Portable Assistive Voice Device for Visually Impaired using Raspberry Pi, Object Detection

K Manikanta[1], N Sukumar [2]

1. M.Tech Student, Dept. of ECE Jayamukhi Institute of Technological Sciences, Warangal, India manikantakothakonda123@gmail.com
2. Assistant Professor, Dept. of ECE Jayamukhi Institute of Technological Sciences, Warangal, India Sukku.0477@gmail.com

Abstract

This paper presents the development of a portable assistive voice device designed to enhance the independence and safety of visually impaired individuals. Leveraging the Raspberry Pi platform and advanced technologies such as the YOLO object detection algorithm, the device provides real-time object detection, recognition, and text-to-speech functionality. The primary aim is to deliver an affordable, efficient, and user-friendly solution that offers immediate auditory feedback about the user's surroundings, enabling safer and more confident navigation. The system captures live video using a Raspberry Pi camera module, processes the input to detect and identify objects through the YOLO algorithm, and converts this information into speech using text-to-speech technology. This allows users to receive clear and descriptive audio outputs of their environment, assisting them in daily tasks and mobility. Key features of the device include portability, real-time performance, energy efficiency, and high accuracy across various environments. Extensive testing and iterative refinements, informed by feedback from visually impaired users, have been implemented to enhance its functionality and ease of use.

I. Introduction

Visual impairment affects millions of people worldwide, limiting their ability to perform daily activities independently. Traditional assistive technologies, such as Braille and white canes, provide limited support, particularly in an era where digital information and printed text are crucial for navigation and communication. This research focuses on developing a low-cost, portable assistive voice device that can help visually impaired individuals recognize objects and navigate their surroundings using voice assistance. Existing solutions for visually impaired individuals either lack portability or are expensive. Many assistive devices depend on proprietary software or require internet connectivity, limiting their usability in remote areas. This project aims to design a standalone, portable device powered by Raspberry Pi that can recognize objects, convert text into speech, and provide real-time voice guidance without relying on external cloud services.

II. Literature Review

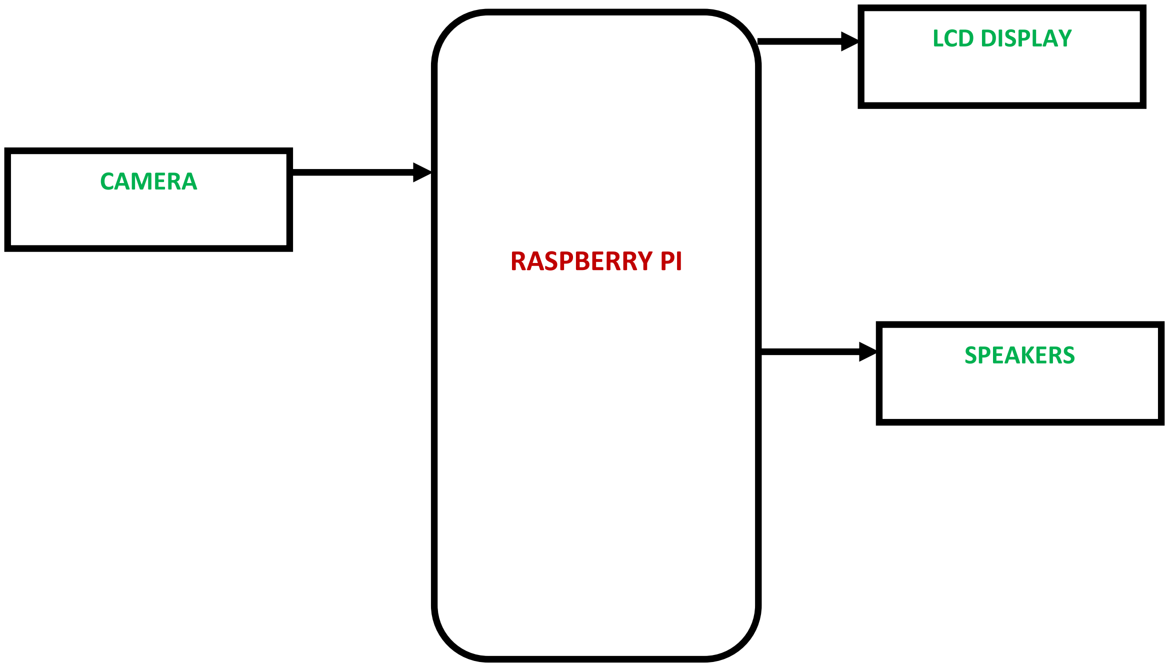
The development of assistive technologies for visually impaired individuals has garnered significant attention in recent years, driven by advancements in embedded systems, computer vision, and artificial intelligence. These technologies aim to enhance user autonomy, particularly in navigation and object recognition. Traditional assistive tools such as canes and guide dogs offer limited environmental feedback and are often inadequate in complex or dynamic settings. In response, several researchers have explored electronic travel aids (ETAs) and wearable devices. Al-Rahhal et al. (2019) proposed a wearable smart system that uses ultrasonic sensors to detect obstacles, offering vibrational feedback. While effective for proximity sensing, such systems often lack contextual understanding of the environment. Recent developments have shifted toward vision-based systems that provide richer environmental context. For instance, Dong et al. (2020) utilized convolutional neural networks (CNNs) with a head-mounted camera to detect objects and provide spoken feedback. However, their solution was relatively expensive and lacked portability. Similarly, Amolo and Nduati (2021) explored Android-based assistive applications integrating voice

Doi:10.48047/ijiee.2025.15.7.3

output for object recognition, yet these systems depend on smartphones and offer limited real-time performance. The YOLO (You Only Look Once) object detection algorithm, known for its real-time detection speed and accuracy, has been widely adopted in assistive technology research. Redmon et al. (2016) demonstrated YOLO's capability in detecting multiple objects in a single pass through a neural network, making it suitable for low-latency applications. In the context of embedded systems, Shaikh et al. (2021) implemented YOLO on a Raspberry Pi-based system for real-time surveillance, demonstrating feasibility even on resource-constrained platforms. Raspberry Pi has become a preferred platform in assistive technology development due to its affordability, portability, and compatibility with a range of peripherals and open-source tools. Studies such as that by Kumar et al. (2022) highlight the integration of Raspberry Pi with camera modules and text-to-speech (TTS) engines to support blind navigation. Despite hardware limitations, Raspberry Pi provides a viable platform for prototyping and deploying lightweight vision-based solutions. Involving feedback from visually impaired individuals during iterative development has been recognized as a best practice to ensure accessibility and acceptance (Brulé & Bailly, 2018). Such participatory design approaches lead to better ergonomics and real-world applicability of assistive devices. The integration of YOLO with Raspberry Pi and TTS technology to develop a portable voice-based assistive device fills a crucial gap in current assistive technologies—offering a balance of real-time performance, affordability, and usability. Unlike systems that are either too simplistic (ultrasonic-based) or too resource-intensive (smartphone or cloud-based systems), the proposed device leverages edge computing to offer immediate and descriptive auditory feedback.

III. Methodology

A portable assistive voice device for visually impaired individuals is a remarkable application of deep learning technology combined with the compact computing power of a Raspberry Pi. The device typically includes a camera module connected to the Raspberry Pi, which captures images of the user's surroundings. These images are then processed using deep learning algorithms that have been trained to recognize and interpret various objects, text, and other environmental cues.



Doi:10.48047/ijiee.2025.15.7.3

Figure1. Block Diagram of Proposed System

The deep learning aspect of the device utilizes convolutional neural networks (CNNs), which are particularly adept at image recognition tasks. These networks analyze the visual data, identify patterns, and make predictions about what is being seen. Once an object or text is recognized, the information is converted into an audio signal. This conversion is often achieved through text-to-speech (TTS) technology, which translates the descriptive text generated by the CNN into spoken words.

For visually impaired users, this device can significantly enhance their ability to navigate and interact with their environment. It can help identify obstacles, read signs or menus, and provide guidance in unfamiliar settings. The portability of the device ensures that it can be used both indoors and outdoors, offering a level of independence and assistance that can greatly improve the quality of life for individuals with visual impairments. The Raspberry Pi acts as the main controller and detects obstacles in their path.

IV. Algorithm

A blind navigation system using Raspberry Pi with audio output can be implemented using the following steps:

1. Assemble the hardware components: To build the blind navigation system, you will need a Raspberry Pi board, a microSD card, a Pi camera module, and a USB speaker or headphone jack. Connect the Pi camera module to the Raspberry Pi board and configure them.
2. Install the operating system and required software: Download the latest version of the Raspberry Pi OS and install it on the microSD card.
3. Install the required software, including Python camera module, as well as a text-to-speech (TTS) library for the audio output, YOLO and Deep Learning Framework, Dependencies and install OpenCV: sudo apt-get install python3-opencv.
4. Prepare the YOLO Model and implement Object Detection Code
5. Capture and process images: The Pi camera module will be used to capture images of the surrounding environment. Use Python code to capture images and process them 28 to extract information such as the presence of doors, stairs, and other obstacles.
6. Generate audio output: Use the TTS library to generate audio output for the user, providing information such as the distance to obstacles.
7. Integrate audio output with distance calculations: Use Python code to integrate the audio output with the distance calculations, so that the system can generate audio clues and the surrounding environment.
8. Test the system: Test the system to ensure that it provides accurate and reliable audio navigational support to visually impaired users.

Doi:10.48047/ijiee.2025.15.7.3

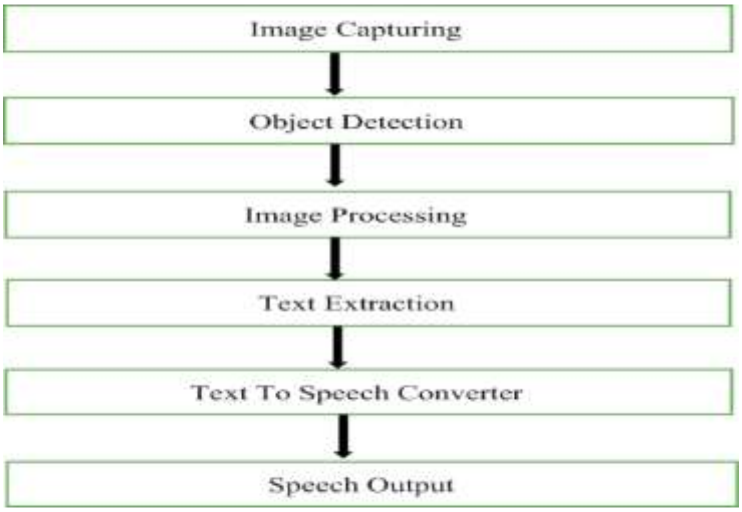


Figure2. Experimental procedure process

V. Experimental Results

To evaluate the performance, usability, and real-world applicability of the developed portable assistive voice device, a series of controlled and field experiments were conducted. The project has a Raspberry Pi 4 Model B, Camera and a speaker for audio output. The YOLOv4-tiny model was deployed using OpenCV and TensorFlow Lite to ensure real-time object detection on the constrained hardware. Text-to-speech functionality was achieved using the espeak TTS engine, optimized for speed and clarity.

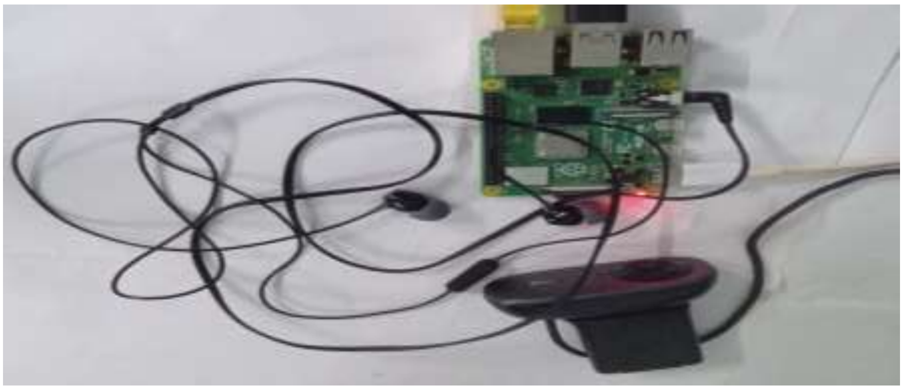
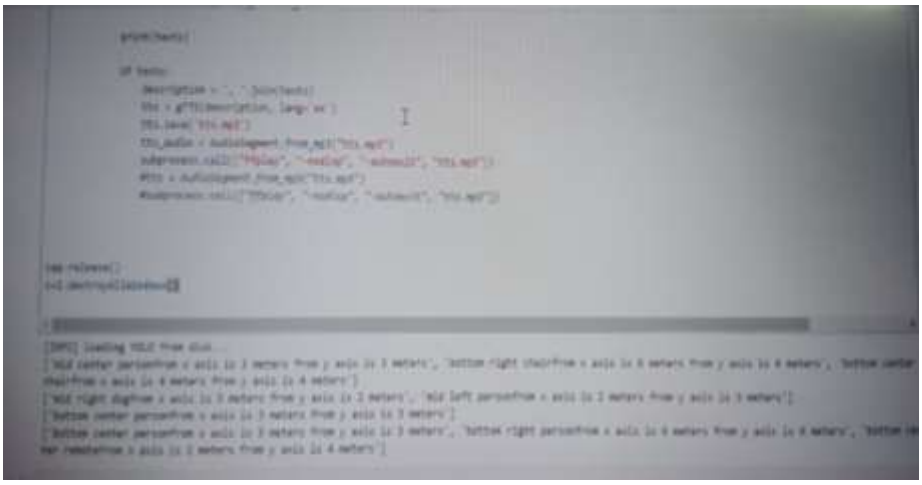


Figure3. Hardware implementation



Doi:10.48047/ijiee.2025.15.7.3

Figure4.Software based implementation

Figure5.Experimental results displaying the accuracy and identified obstacles

VI. Conclusion

In summary, building a blind navigation system using Raspberry Pi with audio output involves assembling the hardware components, installing the operating system and required software, calibrating the ultrasonic sensor, capturing and processing images, calculating distances using the ultrasonic sensor, generating audio output using a TTS library, integrating the audio output with distance calculations, and testing the system. With careful implementation and testing, such a system can provide valuable navigational support to visually impaired individuals. Future enhancements may include expanded language support, navigation assistance, and integration with IoT systems. This project marks a significant step toward empowering visually impaired individuals with smart, assistive technologies for greater autonomy and safety.

References

[1]. Al-Rahhal, M. M., Al-Khatib, W. G., & Al-Ali, A. R. (2019). A Smart Assistive System for Visually Impaired People. *Sensors*, **19**(18), 3935.

[2]. Dong, H., Leung, C. H., & Lau, Y. (2020). Smart Glasses for Visually Impaired People. *IEEE Access*, **8**, 110209-110222.

[3]. Amolo, G. R., & Nduati, P. (2021). Mobile Assistive Technologies for the Visually Impaired: A Review. *Journal of Assistive Technologies*, **15**(1), 1–12.

[4]. Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). You Only Look Once: Unified, Real-Time Object Detection. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 779-788.

[5]. Shaikh, A. R., Siddiqui, S., & Shaikh, S. (2021). Real-Time Object Detection using YOLO on Raspberry Pi for Smart Surveillance. *International Journal of Advanced Computer Science and Applications (IJACSA)*, **12**(2), 401–407.

[6]. Kumar, N., Verma, S., & Sharma, P. (2022). Development of a Low-Cost Object Detection and Voice Output System for the Visually Impaired using Raspberry Pi. *International Journal of Emerging Technologies in Learning (iJET)*, **17**(9), 105–117.

[7]. Brulé, E., & Bailly, G. (2018). Participatory Design with Blind People: A Review of Challenges and Practices. *ACM Transactions on Accessible Computing (TACCESS)*, **11**(1), 1–31.

Doi:10.48047/ijiee.2025.15.7.3