

An Intelligent Voice Assistance System For The Visually Impaired People

Surabhi Suresh

dept. of computer science & engineering
Mahaguru Institute Of Technology
Alappuzha, Kerala.
Email : srbhisrsh2020@gmail.com

Sandra S

dept. of computer science & engineering
Mahaguru Institute Of Technology
Alappuzha, Kerala.

Aisha Thajudheen

dept. of computer science & engineering
Mahaguru Institute Of Technology
Alappuzha, Kerala.

Subina Hussain

dept. of computer science & engineering
Mahaguru Institute Of Technology
Alappuzha, Kerala.

Amitha R

Asst Professor
dept. of computer science & engineering
Mahaguru Institute Of Technology
Alappuzha, Kerala.

Abstract-For those who are fully blind, independent navigation, object recognition, obstacle avoidance, and reading activities are quite challenging. This system offers a fresh kind of assistive technology for persons who are blind or visually impaired. The Raspberry Pi was chosen as an example of the suggested prototype due to its low cost, small size, and simplicity of integration. The user's proximity to the obstruction is measured using the camera and ultrasonic sensors. The technology incorporates an image to text converter after which auditory feedback is provided. So a real-time, accurate, cost-effective, portable voice assistance system for the visually impaired was developed using deep learning. By identifying the things and alerting the users, the goal is to guide blind individuals both indoors and outside. Additionally, a reading aid that helps the blind read is available. When a person who is blind is in danger, the system integrates a pulse oximeter to SMS or contact an emergency number and alert them. This system encourages persons who are blind or have vision impairments to be independent and self-reliant.

Keywords –Deep Learning, Real time object detection, pytesseract

I.INTRODUCTION

Blindness ranks among the most common disabilities. People who are blind are becoming more prevalent today than they were a few decades ago. Individuals who are partially blind experience blurry vision, difficulty seeing in the dark, or tunnel vision. A person who is totally blind, however, has no vision at all. Our system's goal is to support these people in a number of different ways. These glasses, for instance, are quite beneficial in the field of education. Blind or visually impaired people can read, study, and learn from any printed text or image. Our method encourages persons who are blind or have vision impairments to be independent and self-reliant.

Our goal is to develop a real-time, accurate, portable, and affordable voice assistance system for those who are visually impaired using deep learning. The goal of this project is to assist blind people in both indoor and outdoor settings. The user can be appropriately informed when an object is recognized. Additionally, it offered a reading aid that helps people who are visually impaired read. The project offers a wide range of potential applications as a result. The system consists of:

- Navigational aids can be affixed on glasses that also include a built-in reading aid for permanent usage.
- Complex algorithms can be handled by low-end configurations.
- Real-time, accurate distance measurement using a camera-based method
- A built-in reading aid that can turn pictures into words, enabling blind people to read any document.

Therefore, the goal of our problem statement is to develop a real-time, accurate, and portable voice assistance system for people who are visually impaired. The goal of this project is to assist blind people in both indoor and outdoor settings. The user can be appropriately informed when an object is recognized. Additionally, it offered a reading aid that helps people who are visually impaired read. The project offers a wide range of potential applications as a result.

II.OBJECTIVES

The ultimate objective of such a system is to enable people who are blind to access information and carry out tasks on their own, without the aid of others. An intelligent voice assistance system can offer visually impaired users a smooth and personalised experience by utilising the power of deep learning, improving their quality of life and encouraging more independence.

Objective 1: To give the user a more trustworthy and effective navigational aid

Objective 2: To warn the user of obstacles and guide them

around them.

Objective 3: To alert an emergency number if the user's pulse changes.

III. LITERATURE REVIEW

Various systems have been analysed and summarized below:

A smart guiding glass for blind individuals in an indoor setting was suggested by Jinqiang Bai et al. To determine the depth and distance from the barrier, it uses an ultrasonic sensor and depth camera. The depth data and obstacle distance are used by an internal CPU to provide an AR depiction to the AR glasses and audio feedback through earphones. Its flaw is that it is only appropriate for indoor settings.

An RGB-D sensor-based navigation aid for the blind was proposed by Aladrén et al. This system uses a combination of colour and range information to identify objects and alerts users to their presence through audible signals. Its key benefit is that, in comparison to other sensors, it is a more affordable choice and can identify objects with more accuracy. Its shortcomings include its restricted range and inability to operate in the presence of transparent things.

Wan-Jung Chang et al. suggested an AI-based solution for blind persons to assist pedestrians at zebra crossings. This method recommends a system of assistance based on. Blind people can utilise artificial intelligence (AI) tools to help them cross zebras. According to test results, this method is extremely accurate, with an accuracy rate of 90%. This system's primary flaw is that it can only be used for one thing.

For the blind, Rohit Agarwal et al. suggested using ultrasonic smart glasses. A pair of spectacles with an obstacle detecting module mounted in the middle are used to implement this device. The obstacle detection module is made up of a processing unit and an ultrasonic sensor. By using an ultrasonic sensor, the control unit learns about the obstruction in front of the user. It then processes this information and, depending on the situation, provides the output through a buzzer. This technology will be inexpensive, portable, and user-friendly. Only items less than 300 cm can be detected by the technology, which is inaccurate. It does not provide any object information.

A visual aid for blind persons was suggested by Nirav Satani et al. Hardware including a Raspberry Pi processor, batteries, camera, goggles, headphones, and a power bank are included in the setup. Continuous feeds of images from the camera are used for image processing and detection. Deep learning and R-CNN are used to do this. It converses with the user using audio answers.

According to Graves, Mohamed, and Hinton (2013), deep learning techniques for voice recognition. Using deep recurrent neural networks to recognise speech. (pp. 6645–6649) in 2013 IEEE International Conference on Acoustics, Speech and Signal Processing. IEEE.

The idea of deep recurrent neural networks (RNNs) for speech recognition is introduced in this fundamental research. Since then, deep RNNs have been used often in voice recognition systems, serving as the basis for smart voice assistants.

A. C. Tossou, L. Yin, and S. D. Bao (2020). Opportunities and challenges for voice assistants for the blind based on deep learning. 51(4), 318-327, IEEE Transactions on Human-Machine Systems.

The advantages and disadvantages of using deep learning techniques in voice assistants for the blind are covered in this review paper. The potential advantages of adopting deep learning models are highlighted, and important factors including privacy, robustness, and user experience are noted.

S. Dixon, S. Cunningham, and F. Wild. Hey Mycroft, open-source data collection for a voice assistant. In Interactive Media Experiences: Proceedings of the 2019 ACM International Conference (pp. 287–292). ACM. The Hey Mycroft dataset, an open-source dataset for developing and testing voice assistants, is presented in this study. These datasets are essential for developing deep learning models suited for users who are blind or have low vision, allowing researchers to address the unique issues they experience in

D. Trivedi, M. Trivedi, and others (2019). Voice-assistance smart stick for the blind. 8(6):2246–2449 in International Journal of Innovative Technology and Exploring Engineering.

In order to help people who are visually impaired, this article suggests a smart blind stick integrated with a voice assistant. The study shows how it helps user .

IV. PROPOSED SYSTEM

The proposed system consists of both software and hardware elements. The setup shown includes an ultrasonic sensor, camera, Raspberry Pi, and power source. Using Real-time video of the environment can be recorded with the camera. The block diagram of the suggested system is shown in Figure 1.

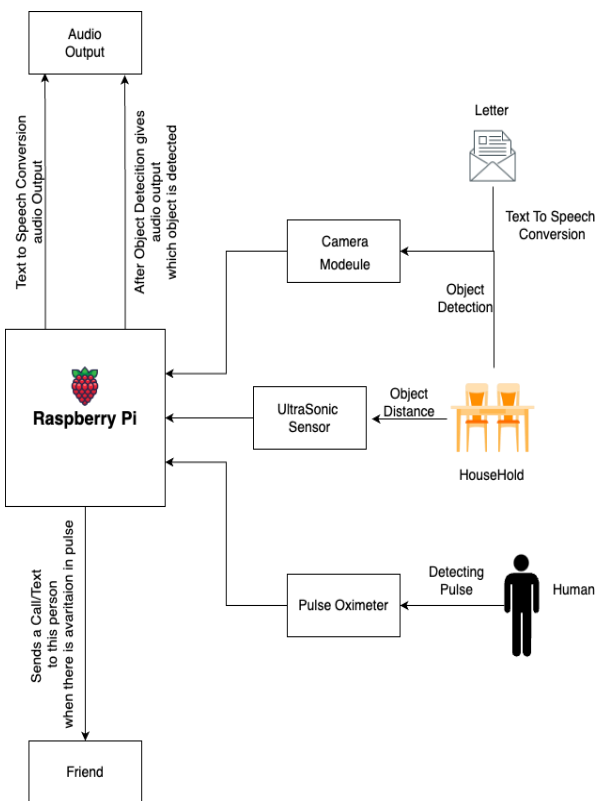


Fig 1. Block diagram of proposed system

Object detection

The Raspberry Pi commonly uses a camera module, machine learning techniques, and software libraries like OpenCV and TensorFlow for object detection of everyday objects. The process for detecting everyday objects with a Raspberry Pi can be summed up as follows:

1. Obtaining and preparing the dataset: A collection of photographs of home goods is gathered, and each object of interest is bound by boxes. Following that, the dataset is divided into training and testing sets.
2. Training the machine learning model: A machine learning model like YOLOv3, SSD, or Faster R-CNN is trained using the training set. Deep learning algorithms and libraries like TensorFlow or PyTorch are used to train the model. The purpose of

training is to improve the model's ability to detect objects.

3. Using the Raspberry Pi to detect objects: The camera module and trained model are installed on the Raspberry Pi. Real-time object detection and classification are performed by the model as the Raspberry Pi receives a live video feed from the camera module. The output can be saved to a file or seen on a screen.

Text to speech

The workflow for text-to-speech conversion for letters shown in front of a camera can be summarized as follows:

1. Image acquisition: A camera module captures an image of the letter(s) displayed in front of it.
2. Text recognition: Optical character recognition (OCR) algorithms are applied to the image to recognize the text content of the letter(s) and convert it into machine-readable text.
3. Text-to-speech conversion: The recognized text is passed through a text-to-speech (TTS) engine to generate an audio output. The TTS engine converts the text into a natural-sounding voice using pre-recorded human voice samples or computer-generated speech synthesis.
4. Audio output: The synthesized speech is played back through a speaker or headphones, making the letter(s) accessible to individuals who are visually impaired.

Sending an alert when pulse varies (pulse oximeter)

The process for a pulse oximeter that recognises changes in pulse and texts or calls an emergency number is as follows:

1. Sensor acquisition: To continuously track the person's heart rate and oxygen saturation levels, a pulse oximeter sensor is fastened to their finger.
2. Data processing: An algorithm processes the sensor data to find variations in pulse rate that deviate from a set range. An alert is generated by the algorithm if such a variation is found.
3. Creation of the alert: The alarm launches a script that calls or texts a predetermined emergency number. To transmit the alarm, the script may make use of text messaging or call service APIs.
4. Emergency response: The alarm is received by the emergency contact, who can then take the necessary steps to safeguard the person's safety.

Distance from the object calculation using ultrasonic Sensor

The process for determining the separation between an object and a Raspberry Pi using ultrasonic sensors can be summed up as follows:

1. Sensor acquisition: To gauge a distance from an item, a Raspberry Pi is connected to an ultrasonic sensor module.
2. Pulse generation: To begin the distance measurement, the Raspberry Pi sends an ultrasonic sensor a trigger pulse.
3. Echo detection: The ultrasonic sensor pulses an object, which the object then bounces back to the sensor. It is measured how long it takes the pulse to return to the sensor.

V.SYSTEM IMPLEMENTATION

The implementation view of the proposed system is shown below :

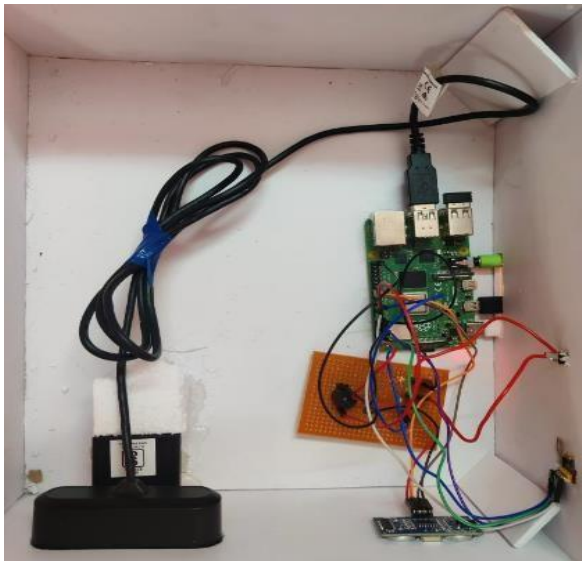


Fig 2. Implementation view

Real-time video is being captured by the camera module, and the item is being recognized using a deep learning algorithm based on a trained model. The pulse oximeter is operating concurrently and, if the pulse is not detected or if it exceeds the preset value, will place an alert call to an emergency number provided by the user. Real-time object detection will halt and text to speech will start when the button is pushed. Any image with text on it may be presented to the camera, and using Pytesseract, it will speak the text and provide the user with audible feedback. The echo

and trigger pin of the ultrasonic sensor are present. The echo pin receives the ultrasonic waves that the trigger pin transmits. When the user is less than 30 cm from the object, the buzzer will sound every 2.5 seconds, and when it is less than 15 cm, it will sound every 0.5 seconds, alerting the user to the obstruction in front of him. Then, calculate the distance using this value.

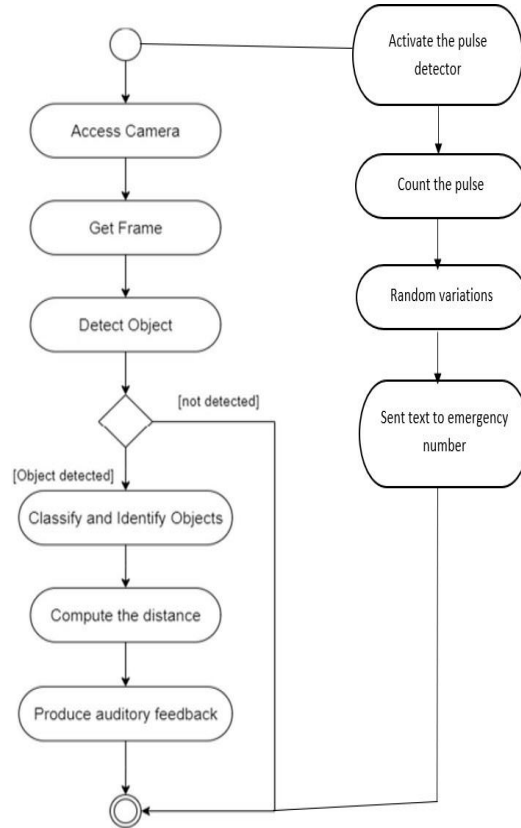


Fig 3. Proposed system workflow

VI.RESULTS

The user received audio feedback from the labels on the bounding boxes when the object detection process was successful. The pulse oximeter sends infrared rays into our blood to measure oxygen levels and sends an emergency call if random fluctuations were discovered.

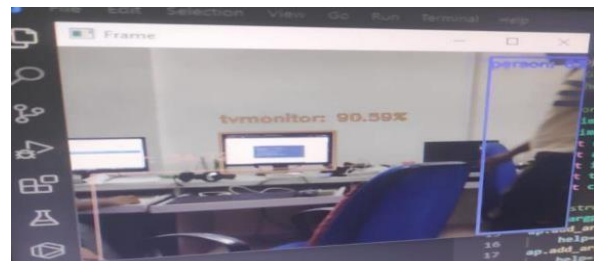


Fig 4. Object detection

When there was an obstruction in front of the user, the ultrasonic sensor also successfully activated the buzzer.

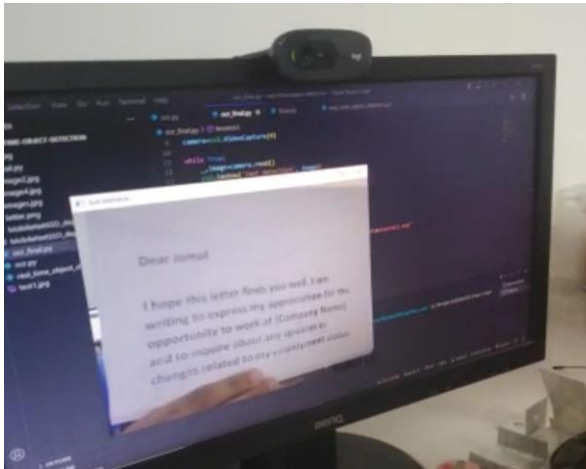


Fig 5. Image to text conversion

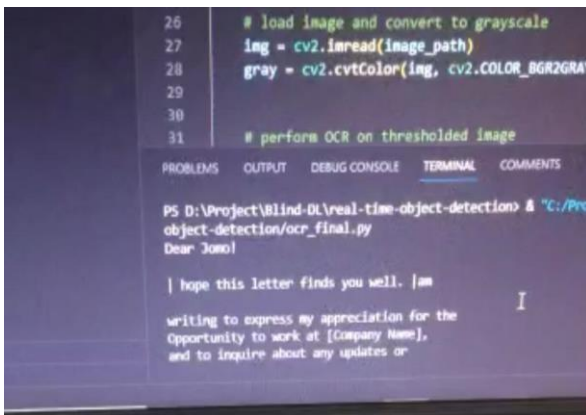


Fig 6. Converted text

VII.CONCLUSION

Using ML, we created a Voice Assistance for Visually Impaired. The system can be added as a tool for those who are blind. Any obstructions in the user's route are identified by the ultrasonic sensor that is implanted inside the device. If a barrier is discovered, the appropriate warning is sent. Despite the absence of current advanced technologies like wet-floor sensing or the use of GPS and mobile connection modules, flexible architecture enables future updates and improvements. The system may also be developed and tested outside, owing to improved machine learning algorithms and a new user interface.

VIII.ACKNOWLEDGEMENT

To everybody who helped us complete this project, we would like to extend our deepest gratitude. We had several challenges while working on the project due to a lack of information and skill, but these people helped us get beyond all of these challenges and successfully realise our idea to make a sculpture. We would like to thank Assistant Professor Mrs. Amitha R for her advice and Assistant Professor Mrs. Namitha T N for organising our project so that everyone on our team could comprehend the finer points of project work.

We would like to thank our HOD, Mrs. Suma S G, for her constant guidance and support during the project work. Last but not least, we would like to thank the Mahaguru Institute of Technology's management for providing us with this learning opportunity.

IX.REFERENCES

[1] Renju Rachel Varghese; Pramod Mathew Jacob; Midhun Shaji; Abhijith R; Emil Saji John; Sebin Beebi Philip, "An intelligent voice assistance system for visually impaired" IEEE international conference 2022
 [2] J. Bai, S. Lian, Z. Liu, K. Wang, and D. Liu, "Virtual-blind-road following based wearable navigation device for blind people," IEEE Trans. Consum. Electron., vol. 64, no. 1, pp. 136–143, Feb. 2018.
 [3] J. Xiao, S. L. Joseph, X. Zhang, B. Li, X. Li, and J. Zhang, "An assistive navigation framework for the visually impaired," IEEE Trans. HumanMach. Syst., vol. 45, no. 5, pp. 635–640, Oct. 2015.
 [4] J. Bai, S. Lian, Z. Liu, K. Wang, and D. Liu, "Smart guiding glasses for visually impaired people in indoor environment," IEEE Trans. Consum.
 [5] X. Yang, S. Yuan, and Y. Tian, "Assistive clothing pattern recognition for visually impaired people," IEEE Trans. Human-Mach. Syst., vol. 44, no. 2, pp. 234–243, Apr. 2014.
 [6] S. L. Joseph, "Being aware of the world: Toward using social media to support the blind with navigation," IEEE Trans. Human-Mach. Syst., vol. 45, no. 3, pp. 399–405, Jun. 2015.

- [7] A. Karmel, A. Sharma, M. Pandya, and D. Garg, "IoT based assistive device for deaf, dumb and blind people," *Procedia Comput. Sci.*, vol. 165, pp. 259–269, Nov. 2019.
- [8] C. Ye and X. Qian, "3-D object recognition of a robotic navigation aid for the visually impaired," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 26, no. 2, pp. 441–450, Feb. 2018.
- [9] Y. Liu, N. R. B. Stiles, and M. Meister, "Augmented reality powers a cognitive assistant for the blind," *eLife*, vol. 7, Nov. 2018, Art. no. e37841.
- [10] A. Adebisi, "Assessment of feedback modalities for wearable visual aids in blind mobility," *PLoS One*, vol. 12, no. 2, Feb. 2017, Art. no. e0170531
- [11] J. Bai, S. Lian, Z. Liu, K. Wang, and D. Liu, "Smart guiding glasses for visually impaired people in indoor environment," *IEEE Trans. Consum.*
- [12] J. Villanueva and R. Farcy, "Optical device indicating a safe free path to blind people," *IEEE Trans. Instrum. Meas.*, vol. 61, no. 1, pp. 170–177, Jan. 2012.
- [13] P. M. Jacob and P. Mani, "A Reference Model for Testing Internet of Things based Applications," *Journal of Engineering, Science and Technology (JESTEC)*, vol. 13, no. 8, pp. 2504-2519, 2018.
- [14] P. M. Jacob, Priyadarsini, R. Rachel and Sumisha, "A Comparative analysis on software testing tools and strategies," *International Journal of Scientific & Technology Research*, vol. 9, no. 4, pp. 3510- 3515, 2020.
- [15] Iannizzotto, Giancarlo, Lucia Lo Bello, Andrea Nucita, and Giorgio Mario Grasso. "A vision and speech enabled, customizable, virtual assistant for smart environments." In 2018 11th International Conference on Human System Interaction (HSI), pp. 50-56. IEEE, 2018.
- [16] Rama M G, and Tata Jagannadha Swamy. "Associative Memories Based on Spherical Separability." In 2021 International Conference on Emerging Techniques in Computational Intelligence (ICETCI), pp. 135- 140. IEEE, 2021.
- [17] S. Qiao, Y. Wang and J. Li, "Real-time human gesture grading based on OpenPose," 2017 10th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI), 2017, pp. 1-6, doi: 10.1109/CISP-BMEI.2017.8301910.
- [18] Philomina, S., and M. Jasmin. "Hand Talk: Intelligent Sign Language Recognition for Deaf and Dumb." *International Journal of Innovative Research in Science, Engineering and Technology* 4, no. 1 (2015): 18785-18790.
- [19] Valsaraj, Vani. "Implementation of Virtual Assistant with Sign Language using Deep Learning and TensorFlow."
- [20] Sangeethalakshmi, K., K. G. Shanthi, A. Mohan Raj, S. Muthuselvan, P. Mohammad Taha, and S. Mohammed Shoaib. "Hand gesture vocalizer for deaf and dumb people." *Materials Today: Proceedings* (2021).
- [21] D. Someshwar, D. Bhanushali, V. Chaudhari and S. Nadkarni, "Implementation of Virtual Assistant with Sign Language using Deep Learning and TensorFlow," 2020 Second International Conference on Inventive Research in Computing Applications (ICIRCA), 2020, pp. 595-600, doi: 10.1109/ICIRCA48905.2020.9183179.
- [22] M. R. Sultan and M. M. Hoque, "ABYS(Always By Your Side): A Virtual Assistant for Visually Impaired Persons," 2019 22nd International Conference on Computer and Information Technology (ICCIT), 2019, pp. 1-6, doi: 10.1109/ICCIT48885.2019.9038603.
- [23] Ramamurthy, Garimella, and Tata Jagannadha Swamy. "Novel Associative Memories Based on Spherical Separability." *Soft Computing and Signal Processing: Proceedings of 4th ICSCSP 2021*: 351.