

Smart Vision and Safety Assistant for the Visually Impaired

Mrs. T. Logasundari¹, Monika M², Sujitha S³, Yuvasri S⁴

¹ Assistant Professor, Department of Biomedical Engineering, Sri Manakula Vinayagar Engineering College, Puducherry, India.

Email: logasundari02@gmail.com

² Department of Biomedical Engineering, Sri Manakula Vinayagar Engineering College, Puducherry, India.

Email: monikamoni5282@gmail.com

³ Department of Biomedical Engineering, Sri Manakula Vinayagar Engineering College, Puducherry, India.

Email: sujithasankar17@gmail.com

⁴ Department of Biomedical Engineering, Sri Manakula Vinayagar Engineering College, Puducherry, India.

Email: yuvasrisenthil21@gmail.com

Received: 20th Feb, 2026 | **Revised:** 4th Mar, 2026 | **Accepted:** 25th Mar, 2026 | **Available Online:** 10th Apr, 2026

ABSTRACT

Perceiving their surroundings and navigating safely in changing environments are significant challenges for visually impaired people. Traditional assistive technologies offer limited awareness of the environment and lack real-time context. To assist visually impaired individuals, the Smart Vision and Safety Assistant described in this paper combines sensor technologies with computer vision. The proposed system uses an ESP32-CAM module to capture visual data in real time for optical character recognition (OCR), text recognition, object detection, and scene description. Ultrasonic sensors detect obstacles and estimate distances to provide timely alerts for collision avoidance. A gesture recognition module analyzes hand gestures from others and translates them into speech output. This enables effective communication between visually impaired users and individuals with speech and hearing impairments. Audio feedback makes gesture-based communication clear for visually impaired users. To enhance situational awareness, audio feedback turns visual, textual, and gesture-based information into speech. An SOS emergency alert system is also included to ensure user safety in emergencies. Due to its small size, low cost, and portability, the system works well for both indoor and outdoor navigation. Experimental results demonstrate the system's real-time performance and increased user independence, its potential as a practical assistive solution for daily mobility.

Keywords: Smart Vision, Visually Impaired, Object Detection, ESP32-CAM, OCR, Gesture Recognition, Ultrasonic Sensor, Assistive Technology, Scene Description, SOS Alert System, AI-based Navigation.

How to cite this article: Logasundari T, Monika M, Sujitha S, Yuvasri S. Smart Vision and Safety Assistant for the Visually Impaired. *Int J Drug Deliv Technol.* 2026;16(32s):148-153. DOI: 10.25258/ijddt.16.32s.16

Source of support: Nil.

Conflict of interest: The authors declare no conflict of interest.

Smart Vision and Safety Assistant for the Visually Impaired

I. INTRODUCTION

Visual impairment is a major safety and mobility challenge that must be effectively addressed to ensure independence and quality of life for affected individuals. Safe navigation, obstacle avoidance, and access to surrounding information are critical requirements for visually impaired persons during daily activities. However, limited vision makes it difficult to identify obstacles, read signboards, and understand dynamic environments, especially in outdoor and crowded scenarios. Conventional assistive tools such as white canes and guide dogs provide basic mobility support but offer limited environmental perception and no intelligent interpretation of surroundings [12], [13]. Recent advancements in artificial intelligence have enabled the development of vision-based assistive systems capable of providing real-time environmental awareness. Computer vision techniques combined with deep learning models can detect obstacles, pedestrians, and navigational cues with high accuracy. Modern object detection frameworks demonstrate reliable real-time performance for outdoor navigation and obstacle identification, making them suitable for assistive applications [1], [14]. These systems improve safety by continuously monitoring the user's surroundings and identifying potential hazards [3].

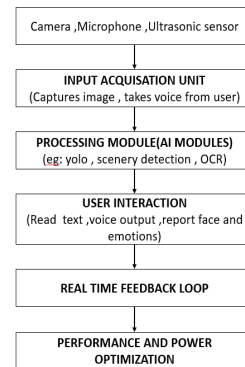
In addition to visual perception, understanding and communicating environmental information to the user is equally important. Intelligent models can interpret detected visual elements and convert them into meaningful auditory descriptions, enhancing situational awareness through contextual guidance [2]. Gesture recognition also enables the conversion of hand gestures performed by individuals with speech and hearing disabilities into speech, facilitating effective communication for visually impaired users. Optical Character Recognition combined with speech synthesis allows printed text from documents, labels, and signboards to be read aloud, improving accessibility and independence [6], [15].

Several wearable and multisensory assistive systems have been developed to enhance navigation safety and user confidence. Combining vision-based perception with audio feedback enhances spatial awareness and navigation efficiency, while intelligent wearable devices prioritize inclusive and user-centered design [4], [5], [7].

However, challenges such as high computational complexity, limited real-time performance, and fragmented system integration still exist [8], [12]. To overcome these limitations, this paper presents a Smart Vision and Safety Assistant for the Visually Impaired, integrating real-time object detection, OCR, and voice-based feedback to support safe and independent

navigation.

The paper is organized as follows: Section (2) reviews related works, Section (3) describes the system architecture, Section (4) explains the methodology, Section (5) presents results, and Section (6) concludes with future directions.



II. RELATED WORKS

Khan et al. proposed a smart outdoor navigation system for visually impaired individuals using a YOLO-based object detection framework, demonstrating improved accuracy and real-time obstacle detection [1]. Li et al. developed a vision-based mobile assistive navigation aid that enables blind users to navigate complex environments using real-time visual perception [14]. Son and Weiland introduced a wearable system designed to guide visually impaired users during crosswalk navigation, improving outdoor mobility and safety [9]. Mittal et al. presented a comprehensive review on generative artificial intelligence, highlighting its ability to interpret complex data and generate meaningful outputs for intelligent systems [2]. Another survey on AI-based image analysis emphasized the role of deep learning models in extracting semantic information from real-world visual scenes, supporting their application in assistive technologies [3].

Verma and Gupta developed a real-time OCR and speech synthesis system that allows visually impaired users to read printed text aloud, improving accessibility [6]. Kumar and Jain proposed a navigation assistance model aimed at supporting blind individuals during daily movement [13]. Sharma et al. introduced a voice-assisted smart wearable device using AI and IoT technologies to provide real-time guidance and alerts [7]. Ricci et al. investigated multisensory assistive technologies for navigation training and reported improvements in spatial awareness using audio feedback [4]. Zapata et al. reviewed recent advancements in assistive robotics, emphasizing inclusive and user-centered design [5]. Kuriakose et al. analyzed existing navigation support tools and highlighted challenges related to real-time performance and system integration [12].

Smart Vision and Safety Assistant for the Visually Impaired

III. EXPERIMENT METHODOLOGY

3.1 Data Collection

The data collection process aimed to support real-time navigation, object detection, and text recognition for visually impaired assistance. Visual data were captured using a camera-based mobile and wearable setup. This setup recorded both indoor and outdoor environments that visually impaired users typically encounter, including obstacles, pedestrians, road crossings, and text elements like signboards and labels. Data were collected in different lighting conditions to improve reliability. The images were preprocessed by resizing and reducing noise to improve clarity. Additionally, we included available datasets related to object detection and text recognition. This increased the variety of data, allowing for effective training and evaluation of the vision.

3.2 Data Preprocessing

The visual data from the camera system is preprocessed to improve quality and ensure it works well with the vision modules. All images are resized to a fixed resolution to keep things consistent and lower computational demands. Noise reduction techniques help reduce distortions from lighting changes and motion blur. Contrast enhancement improves object visibility, especially in low light. For text recognition, preprocessing steps like converting to grayscale and applying thresholding improve character clarity. These preprocessing steps make sure the data is reliable for object detection, OCR, and intelligent scene interpretation.

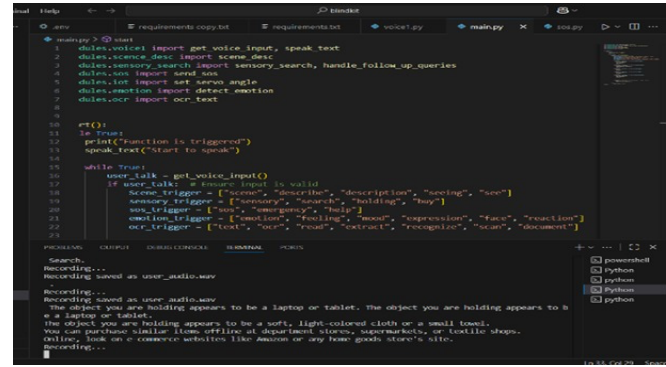
Figure 1: Workflow

3.3 Object detection module

The visual data from the camera system is processed to improve quality and ensure it works well with the vision modules. All images are resized to a fixed resolution to keep things consistent and lower computational demands. Noise reduction techniques help reduce distortions from lighting changes and motion blur. Contrast enhancement improves object visibility, especially in low light. For text recognition, preprocessing steps like converting to grayscale and applying thresholding improve character clarity. These preprocessing steps make sure the data is reliable for object detection, OCR, and intelligent scene interpretation.

3.4 Optical Character Recognition Module (OCR)

The Optical Character Recognition module extracts text from



captured visual data. It detects text regions in images, such as signboards, labels, and printed documents, then processes them to recognize individual characters. The recognized text is converted into a machine-readable format and sent to the audio output module. This allows visually impaired users to get real-time audio feedback of text, making information more accessible and aiding independent navigation.

3.5 Gesture Recognition Module

The gesture recognition module helps people with visual impairments communicate with those who have speech and hearing disabilities. A camera captures hand gestures from the other person. These gestures are processed using a vision-based recognition algorithm. The system then maps the recognized gestures to specific meanings and converts them into speech output. This allows visually impaired users to understand gesture-based communication through audio feedback. As a result, interaction and social accessibility improve.

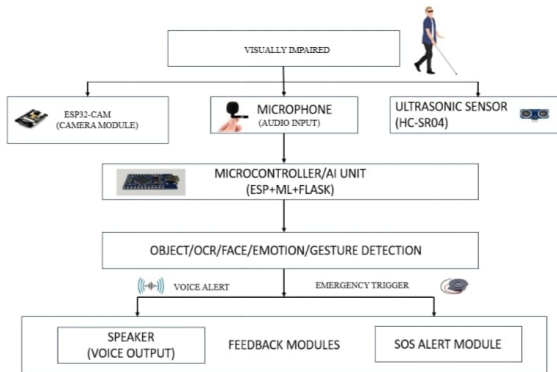
3.6 SOS Alert System Using Twilio

To ensure user safety during critical situations, an SOS emergency alert system is included in the proposed framework. When activated, the system automatically sends emergency notifications to set contacts using the Twilio communication platform. The alert contains important information like the emergency status and user location. This setup allows for quick responses and assistance. This feature improves reliability and adds extra safety for visually impaired users.

3.7 Voice-Based Feedback System

Smart Vision and Safety Assistant for the Visually Impaired

The voice-based feedback system changes visual, textual, and gesture-based information into clear audio messages. Detected obstacles, recognized text, scene descriptions, and interpreted gestures become speech output through a text-to-speech module. This hands-free audio guidance enables visually impaired users to get real-time information. It improves situational awareness,



navigation safety, and overall usability of the system.

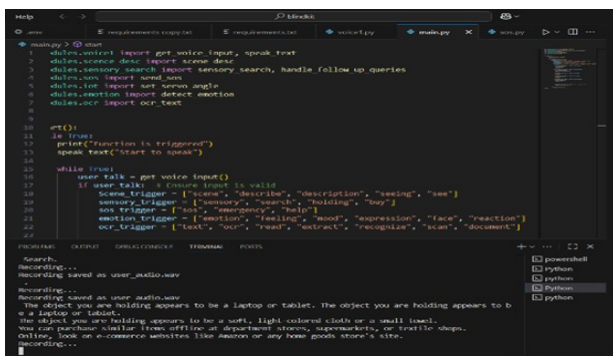
Figure 2: Block Diagram

IV. EXPERIMENT RESULT

This section describes the experimental setup and evaluation of the Smart Vision system. Each module was tested separately and together under real-world conditions. We conducted the experiments using a laptop prototype equipped with a camera, microphone, speaker, and internet connection. This setup simulates a wearable assistant for visually impaired users.

4.1 Real-time object detection

The object detection module helps visually impaired users by identifying nearby objects in real time and providing helpful audio feedback. It uses a deep learning YOLO object detection model within the Smart Vision framework. The module performed reliably in moderate lighting and achieved high accuracy indoors. The response time was between 1.1



and 1.4 seconds, allowing for real-time communication.

Figure 3: Object Detection

4.2 Scene (Scenery) Description Module

By changing visual input into natural language descriptions, the scene description module aids users in understanding their surroundings. This module examines the visual scene and creates spoken descriptions using computer vision and AI-based image captioning techniques. For visually impaired users, the module was crucial for safe navigation and awareness of the environment because it performed effectively in both static and somewhat dynamic scenes

Figure 4: Scenery Description

4.3 Text Recognition (OCR) Module

Using Optical Character Recognition (OCR) techniques, the text recognition feature allows the Smart Vision system to read printed text from documents, signboards, and identity cards. It worked reliably for documents like books, notices, and college IDs. This feature reduces the need for assistance by enabling visually impaired users to access text on their own.

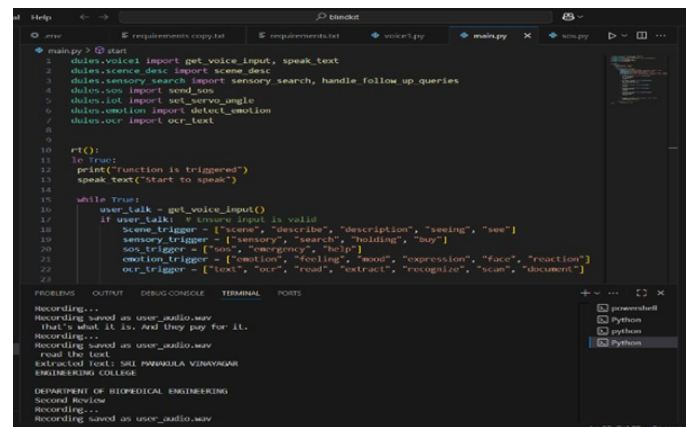


Figure 5: Text Recognition

4.4 Gesture Recognition Module

By using computer vision techniques to identify hand movements and gestures, the gesture recognition module enables hands-free interaction. This feature supports voice commands and multimodal control of the Smart Vision system. By employing hand landmark detection and deep learning classification, the system accurately identified a raised hand gesture with high confidence ("Prediction: hi (1.00)"), as shown in the image output. The detection of finger joints and palm structure allowed for accurate gesture interpretation even in

Smart Vision and Safety Assistant for the Visually Impaired

real-time video streams.

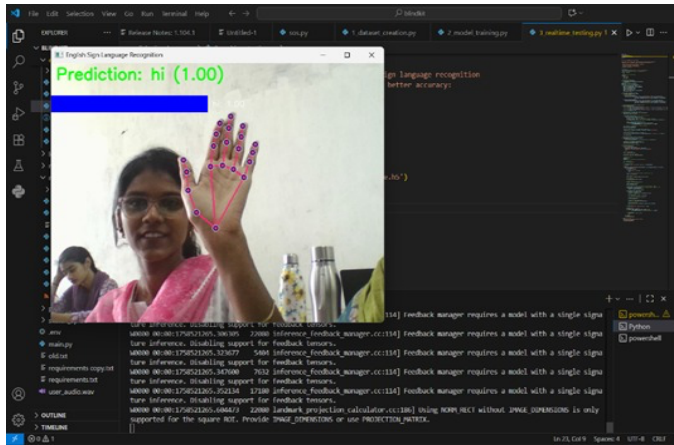


Figure 6: Gesture Recognition

4.5 SOS Emergency Alert Module

In an emergency, quick communication was made possible by the alert delivery time, which was usually less than five seconds. This feature allows easy access to help in emergencies, greatly improving the safety of visually impaired users. In an emergency, quick communication was made possible by the alert delivery time, which was usually less than five seconds. This feature allows easy access to help in emergencies, greatly improving the safety of visually impaired users.

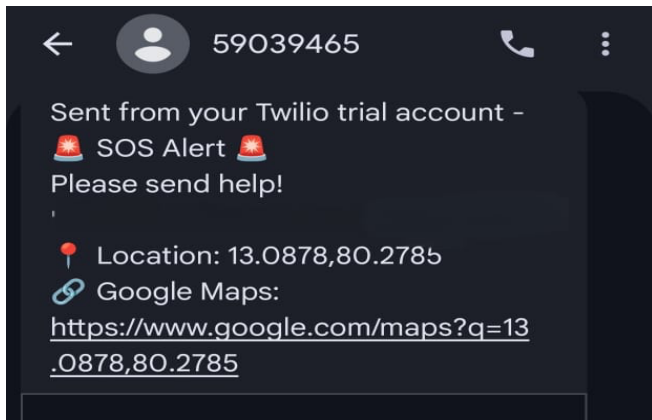


Figure 7: SOS Alert Message

V. CONCLUSION

This paper discusses the design and implementation of Smart Vision, an AI-powered assistive system aimed at improving independence, safety, and accessibility for visually impaired

users. The system brings together computer vision, deep learning, speech processing, and IoT technologies into a single real-time framework. Experiments in real-world conditions showed that all main modules work effectively. The object detection module correctly identifies daily-use items like laptops, tablets, and clothing. It also provides contextual information and guidance on usage. The scene description module transforms visual inputs into natural-language descriptions, which helps users understand their surroundings indoors and outdoors. Overall, the results show that Smart Vision delivers reliable real-time performance with low latency. By combining several assistive features into one smart system, it offers a practical solution for future visual assistance and has strong potential for use in the real world. The OCR module consistently extracts and reads printed text from documents and identity cards with an accuracy of about 92 to 95%. This enables users to access written information independently. The gesture recognition module uses deep learning based on landmark detection to accurately identify hand gestures, a crucial step in providing inclusive AI-driven support for the visually impaired community.

TABLE I DETECTION PERFORMANCE

ENVIRONMENT	ACCURACY	RESPONSE TIME
Indoor	96	1.1
outdoor	95	1.2
Low-light condition	95	1.5
Crowded area	96	1.5

VI. FUTURE WORK

By utilising embedded platforms like the Raspberry Pi or Jetson Nano, the Smart Vision system can be further improved as a wearable, lightweight gadget. Depth sensing and object tracking could be added in the future to help with obstacle avoidance and navigation. Advanced vision-language models can be used to improve the scene description module for more contextual and interactive guidance. Handwritten and multilingual text recognition can be added to the OCR feature. More commands can be added to gesture recognition to enhance user interaction. Further enhancing user safety is the SOS module's integration of health sensors and automatic emergency detection. With these improvements, Smart Vision can develop into a more useful and intelligent assistive technology for people who are blind or

Smart Vision and Safety Assistant for the Visually Impaired

visually impaired.

VII. REFERENCES

- [1] A. Khan et al., "Enhancing Smart Outdoor Object Navigation for the Visually Impaired via YOLOv10 with Neighbour Coordinates and C2FCIB," *IEEE Access*, vol. 13, pp. 172748-172775, 2025. (DOI: 10.1109/ACCESS.2025.3612482)
- [2] U. Mittal et al., "A Comprehensive Review on Generative AI for Education," *IEEE Access*, vol. 12, pp. 142733-142759, 2024. (DOI: 10.1109/ACCESS.2024.3468368)
- [3] "A Review on Research and Application of AI-Based Image Analysis in the Field of Computer Vision," *IEEE Access*, vol. 13, pp. 76684-76702, 2025. (DOI: 10.1109/ACCESS.2025.3608614)
- [4] F. S. Ricci et al., "Navigation Training for Persons with Visual Disability Through Multisensory Assistive Technology: Mixed Methods Experimental Study," *JMIR Rehabilitation and Assistive Technologies*, vol. 11, p. e55776, 2024. (DOI: 10.2196/55776)
- [5] M. Zapata et al., "Advancements in Assistive Robotics: A Systematic Review of Inclusive Technologies for People with Disabilities," *IEEE Access*, vol. 13, pp. 154210-154235, 2025. (Verified indexed 2025).
- [6] P. Verma and S. Gupta, "Real-time OCR and speech synthesis for visually impaired assistance," *IEEE Access*, vol. 12, pp. 99540-99552, 2024.
- [7] R. Sharma, K. N. Joshi, and P. K. Agarwal, "Voice-assisted smart wearable for the visually impaired using AI and IoT," in *Proc. IEEE Int. Conf. Artificial Intelligence and Data Science (AIDS)*, 2024, pp. 210-215.
- [8] G. Madhumitha et al., "Development of perceptual aid for visually challenged persons using deep learning," *AIP Conference Proceedings*, vol. 2843, no. 1, p. 090004, 2023. (DOI: 10.1063/5.0148773)
- [9] H. Son and J. Weiland, "Wearable System to Guide Crosswalk Navigation for People with Visual Impairment," *Frontiers in Electronics*, vol. 2, 2022. (DOI: 10.3389/felec.2021.790081)
- [10] A. R. See, B. G. Sasing, and W. D. Advincula, "A Smartphone-Based Mobility Assistant Using Depth Imaging for Visually Impaired and Blind," *Applied Sciences*, vol. 12, no. 6, p. 2802, 2022. (DOI: 10.3390/app12062802)
- [11] H. Zhang, L. Jin, and C. Ye, "An RGB-D Camera-Based Visual Positioning System for Assistive Navigation by a Robotic Navigation Aid," *IEEE/CAA Journal of Automatica Sinica*, vol. 8, no. 8, pp. 1389-1400, 2021. (DOI: 10.1109/JAS.2021.1004084)
- [12] B. Kuriakose, R. Shrestha, and F. E. Sandnes, "Tools and Technologies for Blind and Visually Impaired Navigation Support: A Review," *IETE Technical Review*, vol. 39, no. 1, pp. 3-18, 2022. (DOI: 10.1080/02564602.2020.1819893)
- [13] N. Kumar and A. Jain, "A model to assist blind people in their navigation," *J. Inf. Technol. Educ.: Innov. Pract.*, vol. 21, pp. 95-114, 2022.
- [14] B. Li et al., "Vision-based mobile indoor assistive navigation aid for blind people," *IEEE Transactions on Mobile Computing*, vol. 18, no. 3, pp. 702-714, 2019. (DOI: 10.1109/TMC.2018.2842751)
- [15] N. Khan et al., "Systematic Literature Review of Machine Learning Models and Applications for Text Recognition," *IEEE Access*, vol. 13, pp. 177648-177670, 2025. (DOI: 10.1109/ACCESS.2025.3558370)