Smart Goggles for visually impaired

[1]Alfina Sulfikar B, [2]Akshaya Raghu, [3]Vishnu Vijayan, [4]Akhil A.G.

[1, 2, 3] Engineering Students, [4]Assistant Professor
[1,2,3,4] Dept. Of Electrical and Electronics Engineering
[1,2,3,4] UKF College of Engineering and Technology

Abstract: Human beings are blessed with five senses to experience the world, vision, auditory, olfactory, gustatory, and touch. These senses rendered humanity unpredictable. Of the five, sight is considered the most important. Not all of us have the blessing of vision. So, the aim was to develop goggles for these peoples. We have used machine learning and deep learning for the analysis and interpretation of texts and environments. Using real time object detection and text reading, the visually impaired person can get information about the surroundings. The world of science is growing minute by minute and, it is our responsibility to extend the benefits of science for the advancement of humanity, especially for those who need special care and support. While many inventions exist in this field, many are practically ineffective. So, it is high time to give more light to this area.

I. INTRODUCTION

According to the World Health Organization (WHO) global report, at least 2.2 billion people have near- and far-sighted disabilities. The majority of individuals with vision impairment and blindness are over the age of fifty years. However, vision loss can occur in individuals of all ages. People with visual impairments find it difficult to communicate effectively not only with ordinary people but also with their surroundings. Visually impaired individuals may have difficulty interacting with their environment because it can be difficult to recognize where someone is and to move from one location to another. Movements can become small and limited, resulting in less contact and interaction with the surrounding world. To minimize these restrictions, blind individuals rely primarily on other senses, such as hearing, touch, and smell to understand and interact with their environment. It is difficult for people who are blind to access quality reading materials in accessible formats. The Internet, a treasure trove of information and reading material, is mostly unattainable to the blind. However, there are not a lot of products and gadgets that can help the visually impaired in their everyday life. If blind people can interact with the world much better with the help of smart devices, then they will be able to gain increased independence and freedom, and this is the reason why we built such a smart glass system..

Our project aims to develop an intelligent glass that has both object detection and text-to-speech conversion. The glasses are designed to help visually impaired people read the scanned text which is captured with a camera and converted into audio text. So the person will hear the audio through a headphone that's connected to the glasses. This invention will help to motivate blind pupils to complete their studies despite all their difficulties. The second task of our glasses is to identify the objects in front of the blind people and this is done by using Tensor Flow Object Detection API. In our project the camera is connected to a microcontroller and the live video taken by the camera is fed to Tensor Flow Object Detection API. After the object is identified, the output is converted into audio text and transferred to the headphones connected to the glasses.

II. CASE STUDY

The majority of people with visual impairment and blindness are over the age of 50 but vision loss may affect people of all ages. This billion population includes persons with moderate or severe visual impairment or blindness caused by uncorrected refractive error (88.4million), cataract 94million), glaucoma (7.7million), corneal opacities (4.2million), and trachoma as well as near vision impairment due to unaddressed presbyopia. Visual impairment was an enormous global financial burden with the overall annual costs of lost productivity associated with a visual impairment from uncorrected myopia and presbyopia alone estimated to be US 24.4billion and US 25.4 billion. The experience of visually impaired individuals varies depending on many different factors. The report found that 7.8 million people in India are blind. Overall 20% of people are blind. We can only solve the blindness problem for the other 80%.

With regards to regional differences, it is estimated that the prevalence of remote vision disorders in low- and middle-income areas is four times higher than in high-income areas. While considering the near vision rates of unaddressed, it was estimated to be greater than 80%in western, eastern, and central sub-Saharan Africa, while relative rates in high-income regions of North America, Australasia, Western Europe, and Asia pacific are reported to be lower than 10%. There is substantial variation in the causes between and within countries availability of eye care services, their affordability, and the eye care literacy of the population.

III. RELATED WORK

There are so many inventions on this are some of them are.

3.1 Oton Glass

The product is developed by a Japanese company. The idea behind this product is the smart glasses are designed to help dyslexic reads and the camera will capture images or words that the user wants to read and reads the words for the user via the earpiece. The benefit of this product is the glasses will convert symbols into sounds. It resembles normal glass.

Fig 1 Oton Glass

3.2 Google Glass

This product is designed by Google Company. The idea of this product is to display google information without using hands and communicate in a normal manner using the internet. The advantage of google glass is that you can take pictures, record videos, obtain directions, and send messages. It is not affordable for everyone because of is expensive.

Fig 2 Google Glass

3.3 AIRA

The product is designed by Suman Kanuganti. The idea of Aira is to use smart glasses to scan the environment. The main advantage of Aira is that it helps users interpret their environment with smart glasses. The advantage of this product is that everything has to be done by a blind person.

Fig 3 Aira

IV. METHODOLOGY

4.1 OCR

Artificial intelligence is a field of computing in which a machine is trained to think and act, resembling intelligent human beings. Optical character recognition (OCR) is one branch of AI. It is used to detect and retrieve characters from scanned documents or pictures and convert them into editable forms. Earlier OCR methods used convolutional neural networks, but they are complicated and generally better suited for single traits. The error rate of these methods was also higher. Tesseract OCR Engine uses Long Short Term Memory (LSTM) which is a part of Recurrent Neural Networks. It is open-source and is most suited for handwritten texts. It is also convenient to recognize more text data instead of single characters. Tesseract OCR Engine considerably reduces the errors created in the character recognition process.

Fig 4 Tesseract OCR

Tesseract OCR

Tesseract assumes that the input image is a binary image and the processing is performed step by step. The first step is recognizing the interconnected elements. The outlines are embedded in blobs. The blobs are organized into lines of text. Text lines are broken according to the pitch. If there is a fixed pitch between the characters, then text recognition takes place, which is a two-way process.It is difficult for people who have a visual impairment to read textual information. Blind people need braille for reading. It would make it easier for them to listen to the audio format of the data. This application can be used to convert text data into an audio format to make it easier for people to hear the information. Google Text-To-Speech API can be used to convert text information into audio form.

4.2 Tensor Flow Object Detection API

Object detection is a computer vision technique that enables us to identify and locate objects in a picture or video. Real-time object detection is performed with quick inference while maintaining a basic level of accuracy.

Fig 5 Tensor Flow

A deep learning system can represent the printing of an object by combining simpler concepts, such as dots and lines, which are in turn defined in terms of borders. By using a variety of algorithms, a set of benchmarking data, and correct labeling packets, a system can be formed to achieve the desired result. The model is trained to detect live objects. This can be best achieved by the universal and open-source library-Tensor Flow (TF). In the TF environment, multiple algorithms may be used for a wide range of data sets.

The model we use here is the COCO SSD model, which is a dataset for detecting, segmenting and subtitling large-scale objects. The model is capable of detecting 90 classes of objects.

Fig 6 COCO Dataset

V. WORKING

The proposed system has two parts, one is a goggle and the other is a pocket device. The ESP32 Wi-Fi Cam module is attached to the goggle and a battery source for the module. And, the pocket device features the Raspberry Pi 3 module and its power supply. The power source for the raspberry pi will be a power bank and for the camera module, it is a replaceable battery.

The system is designed to perform two functions; i) real-time object detection and ii) text-to-speech conversion using OCR. To detect objects, the camera is connected to a microcontroller through Wi-Fi, capture the video. Each frame from this video is fed to the Tensor Flow Object Detection API and the objects are detected. And the result is fed to a Bluetooth headset from the Raspberry Pi. For text reading, the same camera is used. The images are fed to the Raspberry Pi, where the Tesseract OCR scans the image. The image is converted into strings by the OCR. And, the strings are read aloud by Google TTS.

VI. ARCHITECTURE

The architecture shows the multiple steps involved in the process. At first, the ESP32 camera record or send the live video to the microcontroller (Pi Raspberry3 B+) then from the microcontroller if the processed image has objects, then with the help of Tensor Flow Object Detection API the objects are identified. The Tensor Flow Object Detection API provides the framework for constructing a deep learning network that solves object detection issues. There are already predetermined models within their framework that they call Model Zoo. When our obtained image is given to the microcontroller the Tensor Flow Object Detection API installed in the Pi Raspberry3 B+ compares our object pre-trained models in their framework and gives suitable output. Then the output is converted to audio text and passed on to the headphone that's connected to the glasses. For text to speech conversion, the scanned image obtained from the ESP32 camera is transmitted to Tesseract that is installed in the microcontroller. Tesseract is an optical character recognition engine for different operational systems. It is free software, published under the Apache License, version 2.0, and the development has been sponsored by Goggle since 2006. We will use this frame to convert the optical character into text and then convert it into a natural voice. After obtaining the output, it is given to the headset that is connected to the glasses. The goal is to assist people with visual impairments to listen to the text or read a book.

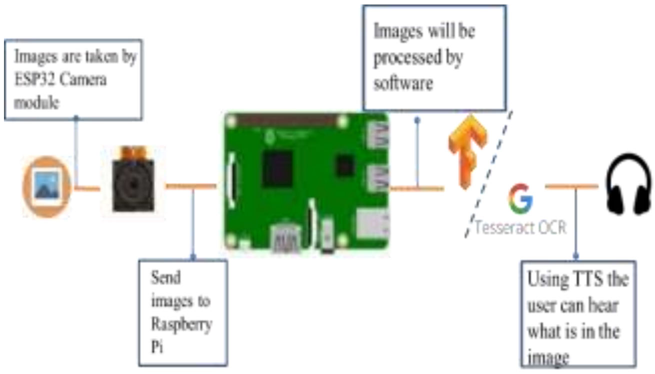

Fig 7 Architecture

VII. FLOWCHART

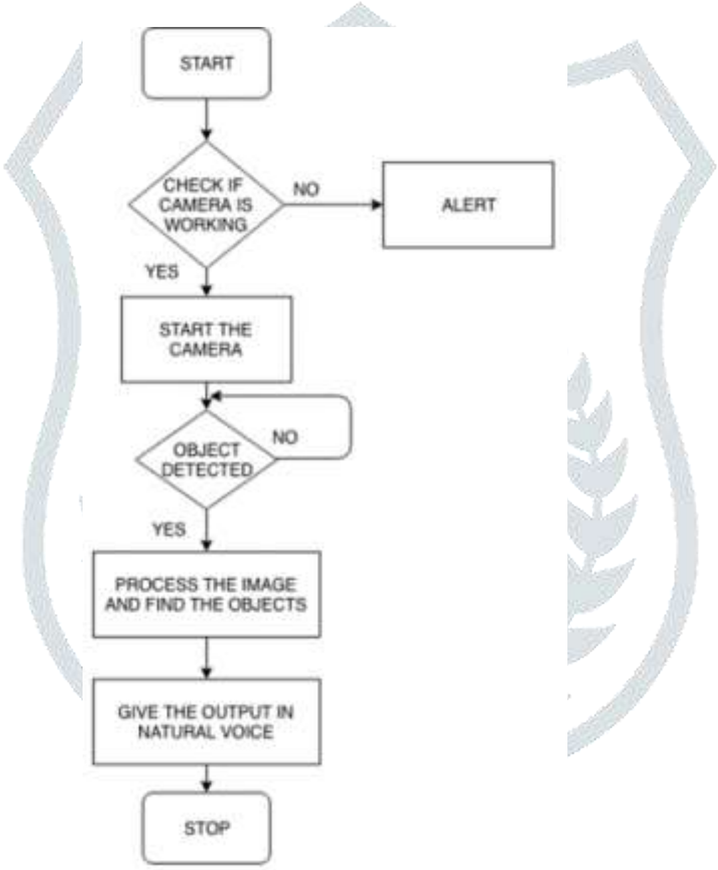

Fig 8 Flow Chart

The flowchart shows the graphical way for the flow of the process of each system model. At first, the Raspberry Pi checks if the camera is working. Once the validation is done the camera starts the live video recording and sends the data to the raspberry for further video processing. Whenever there is a voice command the Microcontroller records it and then sends it for speech recognition and based on AI the output is given.

VIII. RESULTS

We give a sample text to workout OCR. The output string is converted to the audio signal using Google TTS. And the output is represented as the audio waveform using an audio waveform generator.

Fig. 9 Sample input text for OCR

```
import easyocr
from PIL import Image
from pytesseract import *
img = Image.open("ocrpi.png")
output = easyocr.image_to_string(img)
print(output)
```

Fig. 10 program

Fig.11 Audio waveform

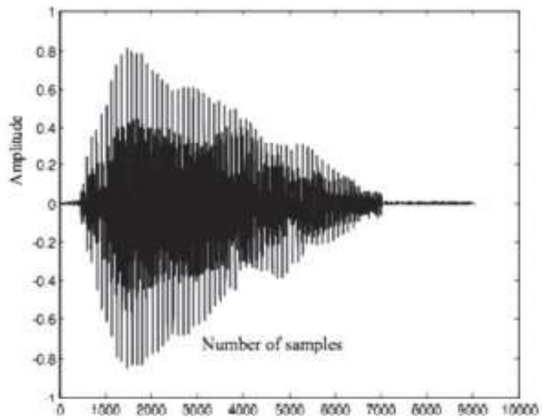

Fig. 12 Output detected for other samples

During, the detection, the algorithm codes will continuously analyze future text which will be used. This is vital because, the time of detecting the word is 1.65 seconds and is much longer than the time taken for processing each word, which is 0.09 seconds. The input can be converted quickly into speech. The speech clarity and quality is remarkably good.

IX. CONCLUSION

The paper introduces a prototype intelligent eyeglass system for the visually impaired. This product can assist many visually impaired people to experience and interact with their environment. Our product comes with object detection and text-to-speech conversion into far more effective and affordable, so that even ordinary people can access it. We use wireless connections to send and receive data, making the product easily transportable. Shortly, we will implement more useful applications in the intelligent glass system in the future.

X. FUTURE SCOPES

The system will be helpful for visually impaired people to read the data without the help of a braille script. In the future, the system can be expanded so that it can be used to read different regional languages, different handwritten notes, and various fonts. Also, in object detection, various developments can be done like training the machines according to the environment of the user. For recording here we used the ESP32 Cam module, in the future it can be replaced by Wi-Fi stream able and cost-effective compact cameras.

REFERENCES

[1] P. Selvi Rajendran, Padhmaveni Krishnan, B. John Aravindhar, "Design and implementation of voice-assisted smart glasses for visually impaired people using Google vision API", 2020 4th International conference on Electronics, Communication, and Aerospace technology

[2] Arivitha, Ambresh G Biradar, M. Chandana,"The Third eye for visually challenged using echolocation technology"2021 International Conference on emerging Smart Computing and informatics.

[3] Suganthi Lakshmanan, Divya B Nirmala, Annamalai M, Pragadeesh T, Sanju Varshini T, "Portable assistive system for visually impaired using Raspberry Pi", 2020 IEEE International Conference on Advances and Development in Electrical and Electronics Engineering (ICADEE).

[4] C Singaravel, E. Vignesh, "Development of prototype spectacles for the blind with digital to voice conversion device," International conference on Power, Energy, Control and Transmission System 2018.

[5] W.C.S.S. Simoes, V.F.de Lucena, "Blind user wearable audio assistance for indoor navigation based on visual markers and ultrasonic obstacle detection" Consumer Electronics(ICCE),2016 IEEE International Conference,2016