Seeing Through Sound: Object Detection and Distance Analyzer for the Visually Impaired Using Audio Feedback

Anishamol Abraham
Assistant professor,Dept of CSE
Amal Jyothi College of Engineering
Kottayam, India
anishamolabraham@amaljyothi.ac.in

Bini M Issac
Assistant professor,Dept of CSE
Amal Jyothi College of Engineering
Kottayam, India
binimissac@amaljyothi.ac.in

Allen Roy
Dept of CSE
Amal Jyothi College of Engineering
Kottayam, India
allenroy2023@cs.ajce.in

Abhijith J Thoppil
Dept of CSE
Amal Jyothi College of Engineering
Kottayam, India
abhijithjthoppil2023@cs.ajce.in

Aswin Raj
Dept of CSE
Amal Jyothi College of Engineering
Kottayam, India
aswinraj2023@cs.ajce.in

Daan Jacob
Dept of CSE
Amal Jyothi College of Engineering
Kottayam, India
daanjacob2023@cs.ajce.in

Christy K Mathews
Dept of CSE
Amal Jyothi College of Engineering
Kottayam, India
christykmathews2023@cs.ajce.in

***Abstract*—A technology called Blind Aid is made to help people who are visually impaired navigate their surroundings. Users receive real-time aural input on their surroundings, including knowledge of adjacent impediments, using a mix of hardware and software. The system's hardware comprises of a compact, portable device with sensors that track the user's surroundings and evaluate them, as well as a speaker for audio output. Users receive a clear and succinct audio depiction of their surroundings because to the software component's usage of machine learning algorithms to understand and process the sensor data. Blind Aid helps to enhance the independence and quality of life of visually impaired people by enabling them to more simply and confidently navigate their surroundings.**

***Index Terms*—Visually impaired people, Visual Impairment, electronic navigation aid, Raspberry Pi, visual aid, wearable system.**

I. INTRODUCTION

Assistive systems for the visually impaired are designed to aid in navigating surroundings and performing daily tasks, and they can range from simple devices with sensors and audio output to advanced systems that utilize cutting-edge technologies like artificial intelligence and machine learning. The white cane, invented in the 1920s, was one of the first types of blind aid assistance systems and has been improved over time with features like reflective tape and built-in sensors. Recent technological advancements have enabled the development of more sophisticated systems that use machine learning algorithms to interpret sensor data or provide real-time auditory feedback. These systems frequently incorporate portable or wearable devices such as smart glasses or smartphones to provide quick and discreet assistance. The primary objectives of these systems are to enhance the quality of life and independence of visually impaired individuals, increase safety by avoiding obstacles and hazards, provide convenient access to assistance, and improve task efficiency.

II. LITERATURE SURVEY

Visual assistance for visually impaired people is an important area of research that aims to improve the lives of visually impaired individuals. In recent years, there has been a significant increase in research on this topic, as technological advancements have made it possible to develop innovative solutions that can help visually impaired people navigate their surroundings more easily. This literature survey presents a review of the research conducted on visual assistance for visually impaired people in IEEE conference papers. The paper [1] provides an overview of the different types of assistive technologies that have been developed to aid visually impaired individuals. The authors discuss various devices and software that can help visually impaired individuals with daily tasks such as reading, navigation, and communication. And in [2] it discusses about an overview of the different types of assistive technologies that have been developed to aid visually impaired individuals. The authors discuss various devices and software that can help visually impaired individuals with daily tasks such as reading, navigation, and communication. Indoor navigation is also an important aspect for the better accesibility fora visually impaired person, an inertial management unit based navigation is discussed in [3] that can assist visually impaired individuals with indoor navigation. The system uses a combination of IMU sensors and a smartphone to provide accurate positioning information to the user. In [4] the indoor navigation is further assisted by giving an auditory feedback input system that can assist visually impaired individuals with navigation in indoor environments. The system uses a combination of Bluetooth beacons and audio feedback to guide the user. Also as discussed in [5] a haptic feedback system that can assist visually impaired individuals with object recognition. The system uses a wearable device that provides haptic feedback to the user when it detects an object in the user's surroundings.

In paper [6] the authors have developed a machine learning-based object recognition system to assist visually impaired individuals in identifying objects in their surroundings. The system uses a convolutional neural network (CNN) algorithm to analyze the images captured by a camera mounted on the user's head. The CNN algorithm has been trained to identify various objects and can recognize them with a high degree of accuracy. Once an object is identified, the system provides an audio feedback to the user about what object is present in their surroundings. The authors evaluated the system's performance using a dataset of real-world objects and found that it achieved an accuracy of over 90% in identifying objects.The proposed system has several advantages over traditional object recognition techniques. It is faster and more accurate than previous approaches and can operate in real-time. The system is also lightweight and portable, making it easy for visually impaired individuals to use it in their daily lives. Overall, this system has the potential to greatly improve the independence and mobility of visually impaired individuals by helping them to more easily navigate their environment and recognize objects around them.

The proposed system's general methodology is shown in Fig. 1. The primary processor is a Raspberry Pi 4. The distance of an object or barrier in front of a person and the classification of the associated object are the major two visual parameters that may be accessible through the input collected from the camera. By reducing the noise surrounding the image and its size, preprocessing techniques can improve the image. The COCO dataset is used in conjunction with the TensorFlow API to classify objects. Via an ear-mounted device, a text to speech converter feeds the user information about the processed image. Fig. 2 depicts the hardware-software combination of the proposed system.

From the extensive literature survey conducted, we have identified the following architecture for the initial prototype.

III. Methodology

The proposed system's general methodology is shown in Fig. 1. The primary processor is a Raspberry Pi 4. The distance of an object or barrier in front of a person and the classification of the associated object are the major two visual parameters that may be accessible through the input collected from the camera. By reducing the noise surrounding the image and its size, preprocessing techniques can improve the image. The COCO dataset is used in conjunction with the TensorFlow API to classify objects. Via an ear-mounted device, a text to speech converter feeds the user information about the processed image. Fig. 2 depicts the hardware-software combination of the proposed system.

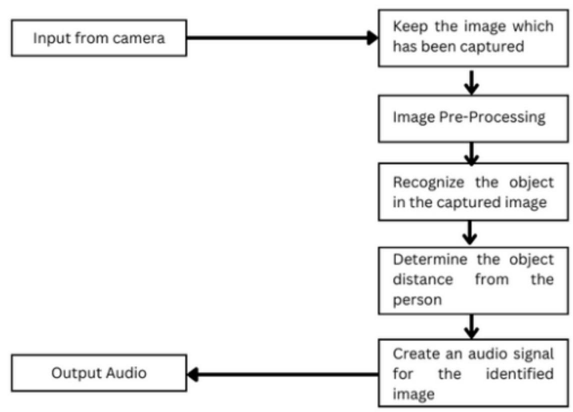

Fig. 1. Basic Architecture of Proposed System

A. Dataset Used

The dataset utilised here is a sample Tflite model from Google.[11]
A sample quantized SSDLite-MobileNet-v2 object identification model, trained on the MSCOCO dataset and optimised for TensorFlow Lite, is provided by Google. It is capable of detecting and recognising 80 typical objects, including people, cars, cups, and other items.
Depthwise separable convolution is the method of convolution that is employed here.
Depth-dependent separable convolution A collection of intermediate feature maps are created by first applying a depthwise convolution that applies a distinct filter to each input channel. Then, to create the final output feature map, a pointwise convolution applies a 1x1 convolution to these feature maps. The number of parameters needed for this procedure is decreased by the use of a single-channel filter in pointwise convolution.

B. Object detection

OpenCV and TensorFlow Lite are commonly utilized in object detection applications. TensorFlow Lite is a scaled-down variant of the TensorFlow platform, specially designed for deploying deep learning models on mobile devices and embedded systems that have limited resources. On the other hand, OpenCV is an open-source library that offers a broad range of image processing and computer vision tools.

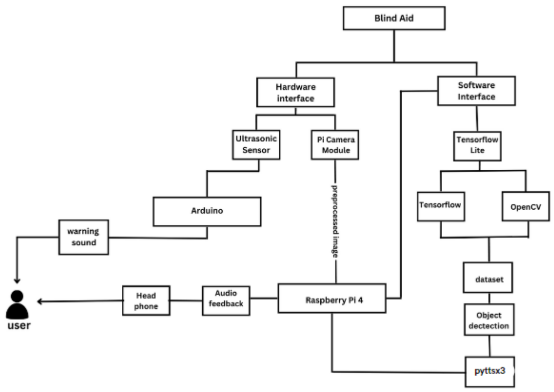

Fig. 2. Software and and Hardware Component visualization

Fig. 3. Object Detection

Fig. 4. Object Detection

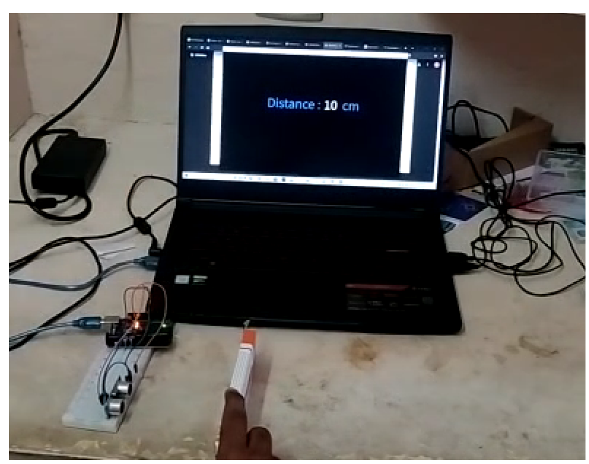

Fig. 5. Distance Analyzer

C. Distance Analyzer

To implement the distance analyzer, an Arduino is used as the functional device along with HC-SR04 ultrasonic sensors. The module is designed to emit a warning sound through an auditory feedback system to alert the user when a safer distance needs to be maintained.

D. Auditory Feedback

An image is captured from the live video feed without interrupting the object detection process. In the background, Tesseract API will extract the texts from the image and save them in a temporary text file. Then it reads out the text from the text file using the Python text-to-speech module.

IV. Expected Outcome

The expected outcomes for the Object Detection and Distance Analyzer for Visually Impaired People system is:

- Object Detection: Fig. 3 and Fig. 4 shows the detection of a single object from a video stream. Although most part of the image contains the background, the model is still able to filter out other bounding boxes and detect the desired object in the frame, with 73% confidence. The device can also detect multiple objects, with different confidence levels, from one video frame.
- Distance Analyzer: A warning or buzzer sound will be given to the user when the safe distance is reached or exceeded (Fig. 5) . A safe distance of 20cm is kept.
- Auditory Feedback: Auditory feedback by means of speaker or headphones are given to the user about the object or objects present in front of the user.
- Data analysis: The collected data can be analyzed to identify trends, patterns, and correlations, providing insights into the factors that influence the surrounding of the user.

V. Conclusion

In conclusion, blind aid assistance systems are a crucial tool for enhancing the independence and quality of life of people who are visually impaired. These systems can come in a variety of shapes, from straightforward devices with sensors and audio output to more intricate systems that incorporate cutting-edge technology like artificial intelligence and machine learning. Blind aid systems can help users navigate unfamiliar areas, carry out tasks more effectively, and live more freely and confidently by giving users real-time information about their surroundings and support with daily tasks. The cost and accessibility of technology, as well as the requirement for continual research to increase these systems' efficacy and usability, are some of the disadvantages and obstacles that these systems face. The potential for blind aid assistance technologies to significantly enhance the lives of visually impaired individuals, and will likely continue to evolve and advance as new technologies emerge.

References

[1] H. Sharma, M. Tripathi, A. Kumar, and M. S. Gaur, "Embedded assistive stick for visually impaired persons," in Proc. 9th Int. Conf. Comput., Commun. Netw. Technol. (ICCCNT), Jul. 2018, pp. 1–6, doi: 10.1109/ICCCNT.2018.8493707.

[2] A. Kocabas and M. Günay, "Assistive technology for individuals with visual impairments: A review," Assistive Technology, vol. 27, no. 2, pp. 65-76, 2015.

[3] D Ankur, N Gangotri, SP Shridevi, AG Ramakrishnan, "Low Resource End-to-end Sanskrit TTS using Tacotron2, WaveGlow and Transfer Learning," Proc. IEEE INDICON 2020.

[4] S. Quraishi and C. Shah, "A survey of assistive technologies for blind and low vision people," ACM Computing Surveys, vol. 52, no. 4, pp. 1-34, 2019.

[5] E. E. Ochen and G. O. Adukwu, "Assistive technologies for people with visual impairments: A review," Disability and Rehabilitation: Assistive Technology, vol. 13, no. 7, pp. 603-618, 2018.

[6] B. C. Koester, "Assistive technology for individuals who are blind or have low vision: A review," Journal of Visual Impairment & Blindness, vol. 110, no. 6, pp. 351-361, 2016.

[7] C. F. Hemmeter, "Assistive technology for people with visual impairments: A review," Journal of Visual Impairment & Blindness, vol. 104, no. 9, pp. 609-620, 2010.

[8] J. Q. Liu, "Assistive technology for individuals with visual impairments: A review," Journal of Visual Impairment & Blindness, vol. 102, no. 6, pp. 399-409, 2008.

[9] A. J. Kirwan and C. T. Best, "Assistive technology for people with visual impairments: A review," Journal of Visual Impairment & Blindness, vol. 101, no. 9, pp. 659-665, 2007.

[10] K. Naqvi, B. Hazela, S. Mishra, and P. Asthana, "Employing real-time object detection for visually impaired people," in Data Analytics and Management (Lecture Notes on Data Engineering and Communications Technologies), vol. 54, A. Khanna, D. Gupta, Z. Pólkowski, S. Bhattacharyya,

[11] Grechikhin, Ivan, and Andrey V. Savchenko. "User modeling on mobile device based on facial clustering and object detection in photos and videos." Pattern Recognition and Image Analysis: 9th Iberian Conference, IbPRIA 2019, Madrid, Spain, July 1–4, 2019, Proceedings, Part II 9. Springer International Publishing, 2019.