Grenze International Journal of Engineering and Technology, Jan Issue

# Third Eye: TTS and Deep Learning-based Assistive Technology for Blind

Asrarulhaq Maktedar[1] and Prathima M N[2]
[1-2]M S Ramaiah Institute of Technology, Bengaluru, India
Email: asrarmaktedar7491@gmail.com, prathima3006@msrit.edu

***Abstract*—According to the World Health Organization (WHO), more than 2.2 billion people throughout the globe are affected by some kind of visual impairment, either in their near or distant vision. Over 12 million people in India are affected by blindness, making it the country with the largest blind population in the world. Traditional techniques, on the other hand, are more difficult and costly to change, and most of the time, blind people end up needing to move about by gripping a wooden stick. People who are blind have very restricted access, or none at all, to any of the many intelligent devices that are presently on the market across the world. The project's objective is to find solutions to the issues so that a blind person using an android application called Smart Third Eye would be able to go to a variety of sites independently. This project assists blind people in recognizing the items that are in their immediate environment and enables them to determine the kind of object with the use of a Deep Neural Network and YOLO. The other objective of the research is to help blind person to use voice instructions to travel to various locations. They also have the option of using voice assistants to make phone calls, and their caretakers are able to monitor their whereabouts using GPS technology and different sensors.**

***Index Terms*— Deep Learning, Visual Impairment, YOLO, Android Application, GPS, Sensors.**

## I. INTRODUCTION

According to research that was carried out by the World Health Organization (WHO), there are approximately 2.2 billion people in the world who are blind or have significantly decreased eyesight. Out of these individuals, 62 million of them reside in India, of which 8 million have total vision impairment, according to the research. The visually impaired have a far greater number of challenges in their day-to-day life than a person with normal vision does, the most significant of which is the inability to move freely around the environment. Visually impaired persons have a difficult time investigating their environment, and mobility is the most major obstacle they face as a result of their lack of a thorough understanding and knowledge about their physical world. They also make use of canes to identify distinct objects and to avoid clashing with such objects. Canes, on the other hand, are not able to provide their users with any assistance in identifying the item, nor are they able to locate an item that is out of their immediate grasp. They rely almost entirely on the experiences they've had in the past in addition to their other senses in order to navigate through their environment. However, because of this reliance, they run the risk of experiencing accidents or injuries if they come across an unknown item in a setting that is unfamiliar to them, which can be hazardous. In recent years, technical breakthroughs have led to the creation of assistive devices that may be used by those who are blind or visually impaired.

*Grenze ID: 01.GIJET.9.1.508*

When we talk about Assistive Technology (AT), we're referring to any piece of hardware or piece of software that makes it simpler for people who are blind or visually impaired to perceive their surroundings. This, in turn, helps these individuals become more capable, independent, and self-sufficient. These technologies help people who are blind or visually impaired become more aware of their environment, which lessens the likelihood that they may hurt themselves or others, as well as the likelihood that others will hurt them. Assistive technology helps people react more effectively to a situation by giving them with information that helps them better understand the circumstance. In contrast to the development of wearable and portable device aids for the visually impaired, recent research indicates that there has been an uptick in the production of mobile application assistance for those who are blind or visually impaired. It is more desirable to have a single device that can accomplish many functions than to have numerous devices, and it should be more cost-effective to have the one device. The growing computational power of smartphones, in addition to the growing availability of smartphones, is creating incentives for the development of many applications with enhanced performance that are geared toward assisting visually impaired individuals in their day-to-day lives. The purpose of the investigation being carried out is to discover a wide variety of helpful technologies that have been developed on a platform based on mobile phones with the intention of assisting visually impaired people in their day-to-day life.

Real-time object identification is provided via the YOLO method, which is based on the utilization of neural networks. The efficiency and precision of this algorithm have contributed to its widespread use. It has found usage in a variety of applications, including those that detect animals, humans, and parking meters in addition to traffic signals. When using the Feature-based Approach, it is important to consider the local aspects of a person's face such as his or her eyes, mouth, nose, and so on. These features are utilized to segment the face and to provide input data for face detection. Holistic Approaches to face detection and identification take into account the full face when determining whether or not to utilize a particular face as an input. The Android application developed will help the blind person to enter into the world of technology and it can be the beginning of a new era for them.

## II. Related Work

Harsur A, et al. [4] created a voice-based system for visually impaired people employing ultrasonic sensors, which allows blind people to roam independently inside the outside world. This project is composed mostly of six modules, which are as follows: Interface, Initialization, Address query translation, Route Query, Transversal of the route, as well as identification of obstacles. Because it is the first module in the system library, it is in the responsibility of ensuring that the initialization of the system library is carried out correctly. Additionally, it is responsible for ensuring that the system library is ready for use. In the next section of the module, i.e. Using a microphone, the user interface gathers the address of the destination from the user at this stage. The speech interface that was developed in this module makes use of services such as Text-to-Speech for the voice outputs and the Google Voice Recognizer API for the recognition of voice inputs and outputs. In the last section of the module, i.e., As we go on to the last part of the module "Address query translator converts geographic coordinates to coordinates that contain latitude and longitude."

According to Manisha Bansode et al. [5], in 2015, they developed a speech-based system for the blind utilising a voice recognition module and GPS module installed on an Arduino board. The speech recognition module will get the name of the destination from a blind person, who will be the input to the system. A GPS receiver continually gets the current position, as well as its longitude and latitude, and compares it to the latitude and longitude of the destination, explains the manufacturer [6].

In 2015, Author Sonal S., et al stated, with this development, the technology of HC-SR04 sensor and sound vibrations came into reality. There are a few things that aren't ideal about it, including the fact that it can't be folded and the fact that it's hard to keep in place. This idea does not have any sensors for detecting water regions, which is another disadvantage [7].

The researchers quoted Muhsin Asaad H., et al. (2013), who said that a white cane is assigned to detect impediments up to knee level within a range of 2 to 3 feet in height (0.91 m). In the case that the cane detects an impediment, it will either vibrate or speak out loud. Because of this, the sensor, as well as the controller, are both housed inside the cane as a consequence of this, which has a battery life of 10 hours on a single charge in all conditions. By calculating the amount of time, it takes for the waves to go to the impediment and the amount of time it takes for the waves to travel back, we can determine the delay time. A buzzer is used to provide the user with distance information [8].

In this paper [9], RFID is a technological advancement that may assist in enhancing the organization and direction of people when participating in daytime activities. The device may be able to deliver a few bits of information to someone who is vision impaired. about the distinguishable and expedite the search; in addition to differentiating between medications, The user may get an audible indication from the device, which will assist them in locating the ideal item in the shortest amount of time feasible. A few chunks of information are made available to those who are visually challenged regarding the distinguishable of a designated item, namely the manner in which it is near or far away.

## III. Proposed Method

The creation of a system plan is one of the critical steps that must be taken in order to successfully deploy the necessary system. Because developing the architecture is part of the system plan, which is done first so that implementation may be done according to the architecture. Following that, data flow diagrams may be created on the basis of the architecture. Because of this, we will get an understanding of how the data flow should take place as well as the modules that are necessary to construct the necessary system. Therefore, the architecture and DFD make up the bulk of the system plan.

### *A. Modules and their descriptions*

The following is a list of the components that make up the system:

1. Messages:
The user may see the messages in two different formats: their inbox and the messages they have sent.
2. Calls:
The user is able to make calls from inside this app, which has a bespoke dialer and call records.
3. Battery:
The user may manually check the battery level, and the system will also provide a notification to the user when the battery is running low.
4. Notes:
The user is able to not only add new notes but also change it whenever it is necessary.
5. Text to Speech:
The programme takes care of everything, from performing the activities to even providing spoken reminders.
6. Read Messages/Notes:
Additionally, the technology assists the user in reading the contents of the message and the notes.
7. Navigation:
The System will take into account both the device's present position and its final destination location before determining the best route for navigation based on the information provided by Google Services.
8. Face Recognition:
The technology is able to recognize who is in the immediate vicinity of you and will announce their name to you as they come into view.
9. Object Detection:
The system is able to identify whether there are any items coming in front of him, and it will also produce a loud sound utilizing a buzzer when it does so. The distance between the ultrasonic sensor and the box would be

$$D = 0.5 \times 0.025 \times 343 \quad (1)$$

10. Object Recognition:
The system takes use of YOLO to detect the things that are located in the immediate vicinity of the user, after which it will play back an audio representation of the kind of object that was discovered.
11. Fall Detection:
The technology is able to determine whether or not the blind person has collapsed, and the individual's loved ones will be able to keep an eye on them by utilizing an application on their Android mobile.
12. GPS Location:
In addition, the caregivers are able to get longitude and latitude readings, which will be of great assistance to them in locating the spot where the blind person was injured.

### *B. System Design*

A Raspberry Pi camera is linked to the CPU in certain way. Python is the programming language used for the processor. The picture is captured in real time by the Pi camera and then sent to the Raspberry Pi-3 CPU so that it may be processed and also used for face recognition. The COCO model is used by the Python code in order to recognize and categorize the items. It will create border boxes around the objects that have been discovered, and it will also display the category index associated with the item. A Raspberry Pi camera is fastened to the frame of the wearer's smart glasses so that it may record the user's surroundings in real time. After some time has passed, the photographs are fed into the trained deep learning model, which then identifies any people who are in the images. As a consequence of this, the user's earphone is responsible for producing the voice output. Because of the tilt sensor that is linked to the system, it will become active if the individual is in danger of falling. an ultrasonic sensor is used in order to determine the distance at which the obstruction or item is located. And using the GPS, we will be able to simply determine the position of the blind person in the event that the fall detection feature is activated.

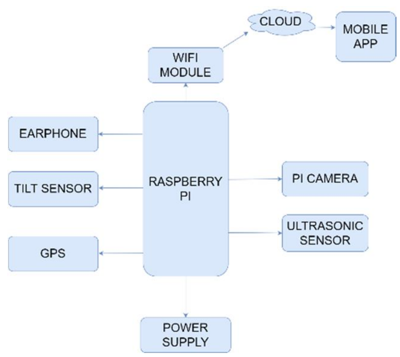


Figure 1. Proposed Model Hardware Architecture

In order to avoid the intelligent assistive system from making a mistake in its evaluation, we have merged the two fall recognition systems. When a fall signal is triggered by the proposed intelligent smart glasses as well as a fall signal is also obtained from the proposed wearable smart glasses, then it is largely decided to also be a legit fall event, and the current geolocation data is directly transmitted to the proposed cloud-based data management platform by a Wi-Fi module as shown in the Figure 1. Therefore, the location and time of the event are transferred to the proposed cloud-based information management platform for safekeeping, and a message is sent to the Android application in order to warn the family members or caretakers of the visually impaired user of the situation. On the other hand, if the proposed smart glasses do not detect a fall event but received a fall signal from the proposed wearable smart glasses, then this fall event will be rejected as the proposed smart glasses made an erroneous judgement.

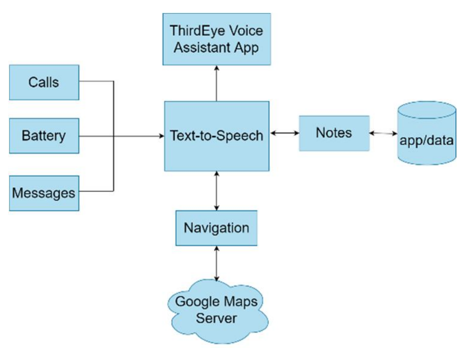


Figure 2. Proposed Model Software Architecture

The android application consists of mainly 5 modules, and these are Message, Calls, Notes, Battery and Google Map Navigation. All the modules work based on Text-to-Speech. The blind person gets all the information from an android mobile in voice output. The application is smart enough to identify the English as well as the Kannada language so that the blind person can be able to understand both the text messages from the receivers.
The application also enables the blind person to make a call to any person by dialing his/her number. The dial-pad is designed in such a way that, A blind person can easily identify the digits placed on screen. In order to navigate across the different places, the user has to shake his mobile phone, then android application will prompt the user to enter the destination address as voice input. The application also warns the user in case of low battery,

despite the fact that application is closed or running. Whereas the blind can also read the notes which are stored in db as shown in the Figure 2. The notes can be read, by just tapping on the notes and the text will be converted into voice.

## IV. Results and Discussion

### *A. Hardware Implementation*

The Raspberry Pi operates in an open-source environment. It uses Linux (a varied distribution), and Raspbian, the main operating system that is supported runs open-source software and is capable of running a variety of open-source programmes. While concurrently releasing a major percentage of its own code as open source.

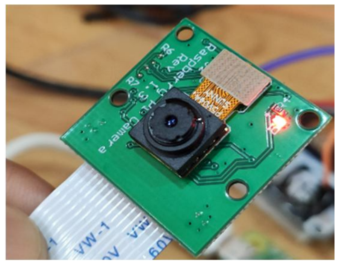

Figure 3. Raspberry Pi Camera

Raspberry Pi Camera has a flex cable in its package as shown in the Figure 3. The connector that is situated between the Ethernet port and the HDMI port is where the flex cable is placed, and the silver connection should be oriented so that it faces the HDMI port. To open the flex cable connection, pull the tabs on the top of the connector upwards and then toward the Ethernet port. This will open the connector. After that, the flex cable is carefully pushed all the way into the connection. and then the upper portion of the connection is pulled toward the HDMI connector and then down, all while the flex cable is retained in its current position. The Raspberry Pi camera is being used in this instance for both the process of face detection as well as face recognition.


Figure 4. HC-SR04 Ultrasonic Sensor

Ultrasonic Sensor is a piece of electronic equipment as shown in the Figure 4, that emits ultrasonic sound waves, then transforms the sound that is reflected back from the target item into an electrical signal in order to calculate its distance from the object. The pace at which audible sound travels is much slower than that of ultrasonic waves.

Figure 5. Smart Glasses for Blind with HC-HR04 and Pi Cam Mounting

The Smart Glasses consists of HC-SR04 Ultrasonic sensor mounted in front of the Smart glasses as shown in Figure 5, in order to detect the objects coming in front of the blind and the Pi Cam will help the blind to identify the type of object detected by HC-SR04. The Pi camera captures the object picture or human face and announce the name of the detected object.

*B. Mobile Application Development*

TTS is compatible with an extremely large number of personal digital devices, including desktop computers, mobile phones, and tablet PCs. It is possible to do text-to-speech conversion on files of any kind, including those written in Microsoft Word and Apple Pages, as well as any other kind of text file. Even the written material on websites accessible through the internet may be read aloud.

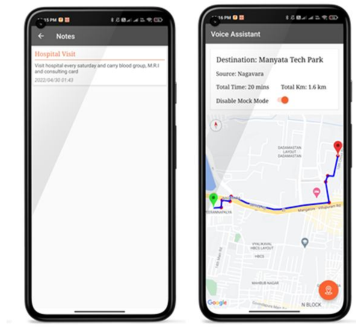


Figure 6. Voice Assitant Android App for Blind

Text is read out loud by the text-to-speech system (also known as TTS) in a computer-generated voice, and the tempo at which the text is read may generally be changed. The voices are of varying quality, but some of them have a human tone to them. In many text-to-speech (TTS) programmes, individual words are highlighted as they are read aloud. Because of this, a person who is blind or has low vision may be able to understand the text via the use of simultaneous auditory processing. The phrase "optical character recognition" refers to a function that is sometimes included in certain TTS programmes (OCR). The optical character recognition (OCR) technique enables text-to-speech (TTS) programmes to read text aloud from images. A syntax may be seen below:

```
ttobj.speak(toSpeak,
TextToSpeech.QUEUE_FLUSH, null)
```

Voice Assistant is an Android application that is based on TTS as shown in Figure 6 and it is used to assist blind people with several choices like as TTS for Messages, TTS for Calls, and TTS for Notes and Battery. Whereas a blind person may also navigate to any location in the globe with his voice commands, the use of which android application will route him to the destination would guide him there. Caregivers are also able to see the status of the blind person in the event that the blind person has an accident. The falls detection sensor is now being used, and it will communicate this information to the caregiver.


Figure 7: Face Recognition with Voice Playback

After the face or object has been identified with the support of the Pi camera that is being used with Raspberry Pi, the system will announce the name of the person whose face has been recognized as shown in Figure 7 or type of the object that has been identified as shown in Figure 8. The system will make the announcement, and a blind person may use headphones to listen to the audio playback of the announcement.

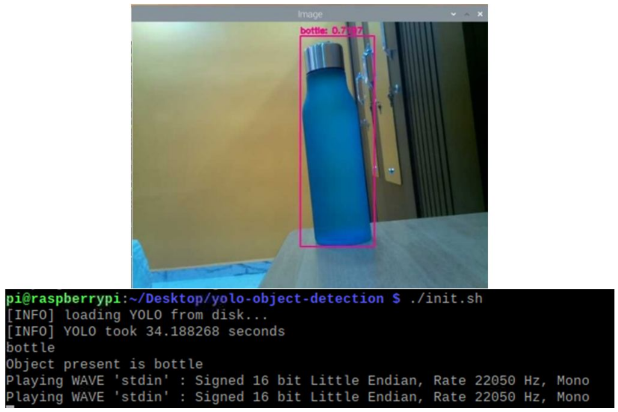


Figure 8: Object Detection with Voice Playback

When opposed to object recognition, face recognition has a higher level of accuracy due to the fact that the system uses CNN to analyze several criteria before making a guess about the person's name.

## V. Conclusions

The study suggests a smart glasses system that is both low-cost and simple to implement, with the goal of assisting those who are blind in leading more independent lives. With the assistance of this programme, the visually impaired person will have no trouble getting about to various locations, and he will also be able to utilize his android phone with the assistance of text-to-speech. The module's mass manufacturing would result in a further reduction in its cost, as well as GPS-based navigation and the sharing of positions with the caretakers, all of which would ensure that the caretakers' loved ones are kept safe and are doing well. Therefore, enabling persons who are blind or visually impaired to "see without the eye" Pre-process Image.

## References


[1] Krishnan, R. Santhana, K. Lakshmi Narayanan, S. Mathumitha Murali, A. Sangeetha, C. Ramasamy Sankar Ram, and Y. Harold Robinson. "IoT based Blind People Monitoring System for Visually Impaired Care Homes." In 2021 5th International Conference on Trends in Electronics and Informatics (ICOEI), pp. 505-509. IEEE, 2021.

[2] Park, Jae Sung, Cheon Lee, Daesung Lim, Seongwoon Jung, Jiman Kim, Junghwa Choi, and YoungSu Moon. "Inclusive Convolutional Neural Network Design Enabling Partially Sighted People to Expand Viewing-Experience on Smart Screens." In 2022 IEEE International Conference on Consumer Electronics (ICCE), pp. 01-04. IEEE, 2022.

[3] Masud, Usman, Tareq Saeed, Hunida M. Malaikah, F. U. Islam, and Ghulam Abbas. "Smart Assistive System for Visually Impaired People Obstruction Avoidance through Object Detection and Classification." IEEE Access 10 (2022): 13428-13441.

[4] Harsur, Anushree, and M. Chitra. "Voice based navigation system for blind people using ultrasonic sensor." IJRITCC 3 (2017): 4117-4122.

[5] Bansode, Manisha, Shivani Jadhav, and Anjali Kashyap. "Voice recognition and voice navigation for blind using GPS." International Journal of Innovative Research in Electrical, Electronics, Instrumentation and Control Engineering 3, no. 4 (2015): 91-94.

[6] Mahdi Safaa, A., H. Muhsin Asaad, and I. Al-Mosawi Ali. "Using ultrasonic sensor for blind and deaf persons combines voice alert and vibration properties." Research Journal of Recent Sciences ISSN 2277 (2012): 2502.

[7] Alam, Syed Tehzeeb, and Sonal Shrivastava. "Smart Device for Blind People." Journal: International Journal of Engineering Research and Technology (IJERT), ISSN (2015): 2278-0181.

[8] Mahdi Safaa, A., H. Muhsin Asaad, and I. Al-Mosawi Ali. "Using ultrasonic sensor for blind and deaf persons combines voice alert and vibration properties." Research Journal of Recent Sciences ISSN 2277 (2012): 2502.

[9] Dionisi, Alessandro, Emilio Sardini, and Mauro Serpelloni. "Wearable object detection system for the blind." In 2012 IEEE International Instrumentation and Measurement Technology Conference Proceedings, pp. 1255-1258. IEEE, 2012.

[10] Alshbatat, Abdel Ilah Nour. "Automated mobility and orientation system for blind or partially sighted people." International journal on smart sensing and intelligent systems 6, no. 2 (2013).

[11] Lakde, Chaitali K., and Prakash S. Prasad. "Review paper on navigation system for visually impaired people." International Journal of Advanced Research in Computer and Communication Engineering 4, no. 1 (2015): 166-168.

[12] Tapu, Ruxandra, Bogdan Mocanu, and Ermina Tapu. "A survey on wearable devices used to assist the visual impaired user navigation in outdoor environments." In 2014 11th international symposium on electronics and telecommunications (ISETC), pp. 1-4. IEEE, 2014.
[13] Parikh, Neel, Ishita Shah, and Safvan Vahora. "Android smartphone based visual object recognition for visually impaired using deep learning." In 2018 International Conference on Communication and Signal Processing (ICCSP), pp. 0420-0425. IEEE, 2018.
[14] Shishir, Md Amanat Khan, Shahariar Rashid Fahim, Fairuz Maesha Habib, and Tanjila Farah. "Eye Assistant: Using mobile application to help the visually impaired." In 2019 1st International Conference on Advances in Science, Engineering and Robotics Technology (ICASERT), pp. 1-4. IEEE, 2019.
[15] Balani, Yogita, Deepa Narayanan, Shashank Parande, Aakash Birari, and Anjali Yeole. "Drishti-Eyes for the blind." In 2019 IEEE 5th International Conference for Convergence in Technology (I2CT), pp. 1-6. IEEE, 2019.