specifications. This board was chosen because of its small size which is suitable for our motivation for making the model more portable and easily wearable. Raspberry pi zero w comes with additional features like Bluetooth Low Energy (BLE). It also has a 1GHz, single-core processor, 512MB RAM, Micro USB power, CSI camera connector, and Mini HDMI. It requires a microSD with an operating system supported by Raspberry pi and a 5V supply to the board. The operating system used is Raspbian OS. The 5MP Raspberry pi camera captures the person's image. Raspberry pi and the camera are put together inside a case which is then pasted in front of the glasses. Additionally, an ultrasonic sensor is also pasted in front of the glasses along with buttons on the left side of the glasses and a 5v 2 A power bank on the right side of the glasses.

2. Literature Review

Hassan and Tang [3] proposed a system that uses Tesseract OCR on a Raspberry pi 2 system which detects the red boundaries and the orientation of frame around an object which contains the text. The image is enhanced with contrast enhancement (histogram matching approach), noise filtering, and morphological modifications after the text region has been localized and cropped. AlSaid et al. [4] proposed a system that takes the help of text detection methods and Deep Learning [5] concepts to find text in a localized area. The text that was recognized by the system is then processed by Google's Text To Speech which is then used as an audio signal. The system used was Raspberry pi 3B+.

Satani et al. [6] built a model which detects around 6 classes of animals like cats, cows, etc. using a CNN model. When an object is detected, an internal thread is triggered to alarm the user saying "Thing ahead". The main disadvantage of these types of systems is that they can detect only a few different classes and it does not say the position or the type of the object that is present in front of the person. Additional feature such as facial recognition is also built as a part of the system. Miah and Hussain [7] have proposed a system to address challenges while travelling. The model measures the distance between the obstacle and the user. The information obtained from the surroundings is transmitted to the person via an earphone. To acquire online data, a SIM900A GSM/GPRS module is used. Mangayarkarasi et al. [8] proposed a prototype that includes several functionalities like news headlines, unread emails, dominant color detection, current date and time, current location, weather information, and a calculator which gets input from a handheld device. It also performs OCR using tesseract API. The major limitation of this proposed system is that it does not convert the text to braille or perform object detection and it also uses a bulky device compared to Raspberry pi Zero W.

There are some of the products which are already readily available in the market. One of the products is eSight3 [9] which is developed by CNETs. It contains a camera for capturing images and videos. This is just an aid for low vision people and helps them avoid surgery. This does not help blind people. Another such product is Oton glasses [10] which supports dyslexic people and helps in converting images to words and then to audio. This supports only for people having reading difficulty and does not have any support for blind people. Eyesynth [11] is also one such product that contains 3D cameras to turn the scene that is present in front of the person into a sound signal providing information on position, size, and shape. The major drawback is this product is costly (575.78$) and it only identifies the directions and objects.

Finally, there is Google's glass [12] where voice commands can be used to communicate with the internet. With the word lens app, it captures photographs and motion pictures, receives directions, sends information, makes voice calls, and translates instantaneous. But this product is more expensive (1,349.99$) than Eyesynth and it was not built for blind people.

3. Materials and Methods

Figure 1 shows the work flow of the proposed smart glass device. The user has the flexibility to choose the mode according to their choice. One button is used to choose the mode. By default, the reading mode is started. If the button is pressed once, it moves on to seeing mode. If it is pressed again, it switches to writing mode. When the button while the system is processing, it first completes the process and then switches the modes. For a demonstration of the proposed system, daily use objects are used for the seeing mode, books or newspapers for reading mode, and the audio recorder by us for the writing mode. Raspberry pi interacts with google drive to upload the braille codes or to get the recently uploaded audio file after the writing mode is chosen. The process starts by welcoming the user by saying the current date and time. By default, mode1 (Seeing Mode) is chosen. The Raspberry pi camera turns on and starts capturing the image frame by frame. Each frame that is captured is then sent to google vision API for detecting objects. The result of the objects that are present in the image is then collected. The position of the object is calculated by finding the center of the bounding boxes of each image that are detected. Then it will classify whether the object is present in the left, right, or in front by checking the center points.

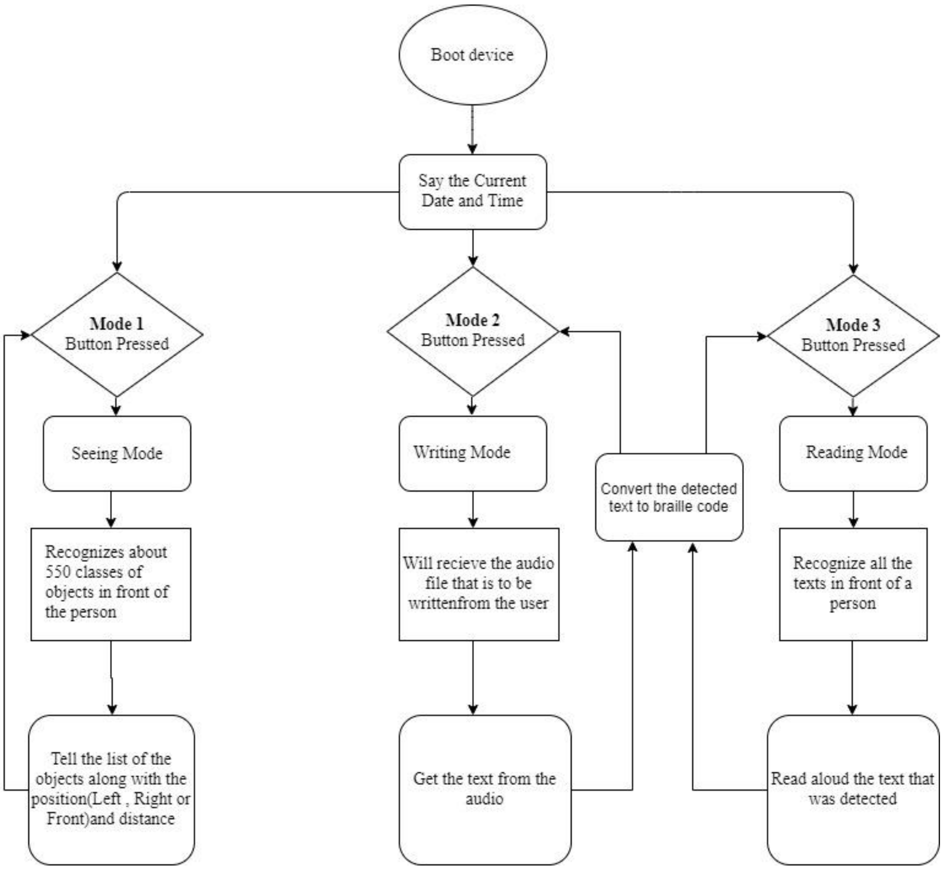


Figure 1. Work flow of the proposed smart glass device.

Each frame is divided into 3 parts. If the center point is present in the first 1/3rd of the frame, it will classify it as left. If the center point is present in the last 1/3rd of the frame it will classify it as right. Any other points are classified as in front of the person. Hence, with all these results the device says the object name that is present along with the position of the object. When the user presses the button, the device automatically switches the mode from 1 to 2. This mode helps visually impaired people to write in braille code. To do that, the user has to upload an audio file to their google drive. When the button is pressed, the device automatically downloads the latest file that is uploaded and uses that audio file to get the text that is said in that file. The text is then converted to braille code which is then uploaded back to google drive automatically as a braille code file.

If the user pushes the button again, the device switches mode from 2 to 3. This mode helps visually impaired people to read. Similar to seeing mode, the pi camera begins taking pictures. With the help of google vision API, the text that is present in front of the person is spotted. If no text is detected the device says the person to either turn on the light or stay stable for a while to capture the image. The text which was detected is then converted to braille code and stored as a file. This file is then uploaded to google drive automatically.

3.1. Hardware System Design

The main hardware components that were used in building this project were Raspberry pi, camera, ultrasonic sensor, and buttons. The headphones are connected via Bluetooth where the Raspberry pi automatically searches and connects to the Bluetooth headsets. The camera that is used is Raspberry Pi standard camera module version 1.3. The Dimensions of the camera are 25mm x 20mm x 9mm. It has a 5 MP resolution with 2592 x 1944pixel static images. It can capture the video in 1080p30 and this module can only take photographs and videos; it cannot record sound. The module is connected to the camera serial interface of Raspberry pi. The distance between the object and camera depends the object itself, and how visible the object is, in the image taken by the Raspberry pi camera. Moreover, it is weight is around 3g which makes it perfect for our application. Raspberry pi zero w does not come with GPIO pins soldered on the board. The developer is required to solder it. The GPIO pins are soldered on the backside of the board so that there would be enough space for the camera to fit inside the case. The top case has a hole where the camera's lens can easily fit through and the bottom case has space for GPIO pins if the developer needs to connect any external hardware to the Raspberry pi. The ultrasonic sensor HC-SR04 emits an ultrasonic sound of frequency 40 kHz. This ultrasound signal travels through the air and bounces of objects that are in its path back to the module. It has 4 pins which are to be connected to the GPIO pins. A lightweight power bank is chosen for this application. The case which contains the Raspberry pi and its camera is attached in front of the glass along with the ultrasonic sensor. The breadboard which contains the buttons is connected to the left of the glass. The power bank is attached to the right of the glasses.

3.2. Software System Design

The first step is to boot the Raspberry pi zero w with Raspbian OS. All the required packages are installed through the command prompt using the pip3 command. All the required modules and functions are built using python 3.7. Initializing and capturing pictures from the Raspberry pi camera module is done using the pi Camera module. This module helps to capture images at a required instant. Capturing of images takes place inside a while loop to capture and process images continuously.

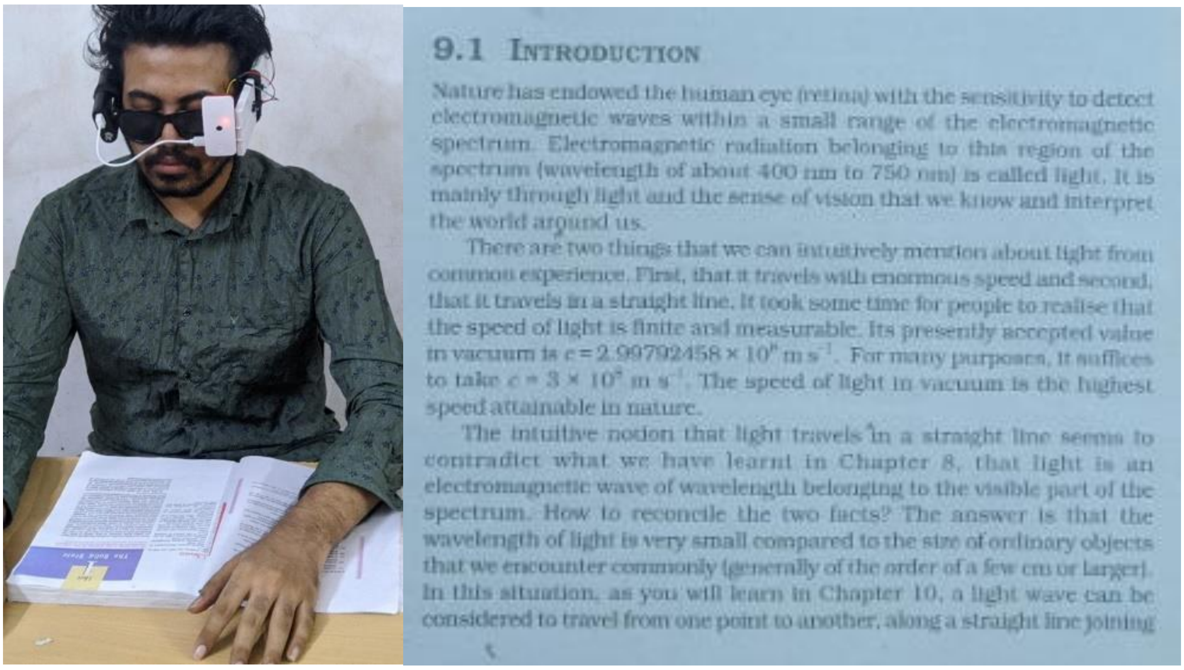
For reading mode and seeing mode, the image that is captured is processed and sent to the google vision application programming interface (API). This applications helps us in classifying about 550+ classes and detecting text in an object with just limited power consumption. Google Vision API was chosen because of the less amount of RAM that is present in Raspberry pi zero w and using such powerful online models is the way for this application. The returned result from the API is then converted to speech and is stored as an mp3 file in the system. This stored mp3 file is then played using the pygame module which lets us play any mp3 file format for any particular amount of time.

Additionally, for the reading mode, the result from the API is converted to braille code. This conversion is done using a predefined symbol for a particular character. The character that is processed is converted to their respective braille symbol and is stored as a string. For writing mode, the latest uploaded mp3 file is automatically found and is processed to convert that file to text. This text is then converted to braille code as said above. All these strings containing braille code are then stored as a .txt file in the system and are then automatically uploaded to google drive using the drive API.

4. Results and Discussion

4.1. Reading Mode Output

A sample book was used to show it as an input to the Raspberry pi (Figure 2). A page from this book was chosen and was placed in front of the camera. The device first welcomes the user with a simple date and time message sent through their headphones.



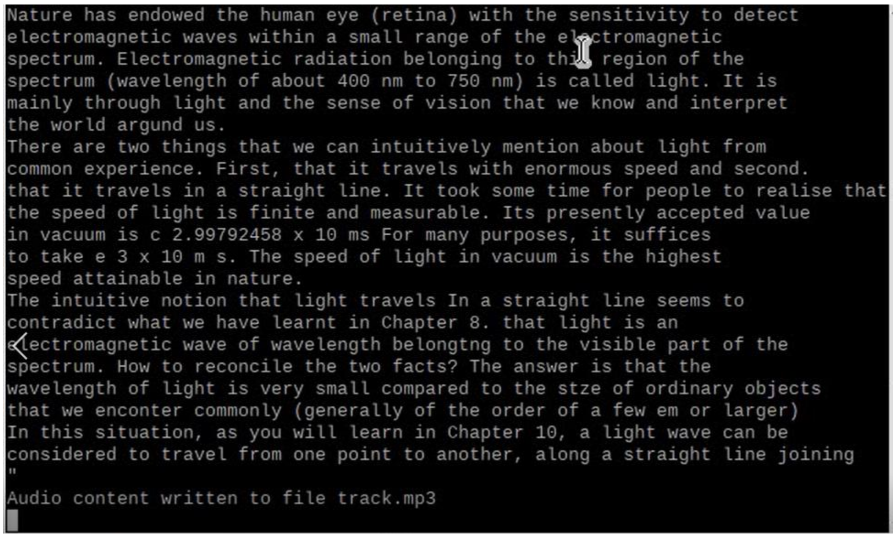
9.1 INTRODUCTION

Nature has endowed the human eye (retina) with the sensitivity to detect electromagnetic waves within a small range of the electromagnetic spectrum. Electromagnetic radiation belonging to this region of the spectrum (wavelength of about 400 nm to 750 nm) is called light. It is mainly through light and the sense of vision that we know and interpret the world around us.

There are two things that we can intuitively mention about light from common experience. First, that it travels with enormous speed and second, that it travels in a straight line. It took some time for people to realise that the speed of light is finite and measurable. Its presently accepted value in vacuum is $c = 2.99792458 \times 10^8$ m s^{-1}. For many purposes, it suffices to take $c = 3 \times 10^8$ m s^{-1}. The speed of light in vacuum is the highest speed attainable in nature.

The intuitive notion that light travels in a straight line seems to contradict what we have learnt in Chapter 8, that light is an electromagnetic wave of wavelength belonging to the visible part of the spectrum. How to reconcile the two facts? The answer is that the wavelength of light is very small compared to the size of ordinary objects that we encounter commonly (generally of the order of a few cm or larger). In this situation, as you will learn in Chapter 10, a light wave can be considered to travel from one point to another, along a straight line joining

Figure 2. Sample text from a book that is captured by the device.



```
Nature has endowed the human eye (retina) with the sensitivity to detect
electromagnetic waves within a small range of the electromagnetic
spectrum. Electromagnetic radiation belonging to this region of the
spectrum (wavelength of about 400 nm to 750 nm) is called light. It is
mainly through light and the sense of vision that we know and interpret
the world argund us.
There are two things that we can intuitively mention about light from
common experience. First, that it travels with enormous speed and second.
that it travels in a straight line. It took some time for people to realise that
the speed of light is finite and measurable. Its presently accepted value
in vacuum is c 2.99792458 x 10 ms For many purposes, it suffices
to take e 3 x 10 m s. The speed of light in vacuum is the highest
speed attainable in nature.
The intuitive notion that light travels In a straight line seems to
contradict what we have learnt in Chapter 8. that light is an
electromagnetic wave of wavelength belongtng to the visible part of the
spectrum. How to reconcile the two facts? The answer is that the
wavelength of light is very small compared to the stze of ordinary objects
that we enconter commonly (generally of the order of a few em or larger)
In this situation, as you will learn in Chapter 10, a light wave can be
considered to travel from one point to another, along a straight line joining
"
Audio content written to file track.mp3
```

Figure 3. Detected text from the book presented in front of the person.

By default, the system starts detecting text in front of a person in reading mode and generates braille code for the detected text as shown in Figure 3 and Figure 4.

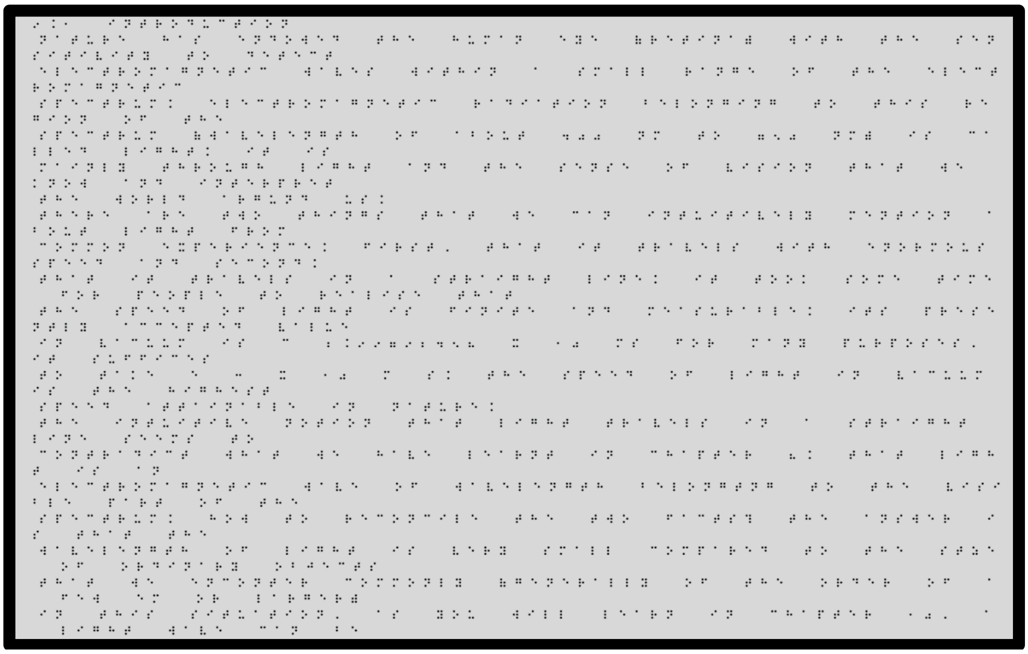


Figure 4. Braille code file generated from the text that was detected.

4.2. Seeing Mode Output

Some daily use sample objects were presented which can be used daily by visually impaired people as shown in Figure 5. These objects were then kept in front of the person to test the detection power of the Raspberry pi. After the button is pressed, the mode changes from reading to seeing mode, and in this, it detects all the objects in the range and their location (left, right or front) as shown in Figure 6.

Figure 5. Sample objects presented in front of the person.

```
Seeing Mode
Number of objects found: 2
Packaged goods in the left and Tubed packaged goods in the center
Audio content written to file track2.mp3
Seeing Mode
Number of objects found: 3
Sunglasses in the right , Packaged goods in the center and Packaged goods in the
 left
Audio content written to file track2.mp3
```

Figure 6. Objects that were detected in front of the person.

4.3. Writing Mode Output

Figure 7 shows a recording.mp3 file that is the latest file uploaded which contains the audio that is to be written in the braille code file. When the button is pressed again, seeing mode changes to writing mode, where it downloads and processes the latest file uploaded and converts audio to text. This text is then converted to braille code and is uploaded to the cloud as shown in Figure 8.

My Drive > Speech files

Name	Owner	Last modified	File size
track (1).mp3	me	Mar 9, 2021 me	96 KB
track (5).mp3	me	Apr 18, 2021 me	129 KB
track (4).mp3	me	Apr 18, 2021 me	129 KB
recording.mp3	me	Apr 20, 2021 me	65 KB

Figure 7. Content to be written is presented as an audio file and uploaded to the drive.

```
Seeing Mode
Number of objects found: 2
Person in the left and Glasses in the left
Audio content written to file track2.mp3
Pressed
Writing Mode
2021-04-20T16:29:52.000Z recording.mp3
I am currently testing the writing mode of my project
Reading Mode
```

Figure 8. Detected text and its conversion to braille code.

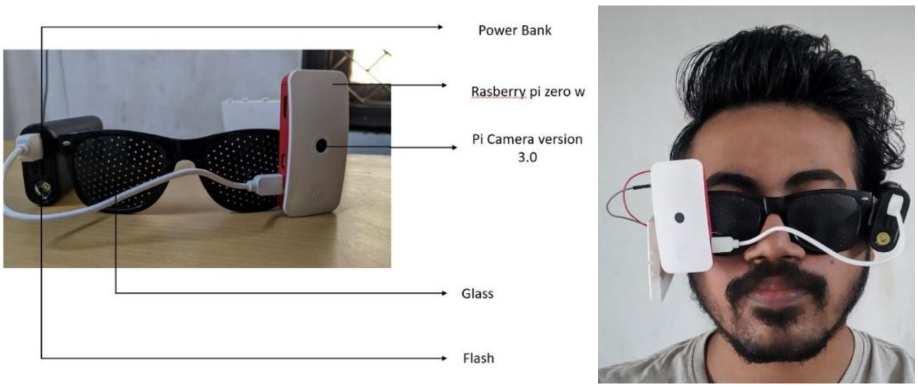


Figure 9. Proposed smart glass for assisting visually impaired people.

Figure 9 shows the final product of our model which contains the main component of Raspberry pi along with its camera inside a case, placed in front of the glasses. It contains a button on the side that helps visually impaired people to choose modes. It also contains a power bank for the power supply to the Raspberry pi along with a torchlight which can be used when there is no light in a room.

5. Conclusion

A new model of smart glass was developed for visually impaired people which contains three different modes: seeing, reading, and writing. In the seeing mode, 550 classes of objects were detected along with the object's position (Left, Right, or Centre). Moreover, in the reading mode, the text in front of the person was recognized and reclaimed (or read out loud). Additionally, in the writing mode, the speech was converted to a file containing braille codes. The cost was reduced to a great extent and the features of the smart glass are to a substantial extent.

Ethical Statement

Authors followed the ethical standards in developing the proposed device according to the guidelines of the institution of the authors. No copyright materials have been used in the manuscript.

Acknowledgements

The authors wish to acknowledge the volunteers for their valuable assistance and the anonymous reviewers for the invaluable inputs in improving the manuscript.

References

[1] Bourne R R A, Adelson J, Flaxman S, Briant P, Bottone M, Vos T, Naidoo K, Braithwaite T, Cicinelli M, Jonas J, Limburg H, Resnikoff S, Silvester A, Nangia V and Taylor H R 2020 Global Prevalence of Blindness and Distance and Near Vision Impairment in 2020: progress towards the Vision 2020 targets and what the future holds. *Investigative Ophthalmology & Visual Science*, **61** (7), p.2317.

[2] Lin B S, Lee C C and Chiang P Y 2017 Simple Smartphone-Based Guiding System for Visually Impaired People. *Sensors (Basel)* **17** (6), 1371.

[3] Hassan E A and Tang T B 2016 Smart Glasses for the Visually Impaired People. In: K Miesenberger, C Bühler, P Penaz (Eds.), *Computers Helping People with Special Needs*. 15th Int. Conf., ICCHP 2016, Linz, Austria, July 13-15, 2016, Proc. , Part II, Lecture Notes in Computer Science, **9759**, pp. 579-582, Springer, Cham.

[4] AlSaid H, AlKhatib L, AlOraidh A, AlHaidar S and Bashar A 2019 Deep Learning Assisted Smart Glasses as Educational Aid for Visually Challenged Students. *Proc. of the 2nd Int. Conf. on New Trends in Computing Sciences*, pp. 1-6.

[5] Lin J Y, Chiang C L, Wu M J, Yao C C and Chen M C 2020 Smart Glasses Application System for Visually Impaired People Based on Deep Learning, *Proc. of the Indo–Taiwan 2nd Int. Conf. on Computing, Analytics and Networks (Indo-Taiwan ICAN)*, pp 1-4.

[6] Satani N, Patel S and Patel S 2020 AI-Powered Glasses for Visually Impaired Person. *Int. Journal of Recent Technology and Engineering* **9** (2), pp. 416-421.

[7] Miah M R and Hussain M S 2018 A Unique Smart Eye Glass for Visually Impaired People. *Proc. of the Int. Conf. on Advancement in Electrical and Electronic Engineering*, pp. 1-4.

[8] Mangayarkarasi P, Khan M S, Sunil K A and Pramod S N 2020 Smart Vision System for Visually Impaired People. *Int. Research Journal of Engineering and Technology* **7** (4), pp. 943-946.

[9] eSight. 2017. eSight3. Website: https://esighteyewear.com/int, Accessed 2022/01.

[10] K. Shimakage, 2014. Oton Glasses. Website: https://www.jamesdysonaward.org/en GB/2016/project/oton-glass/, Accessed 2022/01.
[11] Eyesynth. 2018. Home - Eyesynth. Website: https://eyesynth.com/?lang=en, Accessed 2022/01.
[12] Google. 2013. GoogleGlass. Website: https://www.google.com/glass/start/, Accessed 2022/01.