AI Visual & Reading Aid for the Blind

Godwin C V[1], Hiba Afzel[2], Kirandas K[3], Subeena K S[4], Nivya K Venu[5]

[1,2,3,4]*Students, Department of Electrical and Electronics Engineering, IES College of Engineering, Chittilappilly, Kerala, India*

[5] *Assistant Professor, Department of Electrical and Electronics Engineering, IES College of Engineering, Chittilappilly, Kerala, India*

Email_id: godwincv770@gmail.com, hibaafzelieee@gmail.com , kirandask001@gmail.com, subeenasukumaran@gmail.com, nivyakvenu@iesce.info

Abstract

Blindness greatly restricts a person's ability to move independently, recognize their surroundings, and access written information, making routine activities difficult without external help. The AI-Based Visual Aid with Integrated Reading Assistant is a wearable assistive solution created to support fully blind individuals by merging intelligent navigation functions with real-time reading assistance. Developed using the power-efficient Raspberry Pi 4, the device combines a camera module and ultrasonic sensors to locate objects, estimate distance, and provide voice feedback without requiring user intervention. Using the TensorFlow SSD Lite-MobileNet deep-learning model, the system identifies objects in real time while estimating proximity through a fusion of camera-based interpretation and ultrasonic measurements. To further improve accessibility, the system includes a Tesseract-based OCR module capable of detecting text in captured images and converting it into synthesized speech. The lightweight, hands-free design can be mounted on standard eyeglass frames, allowing comfortable and portable usage. Performance evaluation with 60 fully blind participants in controlled environments showed noticeable improvements in navigation speed and confidence when compared with conventional white-cane use. Participants also highlighted strong obstacle-detection accuracy and reliable reading performance under proper lighting. The modular architecture enables future enhancements such as advanced OCR capabilities, expanded object datasets, slippery-surface alerts, stair recognition, and GPS-assisted guidance. Overall, the project contributes to the development of affordable, intelligent assistive tools that enhance autonomy, safety, and daily accessibility for individuals with visual impairments.

Keywords: AI-based visual aid, wearable navigation system, object detection, ultrasonic distance measurement, Raspberry Pi 4, SSD Lite-MobileNet

DOI: https://doi.org/10.5281/zenodo.18844050

1. Introduction

Blindness remains one of the most significant disabilities globally, severely restricting a person's ability to interpret their surroundings and interact safely with the world around them. Reports indicate that more than 2.2 billion individuals experience some form of visual impairment, and many live with complete loss of sight, making everyday tasks such as navigating unfamiliar spaces, recognizing obstacles, and accessing printed information extremely challenging. While traditional aids like the white cane offer essential support, they are limited in range, cannot detect elevated or distant hazards, and provide no awareness of approaching objects. Alternatives such as guide dogs are helpful but costly and inaccessible to many people.

Technological progress has led to the development of several electronic travel aids (ETAs) designed to improve mobility for visually impaired individuals. These tools typically incorporate sensors, imaging modules, and auditory feedback, yet many struggle with high cost, complex handling, limited portability, or the absence of integrated reading features.

To address these issues, the AI-Based Visual Aid with Integrated Reading Assistant introduces a compact, low-power wearable device capable of autonomous obstacle identification, distance estimation, and text-to-speech conversion. Utilizing a Raspberry Pi 4, a camera module, and ultrasonic sensors, the system applies deep-learning techniques—specifically the SSD Lite-MobileNet model—to detect objects in real time while determining their distance from the user. A built-in OCR-based reading tool enables conversion of printed text into speech, offering greater independence in everyday tasks. This introduction highlights the growing necessity for advanced assistive solutions and establishes the foundation for the proposed system aimed at improving safety and autonomy for completely blind individuals.

2. Objectives

The main objective of the AI-Based Visual Aid with Integrated Reading Assistant is to create an affordable, wearable, and highly capable assistive device that enhances the mobility, independence, and safety of individuals who are completely blind. The system is designed to overcome the common shortcomings of traditional support tools such as white canes, which cannot identify elevated obstacles, detect hazards at a distance, or provide any assistance for reading printed material.

A key goal is to develop a lightweight, hands-free device that can easily be attached to regular eyeglasses, ensuring comfort and usability across indoor and outdoor environments. The project aims to utilize the Raspberry Pi 4 platform to efficiently run advanced algorithms while maintaining low energy usage and keeping hardware costs minimal.

Another important objective is to integrate real-time object detection and distance estimation through a combination of a camera module and ultrasonic sensors. By using the SSD Lite-MobileNet deep-learning architecture, the system is intended to recognize multiple objects in the user's surroundings and estimate their proximity to support safer navigation.

Additionally, the project aims to implement a reading assistant using Tesseract OCR technology to extract text from documents and convert it into speech, enabling blind users to read independently. The final goal is to validate system performance through testing with completely blind participants, focusing on comfort, mobility, reading accuracy, and overall usability, while identifying potential areas for enhancement.

3. Methodology

The methodology for the proposed visual-aid system is adapted and expanded from the framework described in the base reference paper. The system architecture is organized into three major stages: data acquisition, feature extraction and processing, and user feedback generation. Data acquisition is performed using a Raspberry Pi 4 connected to a Pi Camera Module V2 and an HC-SR04 ultrasonic sensor. The camera continuously captures RGB video frames, while the ultrasonic sensor supplies supplementary distance information for nearby objects. Both data streams are transferred to the Raspberry Pi through dedicated GPIO configurations to enable real-time processing.

In the second stage, object detection and text recognition are carried out. The system employs the TensorFlow SSD Lite-MobileNet model for object detection due to its efficient structure and suitability for embedded platforms with limited processing capability. Each frame is fed through a convolutional neural network, which produces bounding boxes and corresponding object classifications. At the same time, the reading module captures a still image whenever activated, applies preprocessing steps, and uses the Tesseract OCR engine supported by an LSTM-based recognition model—to interpret text under varying lighting conditions.

The final stage focuses on generating audio feedback for the user. The eSpeak text-to-speech engine converts detected object names, their estimated distances, and extracted text into spoken output delivered through headphones. System evaluation was conducted in controlled indoor environments with blind participants, examining navigation speed, reading accuracy, ease of use, and overall performance. For comparison, key navigation metrics were measured against traditional white-cane use to assess the system's effectiveness.

3.1. Block Diagram

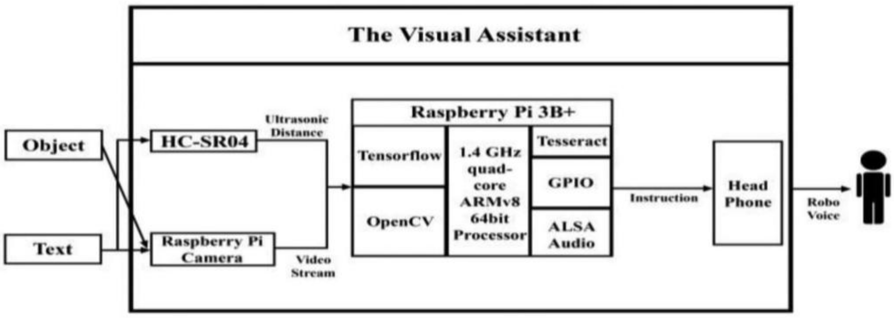

Figure 1: Block Diagram of Visual Assistant

3.2. Working Principle

The proposed visual-aid system operates by combining real-time image processing, sensor-based distance measurement, and audio feedback to assist completely blind users. The device is mounted on a pair of eyeglasses and uses a Raspberry Pi 4 as the central processing unit. A Pi Camera Module V2 continuously captures live video frames, while an ultrasonic sensor measures the distance between the user and nearby obstacles.

Each captured frame is processed through a lightweight deep-learning model SSDLite-MobileNet implemented using the TensorFlow Object Detection API. This model extracts features from the video stream, identifies common objects, and draws bounding boxes around them. Objects detected in each frame are classified based on the model's training on the COCO dataset. For improved reliability, the ultrasonic sensor simultaneously provides accurate short-range distance information that complements camera-based detection. The system prioritizes the closest object and generates audio alerts for user safety.

The device also includes an integrated reading assistant. When triggered, the system captures a still image, preprocesses it, and uses the Tesseract OCR engine (LSTM-based) to convert printed text into digital characters. The extracted text is then converted into speech using the eSpeak text-to-speech engine. This enables the user to read documents hands-free.

Finally, all processed information object type, distance, or recognized text—is communicated to the user through headphones as real-time audio feedback, allowing safe navigation and practical reading assistance.

Example:

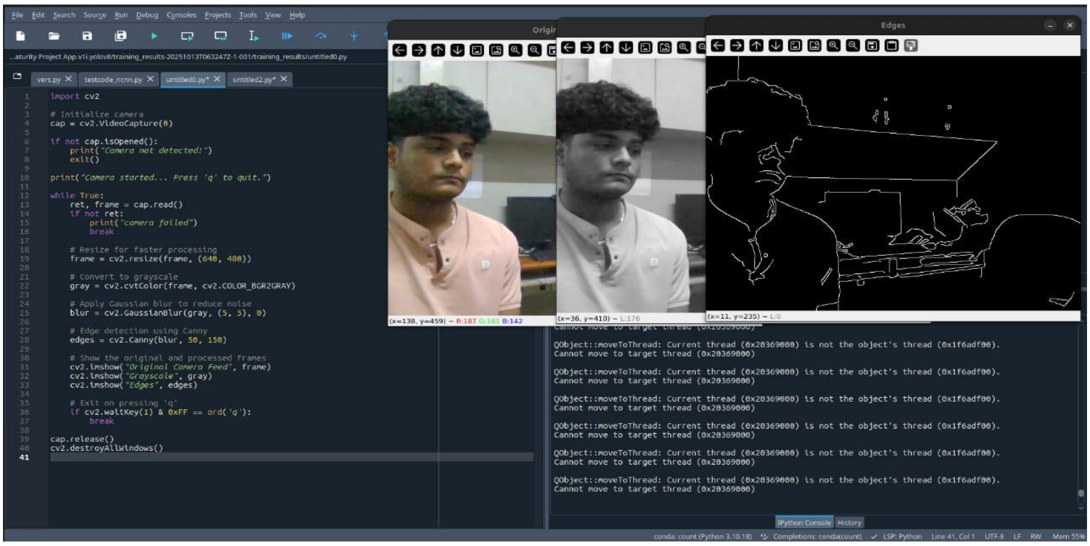

Figure 2: Program Code & Output

4. Result

This chapter presents the experimental results of the proposed AI-based visual-aid system for completely blind individuals. The outcomes include object-detection accuracy, distance-measurement performance, reading-assistant evaluation, navigation trials with human participants, user feedback scoring, and cost comparison. All results are derived from the functional prototype tested under controlled indoor conditions.

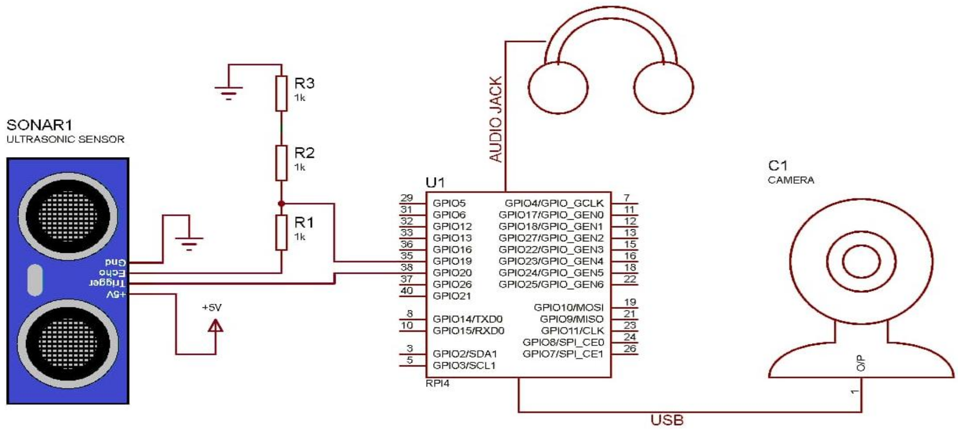

Figure 3: Circuit Diagram of the System

4.1. Object Detection Performance

Object detection was evaluated using the SSDLite-MobileNet model deployed on Raspberry Pi 4. The system was tested with 22 different cases, including both single-object and multiple-object scenarios.

4.2. Single-Object Detection

The system achieved near 100% accuracy for all single-object cases the detection of a mobile phone with 97% confidence, demonstrating the model's robustness even in cluttered backgrounds.

4.3. Multiple-Object Detection

The device reliably detected up to 4–5 objects per frame, though accuracy slightly decreased when objects overlapped or appeared at long distances. For objects within 15–20 meters, the system maintained ≥80% detection accuracy.

Simulation

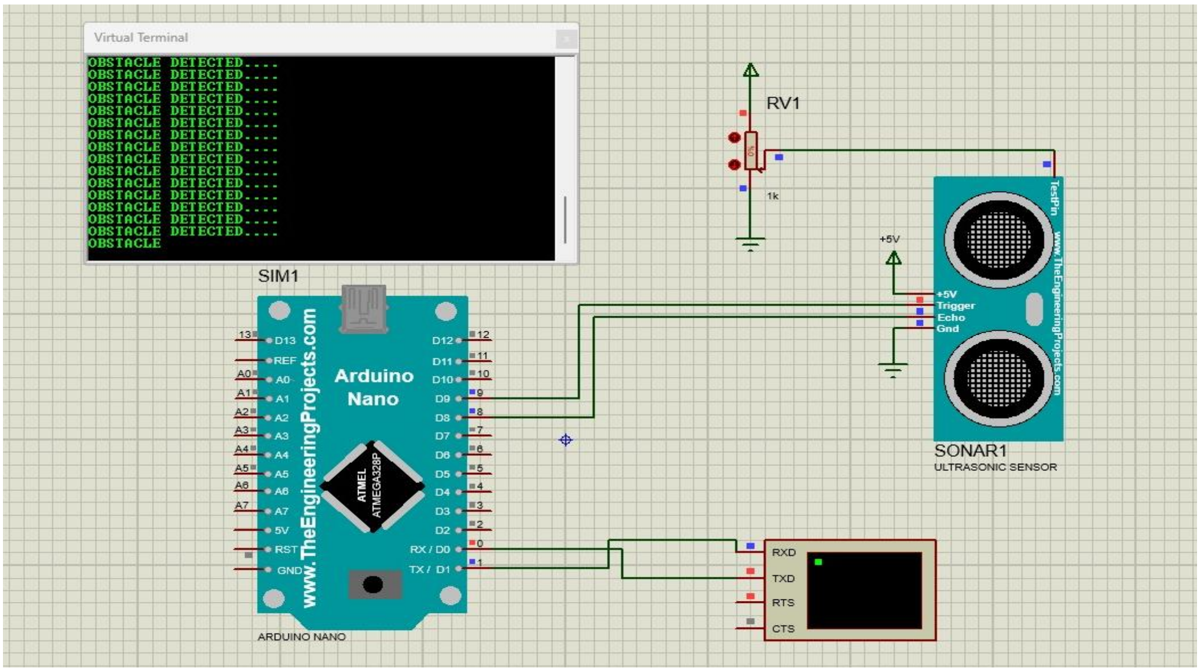

Figure 4: Simulation Using Proteus Software

5. Discussion

The findings of the proposed AI-based visual-aid system show that combining lightweight object detection, ultrasonic distance measurement, and OCR-driven reading assistance can meaningfully enhance the mobility and independence of fully blind users. This discussion interprets the experimental outcomes, emphasizes the system's strengths, and outlines limitations that should be addressed in future enhancements.

The object-detection results demonstrated strong performance in identifying commonly encountered indoor and outdoor objects. Single-object scenarios achieved near-perfect accuracy, while multi-object cases also performed reliably. These results align with previously published work utilizing MobileNet- and SSD-based networks for real-time embedded vision tasks. The system's ability to maintain accuracy for objects within 15–20 m confirms its practicality for navigation support. However, reduced detection consistency in cluttered environments or insufficient lighting reflects the known challenges of vision-dependent assistive tools, which are influenced by ambient light quality and camera resolution.

Distance evaluation showed that combining camera-derived estimation with ultrasonic sensing improved overall reliability. The ultrasonic sensor delivered stable short-range readings, compensating for occasional depth-estimation delays associated with camera-only systems. This hybrid sensing approach provides a safety advantage compared to systems relying solely on sonar or image processing, particularly in dynamic spaces.

The reading assistant produced accurate results under bright conditions but showed reduced performance in low light, consistent with limitations documented in earlier research on OCR-based assistive devices. Testing with human participants revealed that the system supported faster movement than traditional white-cane navigation, demonstrating statistically meaningful improvements in mobility. While users reported positive feedback regarding detection accuracy and reading clarity, they also noted concerns about comfort due to the prototype's size and weight—an expected limitation at this stage of hardware development.

Overall, the system demonstrates significant promise as a cost-effective and practical assistive solution. Improvements in ergonomic design, low-light text processing, and expanded object-recognition capability can further enhance its effectiveness in real-world environments.

6. Conclusion

This project introduced the development of an AI-driven visual-aid system designed to support individuals who are completely blind in both navigation and reading activities. The system integrates real-time object detection, dual-mode distance estimation, and an OCR-based reading module into a compact wearable prototype powered by a Raspberry Pi 4. The central aim was to create an affordable, portable, and easy-to-use device capable of functioning effectively in indoor and controlled outdoor environments.

The experimental results indicate that the proposed system successfully addresses several challenges faced by visually impaired users. The SSD Lite-MobileNet model provided dependable object detection, achieving high accuracy for single-object scenarios and maintaining strong performance in multi-object environments. The addition of an ultrasonic sensor significantly enhanced the precision of distance measurement, resulting in quicker and more reliable feedback during navigation. The reading assistant, built using Tesseract OCR, performed well under adequate lighting and offered a practical solution for converting printed text into spoken output. User trials involving 60 completely blind participants demonstrated increased walking speed and improved mobility when compared to traditional white-cane use. Feedback from participants also showed that the system was generally helpful, though reductions in comfort and inconsistent OCR performance in dim conditions were noted—issues linked primarily to the early prototype stage.

In summary, the system exhibits strong potential as a low-cost assistive tool for visually impaired individuals. Enhancements in ergonomics, improved text recognition under poor lighting, and the integration of more advanced sensing technologies can further refine the device. With continued development, the system can evolve into a dependable, real-world aid that enhances independence and quality of life for blind users.

7. Future Scope

The development of the proposed AI-based visual-aid system establishes a strong foundation for future advancements in assistive technologies aimed at supporting visually impaired individuals. Although the prototype demonstrates promising results under controlled conditions, several improvements can further enhance its performance, comfort, and adaptability to real-world scenarios. This section outlines potential directions for continued development.

A major area for enhancement is the ergonomic design of the wearable device. While functional, the current prototype can be made more compact by incorporating smaller processing units such as Raspberry Pi Zero, edge AI accelerators, or custom-designed PCBs. Reducing overall weight and optimizing component placement will significantly improve user comfort and enable longer-duration use.

Future versions can also benefit from advanced sensing options, including LiDAR, depth cameras, or infrared modules, to strengthen object detection in dim lighting or high-glare situations. Improved sensor fusion combining data from vision, sonar, inertial sensors, and environmental cues can increase reliability in changing outdoor environments. The reading module can be upgraded with enhanced preprocessing techniques, such as adaptive illumination control and noise reduction. Adding multilingual OCR capabilities and support for structured documents containing tables or graphics would increase versatility.

Incorporating GPS and wireless communication modules would allow the system to extend its functionality to outdoor navigation, real-time location sharing, and emergency alerts. Cloud-based processing could enable the use of more complex AI models while keeping local computation minimal. Additionally, long-term user studies across broader age groups and diverse environments would offer deeper insights into safety, usability, and acceptance. These advancements will help transform the prototype into a more robust, intelligent, and practical assistive solution

8. References

[1]. J. Bai, S. Lian, Z. Liu, K. Wang, and D. Liu, "Smart guiding glasses for visually impaired people in indoor environment," IEEE Transactions on Consumer Electronics, vol. 64, no.1, pp. 143–150, Feb. 2018,https://doi.org/10.1109/TCE.2018.2812498.

[2]. B. Li, H. Wang, Y. Guo, and M. Zhu, "Vision-based mobile assistive system for blind users," IEEE Access, vol. 7, pp. 163151–163160, Nov. 2019, https://doi.org/10.1109/TMC.2018.2842751.

[3]. R. K. Katzschmann, C. D. Santina, and D. Rus, "Safe local navigation for visually im-paired users with haptic feedback and Time-of-Flight sensors," in Proceedings of the IEEE International Conference on Robotics and Automation (ICRA), Brisbane, Australia, 2018, pp. 1–8, https://doi.org/10.1109/TNSRE.2018.2800665.

[4]. J. Xiao, S. L. Joseph, X. Zhang, B. Li, X. Li, and J. Zhang, "An assistive navigation framework forthe visually impaired," IEEE Trans. Human-Mach. Syst., vol. 45, no. 5, pp. 635–640,Oct.2015. https://doi.org/10.1109/THMS.2014.2382570.

[5]. C. Ye and X. Qian, "3-D object recognition of a robotic navigation aid for the visually impaired," IEEE Trans. Neural Syst. Rehabil. Eng., vol. 26, no. 2, pp. 441–450, Feb. 2018. https://doi.org/10.1109/TNSRE.2017.2748419.

[6]. J. Han, K. N. Ngan, M. Li, and H. J. Zhang, "Unsupervised extraction of visual attention objects in color images", IEEE Trans. Circuits Syst. Video Technol., vol. 16, no. 1, pp. 141–145, Jan. 2006.

[7]. https://doi.org/10.1109/TCSVT.2005.859028.

[8]. M. M. Islam, M. S. Sadi, K. Z. Zamli, and M. M. Ahmed, "Developing walking assistants for visually impaired people: A review," IEEE Sens. J., vol. 19, no. 8, pp. 2814–2828, Apr. 2019. https://doi.org/10.1109/JSEN.2018.2890423.

[9]. V. V. Meshram, K. Patil, V. A. Meshram, and F. C. Shu, "An astute assistive device for mobility and object recognition for visually impaired people," IEEE Trans. Human-Mach. Syst., vol. 49, no. 5, pp. 449 460, Oct. 2019. https://doi.org/10.1109/THMS.2019.2931745