Smart Glass for Visually Impaired People

Shubash Medu T
Student, Dept of ISE
Atria Institute of Technology
Bangalore, India

Shrilakshmi HS
Student, Dept of ISE
Atria Institute of Technology
Bangalore, India

Poojashree KS
Student, Dept of ISE
Atria Institute of Technology
Bangalore, India

Mr. Srinivas BV
Assistant Professor, Dept of ISE
Atria Institute of Technology
Bangalore, India

Abstract– In order to help visually challenged students with their daily tasks, this study introduces a novel design for smart eyewear. While the price and efficacy of current assistive devices are limited, this wearable technology offers a workable answer. To enable users to access the recorded data and turn text into speech, the research largely focuses on text recognition, object detection, and face detection. A buzzer tone is also used to remind people to take their medications on time. The system uses a Raspberry Pi and a camera to capture images in order to achieve this, producing a prototype that is reasonably priced. The experiment shows that the prototype successfully fulfils its intended function.

Keywords — Python, Text-to-speech, MobileNet SSD, Headphone, Micro camera, Raspberry pi.

1 INTRODUCTION

Over the past few decades, there has been an increase in the number of people with visual impairments. According to the World Health Organisation (WHO), there are 285 million blind individuals worldwide. Unfortunately, many organisations and schools are ill-equipped to satisfy their needs because of a lack of assistive technologies and financial limitations. As a result, almost 90% of people with vision issues continue to live in poverty. Even though new technologies and aids are now available, they can occasionally be either prohibitively expensive (costing over $3000) or reasonably priced (costing around $200), but only have limited functionality, only being able to perform one or a few tasks.

In terms of assistive technology, wearables have shown to be the most successful because they either eliminate the need for hand use or only require a small amount. The most common type of gadget is head-mounted since it naturally points in the user's viewing direction, without the need for further instructions. This study presents a cutting-edge smart eyewear design that can help with several jobs while yet having a minimal construction cost. To communicate with the user, the design makes use of the newest Raspberry Pi, a camera, and an earphone. Due to page restrictions, only the results of the reading task are presented in this study; nevertheless, the possibility of including more tasks is considered.

People who have vision problems frequently need help from others to do daily duties. They have a special capacity for using touch, hearing, and other senses to comprehend their environment. Finding your way around strange areas like residences, subway stations, and restaurants can be difficult. Unfortunately, there are not many accessible tools out there to help them. Their independence and freedom would considerably rise if visually impaired people had access to a technology that could serve as a companion and deliver necessary instructions.

Although many people with visual impairments use the Braille system to read books and documents, there are several drawbacks to the system, including high costs for converting materials to Braille, the inability of people with normal vision to read Braille materials, difficulties in correcting errors in the system, and the requirement for extensive practise and time to read documents and books. Additionally, not all books, documents, and articles are available in Braille, and persuading everyone who is blind to use the Braille approach may be difficult.

To solve the issues mentioned above, our team developed the "Smart Glasses for Visually Impaired People" project, which aims to make wearable assistive glasses for the blind and visually impaired. These smart glasses are designed to assist reading of any printable content, including books, documents, and mobile messages. They do this by converting the text that was detected from a specific image into audio, which can be heard through speakers or microphones. The design, which is portable and fairly priced, uses image processing techniques to identify and extract text from images, which is then converted into a voice for the user. It has a Raspberry Pi module for programming. Two samples—a book page and a mobile document—that were successfully converted into audio format were used to test the final hardware design.

This initiative is incredibly affordable, open to people from all socioeconomic backgrounds, and quite beneficial. Digital image processing has numerous advantages over analogue image processing, including a wider selection of useful algorithms and the prevention of noise and distortion build-up during processing. A computer is used in digital image processing to run an algorithm on digital images. Because images are specified in two dimensions, digital image processing can be treated as a multidimensional system. The development of computers, mathematics, and the need for a wide range of applications in fields including the environment, commerce, agriculture, medical science, and agriculture have all had an impact on the establishment and growth of digital image processing.

Machine Learning: The study of how to develop and interpret algorithms that can "learn" from data to enhance task performance is known as machine learning (ML). Without explicit programming, machine learning (ML) algorithms create models from training data to generate judgements or predictions. These algorithms are useful in a variety of applications, including computer vision, speech recognition, email filtering, medicine, agriculture, and other areas where conventional algorithms fail.

Internet of Things: The "Internet of Things" (IoT) is the interconnection of physical things or groups of objects that are equipped with sensors, computing power, and other technologies and allow them to share data with other devices and systems over the internet or other communication networks. The name "Internet of Things" has come under fire for being misleading, despite the fact that these gadgets only need to be connected to a network and have a particular address for communication. These gadgets just require a network connection, not a connection to the public internet.

Artificial Intelligence: Artificial intelligence (AI) is the ability of machines to exhibit intelligence, as opposed to the intelligence exhibited by non-human animals and people, through the perception, synthesis, and inference of information. AI encompasses a variety of activities, including speech recognition, computer vision, language translation, and other input mappings. AI is used in many different applications, such as sophisticated web searches engines like Google, recommendation systems, self-driving cars like Waymo, automated decision-making systems, and competitive strategic gaming systems.

2 RELATED WORKS

A range of devices can assist the sight impaired read. A popular substitute that produces letters that may be read or written using a system of dots is the Braille reader. Users have a second option, the audiobook, which allows them to listen to books or newspapers that have been saved in audio format. You can read text on a computer screen and turn it into audio by using a screen reader, text to speech software, and e-book readers.

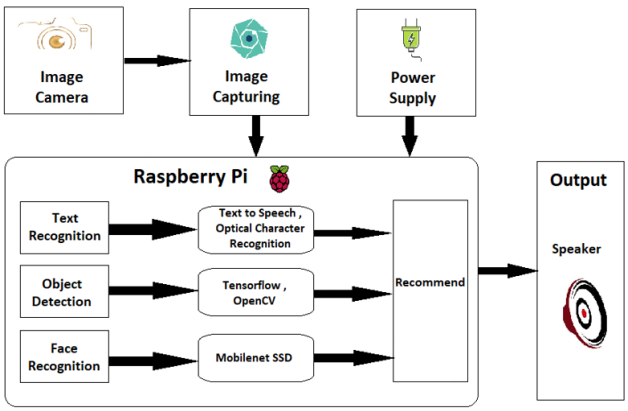
Eyewear technology is a more recent development in this area. For instance, OrCam glasses, which have an integrated computer and a gesture recognition system, can read and show content to the user in audio. Another example of low-vision technology is Esight glasses, which collect and analyse real-time scenes before displaying them in front of the user's eyes on a customised screen.

The blind and visually handicapped may now read text on a number of screens, including computer screens, mobile screens, and textbooks, thanks to the development of straightforward devices that can be carried in a shirt pocket and that emit audio that can be heard using a regular microphone.

3 METHODOLOGY

The capture button is located on the eyeglasses' handle. The camera on the Raspberry Pi device snaps a photo of the document to be read when it is turned on. After employing OCR technology to process the chosen image, the labels are subsequently separated using the CV library. This technique converts the scanned images into printed text. For text to speech conversion in this project, the libraries MobileNet SSD and Tesseract are suggested.

Finally, a text-to-speech engine converts the written text into audio and pronounces it; this audio can then be heard through headphones or speakers. IoT and Sonic wave processes are the foundation of Smart Glasses. To alert the user when a person or item is in front of the glasses and to report the distance, it reports both information. Using the Face dataset folder, the camera module may identify faces and either show a name or mark them as unknown. The battery-operated Smart Glasses utilise Python to function.



3.1 Text-to-speech technology

A text-to-speech (TTS) algorithm is a method for speaking written text. It is a crucial tool for many uses, including the development of audiobooks, language learning, and assistive technology. The text analysis module, natural language processing (NLP) module, voice synthesis module, and audio rendering module make up the structure of the TTS algorithm. To generate speech that sounds natural, the computer analyses the

input text to find linguistic traits and applies linguistic rules. More expressive and natural synthesised speech is now possible because of recent advancements in deep learning techniques like neural networks and deep reinforcement learning. However, difficulties still exist in producing speech that sounds human, especially in emotional expressiveness, voice modulation, and accent changes.

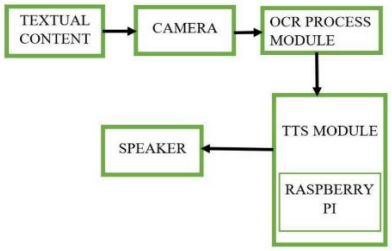


Figure 1: Text-to-speech technology

3.2 Convolution neural network

A CNN is a deep learning algorithm mostly utilised for image recognition and computer vision tasks. It is intended to perform convolution operations to identify patterns and features in images and is composed of many layers. These processes extract smaller traits, which are subsequently processed by other layers to uncover larger and more complex patterns. Additionally, pooling is a function of CNNs, which reduces computation time and output size. Because they are efficient at processing huge quantities of images, they are used in applications including facial recognition, object identification, self-driving cars, and medical image analysis. Additionally, CNNs have shown exceptional performance in the areas of audio processing, natural language processing, and time-series data analysis.

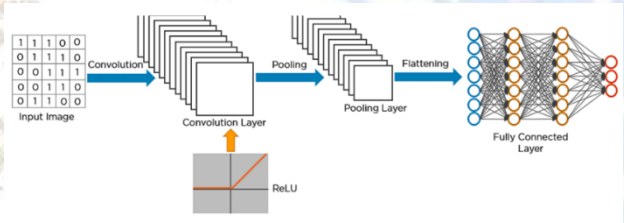


Figure 2: Convolution neural network

3.3 MobileNet SSD

For devices with constrained resources, including mobile phones and embedded systems, MobileNet SSD is a real-time object detection technique. It combines MobileNet and Single Shot Detector (SSD), two deep-learning models. MobileNet SSD uses a deep neural network in a single pass across the input image to quickly identify items. The algorithm can run effectively on low-power devices thanks to its minimal memory footprint. To use fewer parameters while still performing well, the model uses depth-wise separable convolutions. A popular option for real-time object detection in mobile applications, security systems, and robotics, MobileNet SSD has proven to perform exceptionally well in a variety of object detection benchmarks, including the COCO dataset.

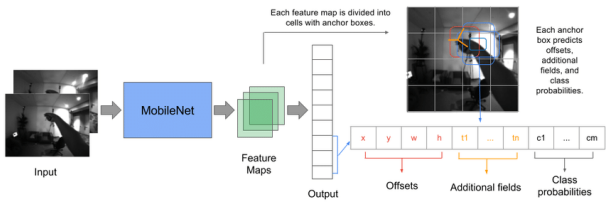
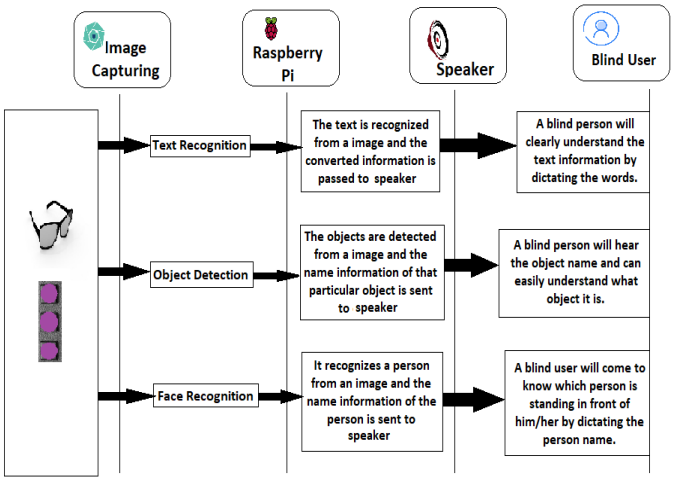


Figure 3: MobileNet SSD

4 SYSTEM ANALYSIS

4.1 Sequence Diagram



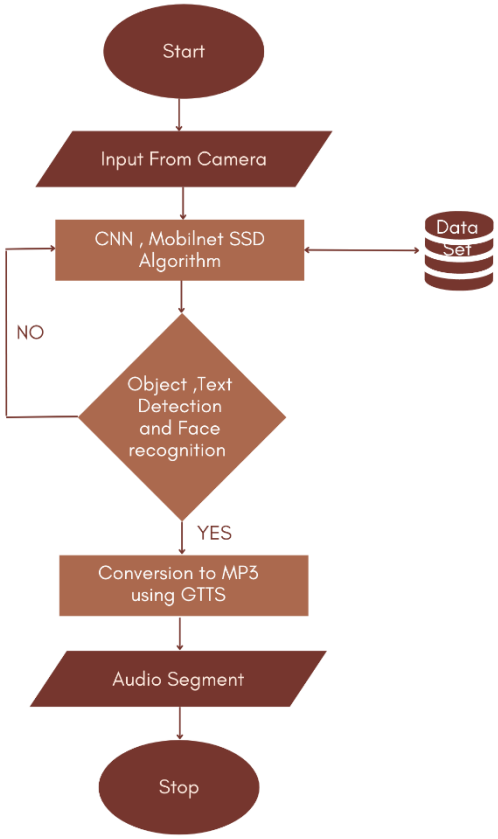
The given sequence diagram depicts a blind user's interaction with devices like spectacles and electronic buttons to choose the desired feature.

The "text recognition" feature button causes the system to run methods like OCR (Optical Character Recognition), which allow text to be extracted from images. Following the extraction of the text, it is turned into speech that the user can hear through headphones. The task of analysing the image and locating the characters existing in it is carried out by the algorithms. Text-to-speech technology is used to identify the characters, convert them to text, and then translate them into speech. This increases the accessibility of visual information for blind individuals by allowing them to access information from images via audio format.

The system starts the execution of pertinent algorithms, including Tensorflow and OpenCV, when the user clicks the "object detection" feature button. These algorithms make it easier to identify the items that are visible in the image, and their names are converted into audio that the viewer may hear through headphones. The algorithms use machine learning and computer vision techniques to analyse the image and find things.

Last but not least, when the user selects the "face recognition" feature button, algorithms like MobileNet SSD are engaged to identify the person in the image and transform their name into an audio format, which the user may hear through headphones.

4.2 Flow chart



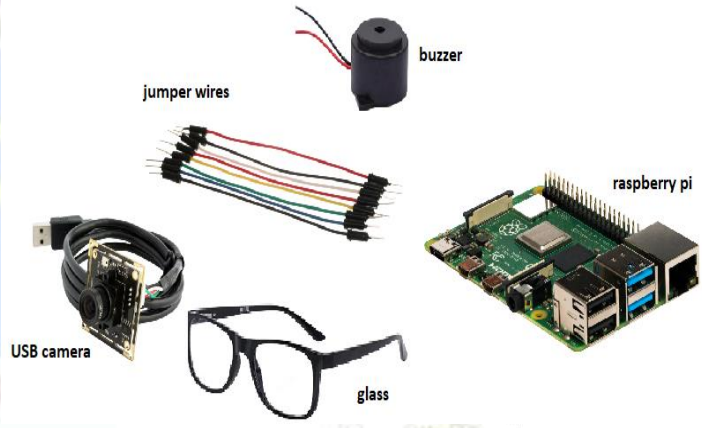
- **Start:** Blind user clicks on the raspberry pi.
- **Input:** The blind user can choose any 3 features as an input button option that is been provided such as object detection, text to speech, face recognition.
- **Process:** When the blind user clicks on either of the option, the respective algorithms will be called for processing the task and provides desired output.
- **Data Set:** Data set consists of images of a person in order to notify the blind user who is the person standing just front of him. If the image of a person is not included in data set, then it says as unknown person.
- **Conversion**: The output of these 3 features is converted into audio format.
- **Audio Segment:** The converted audio format can be heard using headphones.
- **Stop:** The process ends.

5 CONCLUSION

This paper introduces a revolutionary smart eyewear concept that uses a camera and a Raspberry Pi to help those with visual problems. The glasses are made especially for activities like object detection, face recognition, and text recognition, and more models may be added without much difficulty to increase the system's capabilities. Each model runs separately, giving consumers the option to choose the appropriate task. The design, operation, and principles of the system are discussed in the paper together with the findings of the experiments.

Even in low-income environments, the proposed smart glasses have the potential to enhance the lives of visually impaired children. In the future, efforts will be directed towards evaluating user-friendliness and improving power management. The system's straightforward construction yields a user-friendly, economical, and small product. A buzzer tone that reminds users of their prescription regimen is an added feature, and normal mobile phone microphones can access the audio output. The durable construction of the glasses ensures a lengthy lifespan. Overall, anyone with vision impairments should strongly consider using these intelligent reading glasses.

6 RESULT



7 REFERENCES

[1] S. Anbarasi, S. Krishnaveni, R. Aruna, K. Karpagasaravanakumar "Smart Reader Glass for Blind and Visually Impaired People" 2021.

[2] Md. Razu Miah, Md Sanwar Hussain "A Unique Smart Eye Glass for Visually Impaired People" ,2018.

[3] Pavan Hegde, Neha Devathraj, Sushma SK, Pasupuleti Aishwarya "Smart Glasses for Visually Disabled Person" ,2021.

[4] Ali Hassan, E., & Tang, T. B. (2016). Smart Glasses for the Visually Impaired People (pp. 579–582). https://doi.org/10.1007/978-3-319-41267-2_82

[5] Bai, J., Lian, S., Liu, Z., Wang, K., Liu, D.: Smart Guiding Glasses for Visually Impaired People in Indoor Environment. IEEE Transactions on Consumer Electronics, Vol. 63, No. 3, August 2017.

[6] Jafri, R., Ali, S.A.: Exploring the potential of eyewear-based wearable display devices for use by the visually impaired. In: International Conference on User Science and Engineering, Shah Alam, 2–5 September 2014

[7] J. Bai, S. Lian, Z. Liu, K. Wang, and D. Liu, "Smart guiding glasses for visually impaired people in indoor environment," IEEE Transactions on Consumer Electronics, vol. 63, no. 3, pp. 258–266, 2017.

[8] Chayma, Zatout & Larabi, Slimane, "A Novel Output Device for visually impaired and blind people's aid systems", in 2020 IEEE international conference on robotics and automation (ICRA). IEEE, 2017, pp. 6533–6540.

[9] Dr. Sankit Kassa, Manisha Inchanalkar, Amisha Kamble, Mayuri Inchanalkar "Smart vision Glasses for Blind People" ,2018.

[10] Jinqiang Bai, Shiguo Lian,Kai Wang, "Smart Guiding Glasses for Visually Impaired People in Indoor Environment", 2019, IEEE.