An AI-Based Assistive Vision Wearable System Using Raspberry PI Camera Module V2

V.R.Palandurkar[1] Mihir Jadhav[2] Tanvi Kale[3] Aarya Darwatkar[4] Aaditi Mokashi[5]
[1]Lecturer [2,3,4,5]Student
[1,2,3,4,5]Department of Information Technology
[1,2,3,4,5]AISSMS Polytechnic, Pune, India

Abstract — Visual impairment significantly affects an individual's independence and ability to interact safely with the surrounding environment. Traditional assistive tools provide limited environmental awareness and lack object identification capabilities. This paper presents the design and implementation of an assistive vision-based wearable system using Raspberry Pi and Camera Module V2. The system captures real-time video, processes the visual data using an object detection algorithm, and provides audio feedback to the user through a text-to-speech mechanism. The proposed system is compact, affordable, and capable of real-time operation without continuous internet dependency. Experimental evaluation demonstrates reliable object detection, acceptable response time, and improved situational awareness, making the system suitable for daily assistive use by visually impaired individuals.

Keywords: Assistive Technology, Smart Glasses, Raspberry Pi, Object Detection, Computer Vision, Text -to Speech.

I. Introduction

Visual impairment restricts a person's ability to recognize objects, navigate unfamiliar environments, and avoid obstacles, often leading to dependence on others. According to global health reports, a large population suffers from partial or complete vision loss. Traditional tools such as white canes help detect nearby obstacles but do not provide information about object type or context. Guide dogs, while effective, are costly and require long-term training.

Recent advancements in computer vision and embedded systems have enabled intelligent assistive devices capable of analyzing visual information in real time. Wearable vision-based systems offer a promising solution by providing environmental awareness through audio feedback. This paper focuses on developing a wearable smart glasses system using Raspberry Pi and Camera Module V2 to detect and identify objects in real time, thereby improving user safety and independence.

II. Problem Statement

Visually impaired individuals face difficulties in identifying objects and obstacles beyond immediate physical contact. Existing assistive tools provide limited feedback and lack semantic understanding of the environment. Smartphone-based applications depend heavily on internet connectivity and battery life, while sensor-based systems fail to identify object categories. Therefore, there is a strong need for a portable, real-time, vision-based assistive system that can operate independently, identify surrounding objects accurately, and provide immediate audio feedback to the user.

III. Literature Survey

Redmon et al. proposed a real-time object detection system using the YOLO (You Only Look Once) algorithm, which treats object detection as a single regression problem. The system achieved high detection speed and accuracy compared to traditional methods. However, the model required high computational resources, making it less suitable for low-power embedded wearable devices. [1]

Al-Adawi et al. developed a smart assistive system for visually impaired individuals using a camera-based vision system and audio feedback. The system was able to identify objects and provide spoken output. Although effective, the solution relied on external processing units, increasing system cost and reducing portability. [2]

Patil et al. presented a Raspberry Pi-based assistive device using ultrasonic sensors for obstacle detection. The system provided distance-based alerts to users through audio signals. While the system was simple and low-cost, it could not identify object types, limiting environmental understanding. [3]

Gupta et al. proposed a wearable smart glasses system using deep learning-based object recognition techniques. The system improved navigation assistance by detecting obstacles and objects in real time. However, the device consumed high power and required complex hardware, making it unsuitable for long-term wearable usage. [4]

Chen et al. introduced a vision-based assistive technology using a smartphone camera and cloud-based object detection. The system achieved high recognition accuracy but depended on continuous internet connectivity, leading to latency and privacy concerns. [5]

Kumar et al. designed an embedded vision system using Raspberry Pi and OpenCV for real-time image processing applications. The system demonstrated the feasibility of running vision algorithms on low-cost embedded platforms. However, object detection accuracy was limited due to the absence of optimized deep learning models. [6]

Singh et al. developed an assistive navigation system integrating computer vision and text-to-speech technology. The system provided real-time audio feedback to users.

Although promising, the system lacked a compact wearable design and required further optimization for real-time performance. [7]

IV. Comparison with Existing Systems and Novelty of Proposed Work

From the literature survey, it is observed that existing assistive systems suffer from several limitations such as reliance on ultrasonic sensors, dependency on cloud processing, high hardware complexity, and lack of portability. Sensor-based systems provide only distance information without object identification, while cloud-based approaches introduce latency and privacy concerns. High-end deep learning wearables require complex and power-intensive hardware, making them impractical for continuous daily use.

The proposed system overcomes these limitations by integrating a Raspberry Pi Camera Module V2 with a lightweight YOLO-based object detection model and text-to-speech technology into a single wearable smart glasses unit. Unlike earlier systems, the proposed solution performs real-time object detection and audio feedback completely offline, eliminating internet dependency. The use of embedded deep learning enables semantic understanding of the environment by identifying object types rather than just obstacle distance. Additionally, the system is compact, cost-effective, and optimized for low-power operation, making it suitable for long-term wearable assistive applications. This work represents a significant advancement by combining embedded vision, edge computing, and assistive audio feedback into a practical and user-friendly solution for visually impaired individuals.

V. SYSTEM ARCHITECTURE

The proposed system is designed as a wearable smart glasses unit integrating both hardware and software components into a single architecture. The system consists of a Raspberry Pi as the processing unit, Raspberry Pi Camera Module V2 for image acquisition, headphones for audio feedback, and a portable power supply. The software stack includes Raspberry Pi OS, Python programming language, OpenCV library for image processing, a YOLO-based object detection algorithm, and a text-to-speech module for audio output.

The camera continuously captures live video frames, which are transferred to the Raspberry Pi through the CSI interface. The Raspberry Pi processes each frame using computer vision techniques and object detection algorithms. Once an object is detected, its label is converted into speech and delivered to the user through headphones. This integrated architecture ensures real-time performance and portability.

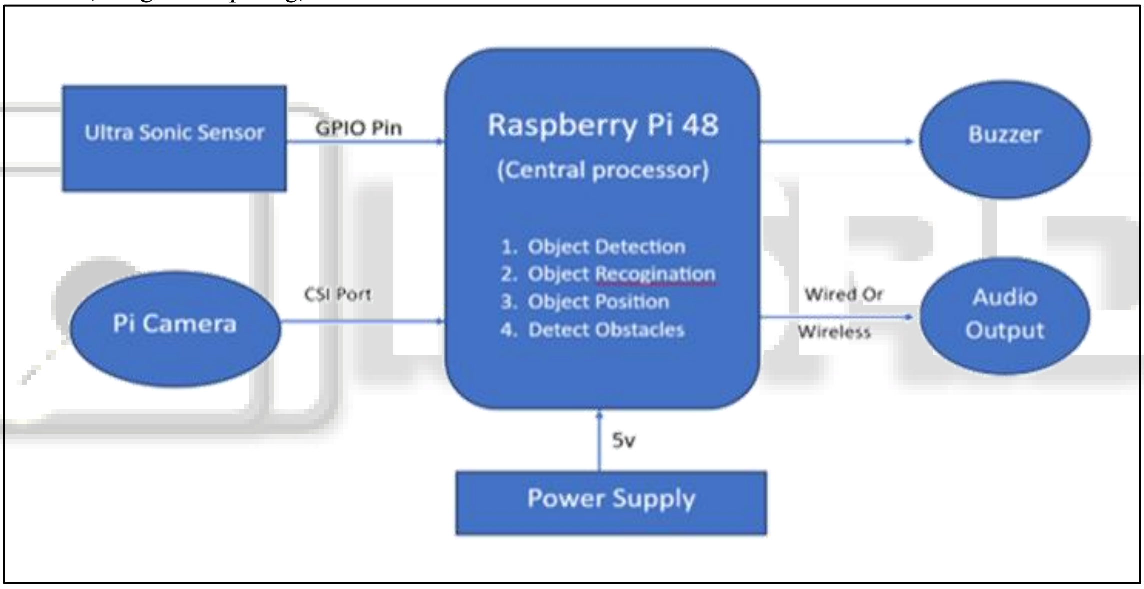


Fig. 1: Block diagram of proposed assistive vision-based wearable system.

VI. METHODOLOGY

The methodology of the proposed system is explained in detailed sequential steps as follows:

1) Step 1: System Initialization

The Raspberry Pi boots with Raspberry Pi OS and initializes all connected peripherals, including the camera module and audio output device. Required Python libraries such as OpenCV and text-to-speech modules are loaded.

2) Step 2: Video Frame Capture

The Raspberry Pi Camera Module V2 continuously captures live video frames. Each frame is extracted from the video stream for processing.

3) Step 3: Frame Preprocessing

Captured frames are resized to a fixed resolution to reduce computational load. Noise reduction and color space conversion are applied to enhance detection accuracy.

4) Step 4: Object Detection Processing

Each preprocessed frame is passed to the YOLO-based object detection model. The model analyzes the frame and identifies objects by generating bounding boxes and class labels.

5) Step 5: Object Classification and Filtering

Detected objects are filtered based on confidence thresholds to avoid false detections. Only relevant objects with high confidence scores are selected.

6) Step 6: Text Generation

The detected object labels are converted into text strings representing object names.

7) Step 7: Text-to-Speech Conversion

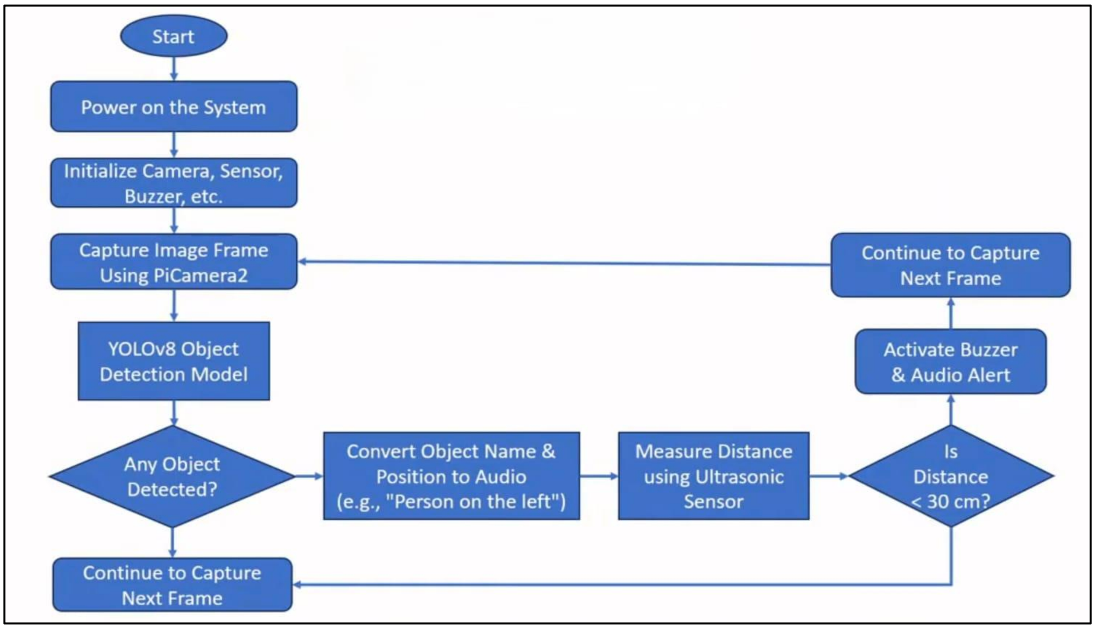
The text strings are processed by the text-to-speech module, which converts them into audible speech.

8) Step 8: Audio Feedback Delivery

The generated audio is delivered to the user through headphones, providing real-time information about surrounding objects.

9) Step 9: Continuous Monitoring

The system repeats the process continuously, ensuring real-time assistance as the user moves.



The working of the proposed assistive vision-based wearable system is illustrated in Fig. X. Initially, the system is powered on and all components such as the Raspberry Pi, camera module, audio output device, and sensors are initialized. The Raspberry Pi Camera Module captures real-time image frames using the PiCamera2 interface.

Each captured frame is processed using the YOLOv8 object detection model to identify objects present in the scene. If no object is detected, the system continues to capture the next frame. When an object is detected, the object name and its relative position are converted into audio output using a text-to-speech module, for example, “person on the left.”

Additionally, the system measures the distance of nearby obstacles using an ultrasonic sensor. If the measured distance is less than a predefined threshold, an audio alert or buzzer is activated to warn the user. After generating the audio feedback, the system continues capturing subsequent frames, enabling continuous real-time operation.

The working of the proposed assistive vision-based wearable system follows a sequential computer vision pipeline as illustrated in Fig. C. The process begins with data acquisition, where the Raspberry Pi Camera Module V2 captures real-time image or video frames of the surrounding environment. These frames are then subjected to pre-processing and image enhancement techniques using OpenCV to improve visual quality and reduce noise.

In the next stage, feature extraction is performed using a YOLO-based deep learning model, which extracts important visual features from the input frames. The extracted features are used for object detection and recognition, where objects present in the scene are detected and classified into predefined categories such as person, chair, bottle, or door.

Once the object is identified, classification and interpretation are carried out to determine relevant object information. Based on this information, the system performs decision making and generates audio output using a text-to-speech module. The detected object name is announced to the user through headphones, providing real-time auditory feedback.

The entire process is repeated continuously, enabling real-time object detection and assistance as the user moves through different environments.

VII. Result and Discussion

The proposed system was tested in multiple indoor and outdoor environments to evaluate its performance. Objects such as chairs, tables, doors, bottles, bags, and people were successfully detected and announced. The system demonstrated consistent detection accuracy under normal lighting conditions.

The response time between object detection and audio feedback was found to be suitable for real-time usage, allowing users to react promptly. The wearable design and low-cost hardware make the system practical for daily use. Unlike sensor-based systems, the proposed solution provides semantic understanding by identifying object types rather than just distance information.

The system operates without continuous internet connectivity, improving reliability and privacy. The integration of camera-based vision and audio feedback significantly enhances situational awareness, enabling visually impaired users to navigate environments more confidently and safely.

VIII. Conclusion

This paper presented a detailed design and implementation of an assistive vision-based wearable system using Raspberry Pi Camera Module V2. By combining real-time object detection with audio feedback, the system enhances environmental awareness for visually impaired individuals. The experimental results demonstrate that the system is accurate, responsive, affordable, and suitable for wearable assistive applications. The proposed solution offers a practical and efficient approach to improving independence and safety for visually impaired users.

References

[1] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You Only Look Once: Unified, Real-Time Object Detection," Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 779–788, 2016.
[2] M. Al-Adawi, A. Al-Hinai, and S. Al-Busaidi, "Smart Assistive System for Visually Impaired People Using Computer Vision and Audio Feedback," International Journal of Advanced Computer Science and Applications (IJACSA), vol. 9, no. 4, pp. 120–126, 2018.
[3] S. Patil, P. Kulkarni, and R. Patil, "Obstacle Detection System for Visually Impaired Using Raspberry Pi," International Journal of Engineering Research and Technology (IJERT), vol. 6, no. 3, pp. 45–49, 2017.
[4] R. Gupta, N. Sharma, and P. Verma, "Smart Glasses for Visually Impaired Using Deep Learning," International Journal of Computer Applications, vol. 182, no. 25, pp. 7–12, 2018.
[5] Y. Chen, X. Zhang, and L. Wang, "Vision-Based Assistive Technology Using Mobile Devices for Visually Impaired," IEEE Transactions on Human–Machine Systems, vol. 47, no. 3, pp. 390–401, 2017.
[6] A. Kumar and S. Rao, "Real-Time Image Processing on Raspberry Pi Using OpenCV," International Journal of Computer Science and Information Technologies (IJCSIT), vol. 7, no. 5, pp. 2341–2345, 2016.
[7] H. Singh, R. Kaur, and M. Singh, "Computer Vision-Based Assistive System with Text-to-Speech for Visually Impaired," International Journal of Emerging Technologies in Engineering Research (IJETER), vol. 6, no. 2, pp. 85–90, 2018.
[8] OpenCV Team, "Open-Source Computer Vision Library," [Online]. Available: https://opencv.org
[9] Raspberry Pi Foundation, "Raspberry Pi Camera Module V2 – Datasheet and Documentation," [Online]. Available: https://www.raspberrypi.com
[10] Python Software Foundation, "Python Programming Language Documentation," [Online]. Available: https://www.python.org
[11] Google Research, "YOLO: Real-Time Object Detection," [Online]. Available: https://pjreddie.com/darknet/yolo/