AI-Powered Object Detection for the Visually Impaired (AODVI)

Divya Potbhare[1], Tanuja Jadhav[2], Sameeksha Meher[3], Mrs. N. R. Dangi[4]

Student, Department of Computer Engineering [1-3]
Guide, Department of Computer Engineering[4]
Rasikalal M. Dhariwal Institute of Technology, Chinchwad

Abstract: *Visual impairment impacts a large number of people around the world and makes daily activities such as navigation and object identification difficult. While conventional aids such as white canes provide basic mobility support, they fall short of giving users a meaningful understanding of their immediate environment. This paper presents the AI-Powered Object Detection for the Visually Impaired (AODVI) system, which combines real-time computer vision with audio feedback to bridge this gap. The system uses a camera to capture live surroundings, processes the frames through a YOLO-based object detection model, and instantly relays detected object information to the user via a text-to-speech engine. By integrating GPS-based location awareness and an emergency alert mechanism, AODVI aims to offer a practical, affordable, and portable solution that enhances the independence, safety, and mobility of visually impaired individuals*

Keywords: Artificial Intelligence, Object Detection, Computer Vision, Deep Learning, YOLO, Assistive Technology, Text-to-Speech, Real-Time Processing, Accessibility, GPS Navigation

I. INTRODUCTION

Visual impairment is a growing global challenge. According to the World Health Organization, over 2.2 billion people worldwide live with some form of visual impairment, and a significant portion face difficulty in performing basic daily activities without assistance. Navigating unfamiliar roads, detecting obstacles in the path, or simply understanding the environment requires visual awareness that these individuals are unable to access through conventional means.

Traditional assistive tools — such as white canes and guide dogs — provide partial support. White canes help detect physical barriers at ground level but cannot identify objects at higher elevations or provide contextual information. Guide dogs, while effective, are not accessible to everyone due to cost and availability. These limitations have motivated researchers and engineers to explore technology-driven solutions.

Recent developments in artificial intelligence, especially deep learning and computer vision, have created new opportunities for developing assistive technologies. Real-time object detection models like YOLO (You Only Look Once) can identify multiple objects within a single camera frame in milliseconds, making them well-suited for time-sensitive assistive applications.

This paper presents the design and development of AODVI — a system that captures the user's environment through a camera, identifies objects using a YOLO model, and communicates findings through voice output. The system also incorporates GPS-based location awareness and an emergency alert feature to provide a comprehensive assistive experience.

II. LITERATURE SURVEY

A substantial body of research has explored real-time object detection and its application in assistive technologies. The following studies form the foundation upon which AODVI was conceptualized.

Jan Hacek [1] demonstrated the feasibility of using a smartphone camera to assist visually impaired users. His system processed captured images to extract useful environmental information and conveyed it to users through audio

Copyright to IJARSCT
www.ijarsct.co.in

DOI: 10.48175/568

191

ISSN
2581-9429
IJARSCT

feedback. The work highlighted that mobile hardware, even with limited processing power, can support real-time detection when paired with efficient algorithms.

Ren et al. [2] introduced Faster R-CNN, a significant milestone in object detection. By combining region proposal networks with convolutional neural networks, Faster R-CNN achieved high accuracy in detecting objects within complex scenes. Although computationally intensive, it established the benchmark for precision in detection systems.

Redmond et al. [3] proposed YOLO (You Only Look Once), which fundamentally changed the approach to object detection by treating it as a single regression problem. Unlike multi-stage detectors, YOLO processes entire images in one pass, enabling real-time performance that is critical for assistive systems.

Liu et al. [4] presented the Single Shot Multifood Detector (SSD), which offered a balance between detection speed and accuracy. SSD uses multiple feature maps at different scales, allowing it to detect objects of varying sizes efficiently — an important quality for real-world assistive applications.

SYSTEM ARCHITECTURE AND DESIGN

The AODVI system is built around four core functional modules that work in real time: image acquisition, object detection, location tracking, and audio feedback. Figure 1 illustrates the overall system architecture.
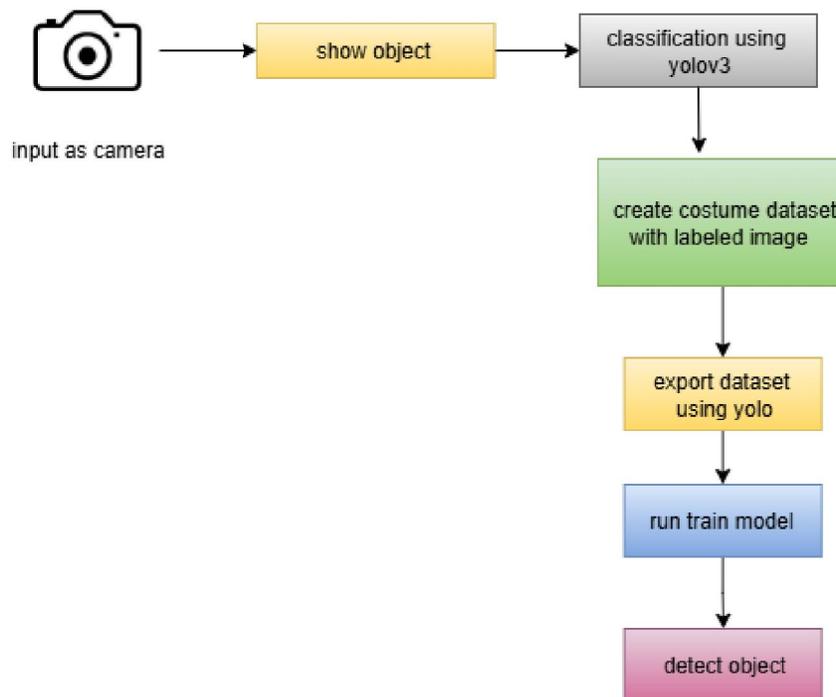
A. Graphical User Interface (GUI)

The interface is built using Python's Tinter library. It provides a clean main window with Start, Stop, and Exit controls, allowing users or caregivers to easily manage the application. A live camera feed is displayed within the GUI using the Opens library, along with real-time GPS coordinates and detection status.

B. Object Detection Module

Camera frames are continuously fed into the YOLO object detection model. The model identifies objects present in the frame, assigns confidence scores, and passes the results to the output module. YOLO was chosen for its exceptional speed-to-accuracy ratio, which allows the system to respond to environmental changes in real time.

C. Location Awareness Module

A GPS module continuously retrieves the user's geographic coordinates. These coordinates are displayed on the GUI and can be included in emergency alerts sent to registered contacts. This feature ensures caregivers or family members can locate the user if needed.

D. Voice Feedback Module

Detected object names and relevant alerts are passed to a text-to-speech (TTS) engine, which reads the information aloud to the user. This enables hands-free, eyes-free interaction — essential for users who rely entirely on auditory cues.

III. TECHNICAL REQUIREMENTS

Software Requirements
- **Platform:** Anaconda Navigator
- **Development Environment:** Spider IDE
- **Programming Language:** Python 3.x
- **Key Libraries:** Opens, YOLOv5/YOLOv8, pyttsx3 (TTS), Tinter, gaps

Hardware Requirements
- **Operating System:** Windows 10 / Windows 11
- **Processor:** Intel Core i5 or above
- **RAM:** 8 GB minimum
- **Storage:** 500 GB HDD / SSD
- **Camera:** USB webcam or integrated camera module

IV. SCOPE AND KEY FEATURES

The AODVI system is designed to serve as a day-to-day assistive companion for visually impaired individuals. Its scope extends across indoor and outdoor environments, and it is intentionally designed to be affordable and easy to operate without technical expertise.

The core features of the system include:

• Real-Time Object Detection and Obstacle Alerts — The system continuously analyses the camera feed using the YOLO model and immediately alerts the user via voice output when an obstacle or object is detected.

• GPS-Based Location Assistance — Real-time GPS tracking enables situational awareness, helping users understand where they are and allowing caregivers to monitor their location remotely.

• Voice-Based Feedback — All detection and location data is converted into natural-sounding speech, allowing users to receive information without requiring any visual interaction with the device.

• Emergency Alert System — When potentially hazardous conditions are detected or the user triggers an alert, the system automatically notifies registered contacts with the user's current GPS coordinates

ProjectImplementation Period: September – March

Activity	Month I 2025	Month II 2025	Month III 2025	Month IV 2025	Month V 2025	Month VI 2026	Month VII 2026	Month VIII 2026
Start date from Group Registration	JUNE 30	AUG 22	SEP 5	SEP 30	OCT 17	JAN 1	MAR 14	MAR 17
Submission of Synopsis (22 Aug 2025)		▨						
Submission of Survey Paper (05 Sept 2025)			▨					
Submission of Design Paper (30 Sept 2025)				▨				
Last Date to Submit Report SEM I (17 Oct 2025)					▨			
Submission of Result Paper with 100% Coding (01 Jan 2026)						▨		
Technical Event & IPR Activity (14 March 2026)							▨	
Last Date to Submit Final Report SEM II (17 March 2026)								▨
Completion of All Documents & Records								▨

V. CONCLUSION

This study introduced AODVI, an AI-based assistive solution developed to support visually impaired people in understanding and navigating their surroundings. more safely and independently. By combining real-time YOLO-based object detection with GPS location tracking and text-to-speech output, the system offers a practical, portable, and cost-effective solution that addresses the limitations of traditional assistive tools.

The integration of an emergency alert feature further strengthens the system's real-world applicability by connecting users with their caregivers in critical situations. The proposed approach is designed to run on standard consumer hardware, making it accessible to a wide user base.

Future work will focus on improving the accuracy of distance estimation between the user and detected objects, incorporating semantic navigation cues, and testing the system in diverse real-world environments. The authors also

plan to explore deployment on lightweight embedded hardware such as Raspberry Pi to make the system fully wearable.

REFERENCES

[1] B. Jan Hacek, "Application of a Camera in a Mobile Phone for Visually Impaired People," Master's thesis, Czech Technical University in Prague, May 2017.

[2] S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks," IEEE Transactions on Pattern Analysis and Machine Intelligence, Dec. 2016.

[3] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You Only Look Once: Unified, Real-Time Object Detection," in Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016.

[4] W. Liu, D. Anguelov, D. Erhan, C. Szegedy, S. Reed, C.-Y. Fu, and A. C. Berg, "SSD: Single Shot Multibox Detector," in Proc. European Conference on Computer Vision (ECCV), 2016.

[5] S. Hijazi, R. Kumar, and C. Rowen, "Using Convolutional Neural Networks for Image Recognition," IP Group, Cadence Design Systems, Technical Report, 2015