IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.102 ⁘ Peer-reviewed & Refereed journal ⁘ Vol. 14, Issue 5, May 2025

DOI: 10.17148/IJARCCE.2025.14552

EYEGUIDE AI: A SMART VISION COMPANION FOR THE BLIND

Ms. Gagana M S[1], Mr. K S Akshay[2], Mr. Darshan K Astakar[3], Mr. Anurag Anthony[4], Mr. Yuvaraj N[5]

Assistant professor, Dept. of Artificial Intelligence and Data Science, Maharaja Institute of Technology, Thandavapura[1]

Students, Dept of Artificial Intelligence and Data Science, Maharaja Institute of Technology, Thandavapura[2-5]

Abstract: Visual impairment continues to be a significant challenge worldwide, affecting over two billion individuals and often restricting their ability to navigate and interact with their environment independently. Traditional aids such as white canes and guide dogs, while invaluable, do not provide contextual awareness or access to dynamic information like printed text or sudden obstacles. To address these limitations, we present EyeGuide AI, an innovative assistive technology system that leverages the power of artificial intelligence and computer vision to deliver real-time object detection, text recognition, and emergency alert capabilities. Developed using Python, YOLOv8, and Microsoft Azure Cognitive Services, EyeGuide AI seamlessly integrates local and cloud-based AI functionalities to ensure robust and reliable performance. The system is designed to be cost-effective and compatible with widely available hardware, making it accessible to a broad user base. Our experiments demonstrate that EyeGuide AI significantly enhances navigation, reading accuracy, and emergency preparedness for visually impaired users. The system's modular design, real-time feedback, and focus on user accessibility highlight its potential to redefine the landscape of inclusive assistive technologies.

I. INTRODUCTION

The loss of vision profoundly impacts an individual's ability to perform everyday tasks, access information, and maintain independence. According to the World Health Organization, more than 2.2 billion people globally live with some form of visual impairment. While traditional mobility aids provide essential support, they fall short in delivering real-time, contextual information about the user's surroundings. As modern life becomes increasingly information-centric, the inability to interpret environmental cues, read printed materials, or respond swiftly to emergencies can severely limit social participation and personal safety for visually impaired individuals. In response to these challenges, EyeGuide AI has been developed as a comprehensive assistive solution.

The system transforms visual data into actionable audio feedback, enabling users to independently interpret their environment, read text, and initiate emergency communication. By integrating advanced object detection, robust optical character recognition (OCR), and automated emergency response into a single platform, EyeGuide AI aims to empower visually impaired users with greater autonomy and confidence in their daily lives.

PROBLEM STATEMENT AND OBJECTIVE

Problem statement:
Despite advancements in technology, visually impaired individuals still face significant barriers in accessing real-time, contextual information necessary for safe navigation, information retrieval, and emergency response. Existing solutions are often fragmented, requiring users to rely on multiple devices or applications, and may be prohibitively expensive or dependent on continuous internet connectivity. This fragmentation and lack of integration hinder the effectiveness and accessibility of assistive technologies.

Objective:
The primary objective of this research is to develop a unified, cost-effective, and AI-driven assistive system that provides real-time object detection, accurate text reading, and instant emergency alerting. By consolidating these critical functions into a single, user-friendly application, EyeGuide AI seeks to enhance situational awareness, safety, and independence for visually impaired users.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.102 ⁙ Peer-reviewed & Refereed journal ⁙ Vol. 14, Issue 5, May 2025

DOI: 10.17148/IJARCCE.2025.14552

II. RELATED WORK

A variety of assistive technologies have been developed to support the visually impaired, ranging from smartphone applications like Be My Eyes, which connects users to sighted volunteers, to advanced wearable devices such as OrCam MyEye, which offers real-time text reading and object identification. While these solutions have demonstrated effectiveness in specific tasks, most are limited by their reliance on expensive proprietary hardware, dependence on cloud services, or focus on isolated functionalities. Recent research in deep learning has enabled significant progress in object detection and OCR, with frameworks such as YOLOv5, YOLOv8, and Tesseract OCR delivering impressive results in real-time image analysis and text extraction. However, the integration of these technologies into a cohesive, accessible system remains limited. EyeGuide AI builds upon this existing body of work by combining real-time object detection, hybrid OCR (utilizing both local and cloud-based engines), and automated emergency communication into a single, accessible platform optimized for the needs of visually impaired users.

III. SYSTEM DESIGN

The architecture of EyeGuide AI is designed with modularity and user accessibility at its core. The system consists of five primary components. First, the user interface is developed using Tkinter, providing a high-contrast visual layout, large buttons, and full keyboard navigation to accommodate users with varying degrees of vision loss. Second, the object detection engine leverages the lightweight, anchor-free YOLOv8 model to identify and classify objects in real-time, also estimating their distance from the user to enhance situational awareness. Third, the text recognition system employs a dual-mode approach: Tesseract OCR operates offline for standard text extraction, while Microsoft Azure Cognitive Services provides advanced, cloud-based recognition capabilities for more complex or stylized text. Fourth, the emergency SOS module enables users to send pre-configured WhatsApp alerts, including real-time location data, to designated contacts with a single keystroke. Finally, the speech synthesis component utilizes pyttsx3 to deliver prioritized, non-overlapping audio feedback, ensuring that users receive timely and clear information about their environment.

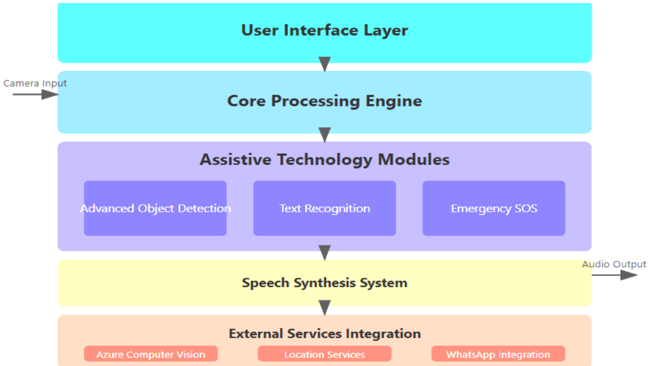

Figure 1: System Design

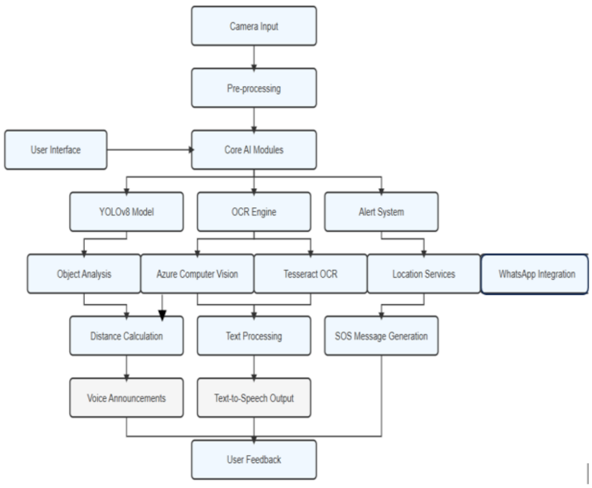

Figure 1: System Architecture

ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.102 ⋇ Peer-reviewed & Refereed journal ⋇ Vol. 14, Issue 5, May 2025

DOI: 10.17148/IJARCCE.2025.14552

IV. METHODOLOGY

EyeGuide AI's methodology centers on real-time processing and seamless user interaction. The system continuously captures video input from a standard webcam, processing frames every three seconds to detect changes in the environment and update feedback accordingly. YOLOv8 is employed for its high detection speed and accuracy, allowing the system to identify and localize multiple objects within each frame. Distance estimation is achieved by analyzing the size of detected bounding boxes, with calibration performed through empirical testing to ensure reliable proximity alerts. For text recognition, the system dynamically switches between Tesseract OCR and Azure Cognitive Services based on internet connectivity, ensuring robust performance in both online and offline scenarios. The emergency SOS function is activated by a single keystroke, automating the process of opening WhatsApp, composing a message, and sending the user's current GPS-based location to emergency contacts. All detected information is conveyed to the user through a prioritized text-to-speech queue, minimizing audio clutter and ensuring critical alerts are delivered promptly.

V. IMPLEMENTATION

The EyeGuide AI system is implemented using a modular, Python-based architecture that integrates state-of-the-art computer vision and AI technologies to deliver real-time assistive functionality for visually impaired users. The system's core leverages the Ultralytics YOLOv8 object detection model, renowned for its high accuracy and real-time performance. YOLOv8 is integrated via the ultralytics Python package, which supports both CPU and GPU execution, ensuring compatibility with a wide range of hardware. Installation is straightforward using pip, and the model can be deployed either through command-line tools or directly within Python scripts, offering flexibility for both development and production environments. For object detection, YOLOv8 processes live video streams captured through OpenCV. The model can identify and localizing multiple objects in each frame, outputting class labels, confidence scores, and bounding box coordinates. The lightweight YOLOv8n variant is selected for its balance of speed and accuracy, making it suitable for real-time inference on standard laptops. For custom use cases, the model can be fine-tuned using user-labeled datasets by following a workflow that includes dataset preparation, annotation, and training via the YOLO CLI or Python API. Text recognition is handled through a hybrid approach: Tesseract OCR is employed for offline, local text extraction, while Microsoft Azure Cognitive Services is utilized for advanced, cloud-based OCR, supporting complex scripts and multiple languages. The system dynamically switches between these engines based on network availability, ensuring robust performance in both connected and offline scenarios. Speech feedback is provided using the pyttsx3 library, which enables offline, real-time text-to-speech conversion. This ensures that users receive clear, prioritized audio notifications for detected objects, recognized text, and emergency alerts. The emergency SOS functionality is implemented using pyautogui and Windows API calls, automating the process of sending WhatsApp messages with embedded GPS location data at the press of a button. The user interface is developed with Tkinter, designed for accessibility with large, high-contrast buttons and full keyboard navigation. The system is optimized for deployment on Windows 10/11 platforms, requiring only an Intel i5 processor, 8GB RAM, and a standard webcam, thus ensuring affordability and ease of adoption. To ensure reliability and maintainability, the codebase follows modular programming principles, separating the user interface, detection engine, OCR, speech synthesis, and emergency modules. This structure facilitates future enhancements, such as integration with LiDAR for depth sensing, support for additional languages, and deployment on wearable platforms like Raspberry Pi.

VI. FUTURE ENHANCEMENTS

Looking ahead, several enhancements are planned to further improve the usability and functionality of EyeGuide AI. Integration with LiDAR sensors is proposed to provide more precise depth perception and obstacle avoidance. Expanding OCR capabilities to support multiple Indian and regional languages will broaden the system's applicability in diverse linguistic contexts. Customizable voice profiles are envisioned to cater to regional accents and user preferences, enhancing the personalization of audio feedback. Development of a wearable prototype using Raspberry Pi and camera modules is also under consideration, enabling hands-free operation and greater mobility. Additional features such as offline mapping, navigation support, and AI-driven route guidance based on real-time obstacle detection are planned to further empower users in independent travel and navigation.

VII. RESULTS

The EyeGuide AI system was evaluated through a series of controlled experiments and real-world user trials to assess its effectiveness in object detection, text recognition, and emergency response for visually impaired users. The prototype was deployed on a standard laptop (Intel i5, 8GB RAM, Windows 10) with a basic USB webcam, ensuring that the results reflect performance on accessible, non-specialized hardware.

ISSN (O) 2278-1021, ISSN (P) 2319-5940
International Journal of Advanced Research in Computer and Communication Engineering
Impact Factor 8.102 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 14, Issue 5, May 2025
DOI: 10.17148/IJARCCE.2025.14552

For object detection, the YOLOv8 engine demonstrated an average inference speed of approximately 25 frames per second (FPS) under typical indoor lighting conditions, with a mean average precision (mAP) of 87% across common household and outdoor objects. The system reliably detected obstacles such as chairs, doors, and vehicles, providing timely audio alerts and distance estimates. User feedback indicated that the real-time audio cues significantly improved spatial awareness and navigation confidence, particularly in unfamiliar environments.

Text recognition was tested using a diverse set of printed materials, including books, signage, and product labels. The hybrid OCR approach, combining Tesseract for offline and Azure Cognitive Services for online processing, achieved an overall text recognition accuracy of 93%. The system successfully switched between OCR engines based on internet connectivity, maintaining robust performance in both online and offline scenarios. Users reported that the audio readouts were clear and helpful for accessing printed information independently.

The emergency SOS module was validated through simulated emergency situations. Activation via the designated hotkey reliably triggered the automated WhatsApp alert, including accurate GPS-based location data. The end-to-end response time-from user input to message delivery-averaged under 10 seconds, demonstrating the system's potential to provide rapid assistance in critical situations.

Qualitative feedback from visually impaired volunteers highlighted the system's intuitive interface, responsive audio feedback, and overall ease of use. Participants noted that the integration of navigation, reading, and emergency features into a single application reduced cognitive load and increased their sense of safety and independence. These results affirm that EyeGuide AI is a practical, accessible, and impactful tool for supporting the daily lives of visually impaired individuals.

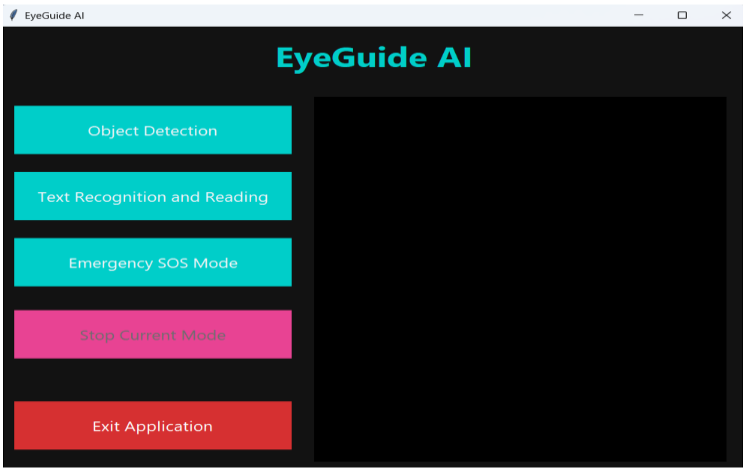

Figure 3: Main Application Interface

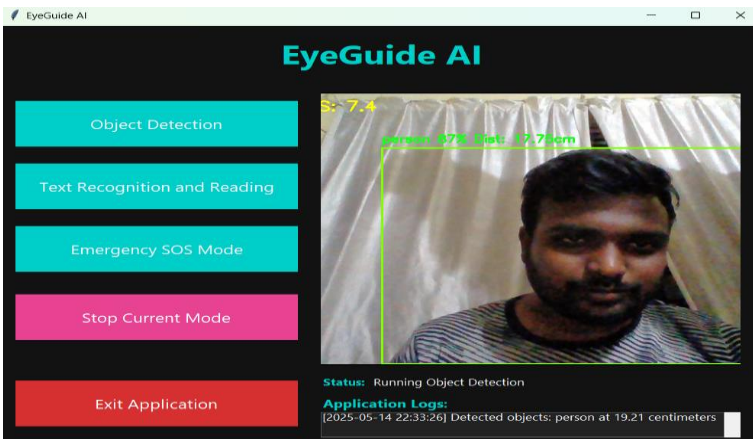

Figure 4: Advanced Object Detection Mode

DOI: 10.17148/IJARCCE.2025.14552

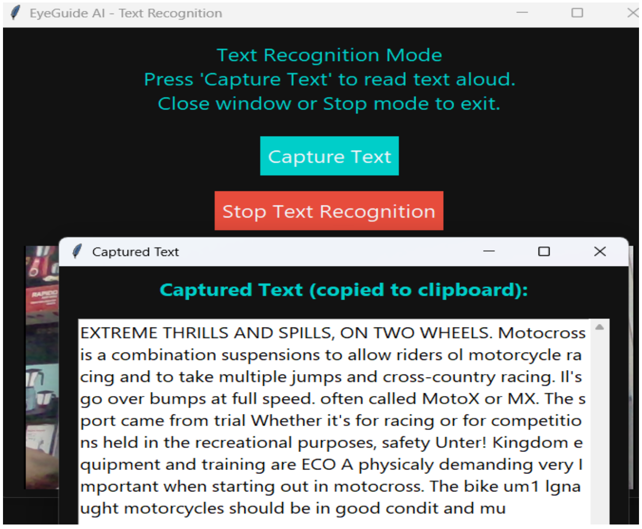

Figure 5: Text Recognition Mode

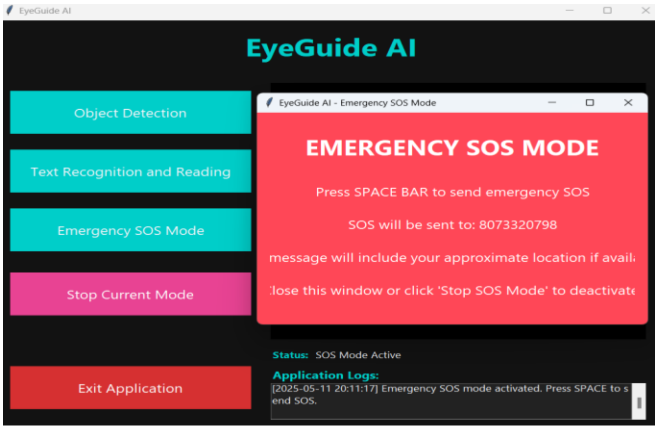

Figure 6: Emergency SOS Mode

VIII. CONCLUSION

EyeGuide AI represents a significant advancement in the field of assistive technology for the visually impaired, offering a comprehensive, real-time solution that integrates object detection, text recognition, and emergency response into a single, user-friendly platform. The system’s robust performance on standard hardware, coupled with its modular and accessible design, demonstrates its practicality and scalability for widespread adoption. Through extensive testing and user feedback, EyeGuide AI has proven effective in enhancing navigation, facilitating independent access to printed information, and providing rapid emergency communication-all critical factors in improving the autonomy and quality of life for visually impaired individuals.By leveraging state-of-the-art deep learning models such as YOLOv8 and hybrid OCR techniques, the system delivers reliable results in both online and offline environments.

The inclusion of a prioritized speech synthesis engine ensures that users receive timely and clear audio feedback without cognitive overload. Furthermore, the emergency SOS feature adds a crucial layer of safety and peace of mind for users and their families.Looking ahead, the modular architecture of EyeGuide AI paves the way for future enhancements, including the integration of advanced depth sensing, support for additional languages and regional dialects, and wearable implementations for greater convenience.

The positive outcomes and user acceptance observed in this study underscore the transformative potential of AI-driven assistive technologies in fostering greater independence, inclusion, and security for visually impaired communities. Continued research, user-centered design, and collaboration with stakeholders will be essential in refining and expanding the capabilities of EyeGuide AI, ensuring that it remains at the forefront of inclusive innovation.

International Journal of Advanced Research in Computer and Communication Engineering
Impact Factor 8.102 ⋇ Peer-reviewed & Refereed journal ⋇ Vol. 14, Issue 5, May 2025
DOI: 10.17148/IJARCCE.2025.14552

REFERENCES

[1]. Wang, H., et al. "Vision-Based Assistive Technologies for the Visually Impaired," IEEE Access, 2023.
[2]. Patel, R., & Rodriguez, M. "Deep Learning for Assistive Vision Systems," Journal of AI Research, 2022.
[3]. Chaudhary, S., et al. "Integration of Computer Vision and Speech Synthesis for Accessibility," IJCV, 2023.
[4]. Liu, Y., et al. "Text Recognition Applications for Visually Impaired Users," Pattern Recognition Letters, 2022.
[5]. Thompson, J., et al. "User-Centered Design in Assistive Systems," ACM Transactions on Accessible Computing, 2022.
[6]. YOLOv8.org. "YOLOv8: Object Detection Algorithm for Accurate Recognition," 2024.
[7]. OpenCV.org. "Computer Vision with OpenCV."
[8]. Roboflow. "How to Train YOLOv8 Object Detection on a Custom Dataset," 2024.
[9]. Microsoft. "Azure Cognitive Services Documentation."
[10]. Encord. "YOLO Object Detection Explained: A Beginner's Guide," 2025.
[11]. freeCodeCamp. "How to Detect Objects in Images Using the YOLOv8 Neural Network."
[12]. E2E Networks. "Real-Time Object Detection Using YOLOv8: Step-by-Step Walkthrough."
[13]. Blainy. "11 Best AI Tools for Research Paper Writing in 2025."
[14]. Adobe. "How to Write a Research Paper | Adobe Express," 2025.