International Journal of Current Science Research
www.drbgrpublications.in

Lightweight Real-Time Object Detection and Navigation System for Visually Impaired Using Deep Learning

Samuel Simutwe[1]* and Sesu Ovid[2]

[1]*UG Student, Department of Computer Science and Technology, DMI-St. Eugene University, Zambia*

[2]*Lecturer, Department of Computer Science and Technology, DMI-St. Eugene University, Zambia*

**Corresponding author email:* ssimutwe0@gmail.com

Abstract

Visual impairment continues to affect millions of individuals worldwide, limiting their independence and access to essential services. Existing assistive technologies often depend on expensive hardware, continuous internet connectivity, or fragmented functionalities that fail to address the diverse needs of visually impaired users. This paper presents a lightweight real-time object detection and navigation system designed to assist blind and visually impaired individuals using deep learning technologies. The proposed Android-based application integrates YOLOv8 object detection, Google Gemini-powered scene understanding, optical character recognition, GPS-based emergency alert services, and a voice-driven user interface. TensorFlow Lite is employed to enable efficient on-device inference for real-time obstacle detection while minimizing computational requirements. The system provides audio-based feedback, allowing users to navigate independently and safely. Experimental evaluations demonstrate high object detection accuracy, low response latency, and improved user experience. The proposed solution offers an affordable and scalable assistive technology suitable for deployment in developing countries where accessibility resources remain limited.

Keywords: Visual Impairment, Deep Learning, YOLOv8, TensorFlow Lite, Assistive Technology, Artificial Intelligence, Mobile Application, Accessibility, Navigation System

Introduction

Visual impairment is one of the leading causes of reduced mobility and social exclusion worldwide. According to the World Health Organization, approximately 2.2 billion people live with some form of vision impairment, with millions experiencing severe blindness. In developing countries such as Zambia, visually impaired individuals face additional challenges due to limited access to assistive technologies, specialized healthcare services, and accessibility infrastructure.

Traditional navigation aids such as white canes and guide dogs provide basic assistance but cannot identify objects, interpret environmental contexts, read textual information, or communicate emergency situations. Recent advancements in artificial intelligence, computer vision, and mobile computing have created new opportunities for intelligent assistive systems capable of providing real-time environmental awareness and navigation support.

Deep learning-based object detection models such as YOLOv8 have achieved remarkable improvements in both detection accuracy and processing speed. Similarly, multimodal artificial intelligence systems such as Google Gemini enable contextual scene understanding and natural language interaction. Integrating these technologies into a mobile platform can significantly enhance the independence and safety of visually impaired individuals.

This study proposes a lightweight AI-powered navigation and object detection system that combines real-time object recognition, scene description, text reading, and GPS-based emergency alerting into a unified Android application. The system is specifically designed for deployment on affordable smartphones commonly available in developing countries.

Literature Review

The application of artificial intelligence in assistive technology has received considerable research attention during the last decade. Early systems relied primarily on image processing techniques and ultrasonic sensors for obstacle detection. Although these approaches provided basic functionality, they often suffered from low accuracy and limited adaptability in dynamic environments.

The emergence of deep learning significantly improved object detection performance through architectures such as Faster R-CNN, SSD, and YOLO. Among these, YOLOv8 has demonstrated superior performance in balancing speed and accuracy for mobile applications. Studies have shown that YOLOv8 can achieve real-time object detection while maintaining high precision across multiple object categories.

Several commercial applications have been developed to support visually impaired individuals. Microsoft Seeing AI provides text reading, object recognition, and scene description services. Google Lookout offers environmental narration and document reading capabilities. Be My Eyes connects users with sighted volunteers for remote assistance, while Envision AI combines text recognition and scene interpretation functionalities.

Despite these advancements, existing solutions exhibit several limitations. Most require continuous internet connectivity, lack integrated emergency alert mechanisms, and are not optimized for deployment in low-resource environments. Furthermore, many applications provide only partial assistance by focusing on a single functionality rather than offering a comprehensive support platform.

The literature reveals a significant gap in developing integrated systems capable of combining real-time object detection, scene understanding, text recognition, and emergency communication within a single mobile application. This research seeks to address this gap by designing a unified AI-driven assistive platform.

Methodology

The development of the proposed system followed the Agile Software Development methodology, enabling iterative design, implementation, testing, and refinement throughout the project lifecycle.

The first phase involved requirement gathering through consultations with visually impaired individuals and accessibility specialists. User needs were analyzed to identify critical functional requirements related to navigation, environmental awareness, emergency assistance, and accessibility.

The second phase focused on system design. A layered architecture was developed comprising a presentation layer, application layer, artificial intelligence processing layer, data management layer, and external service layer. This modular structure improves maintainability, scalability, and integration efficiency.

The third phase involved implementation using Android Studio. The object detection module was developed using YOLOv8 and deployed through TensorFlow Lite for efficient on-device inference. The Google Gemini API was integrated to provide contextual scene descriptions and text interpretation capabilities. Speech recognition and text-to-speech technologies were incorporated to enable voice-based interaction.

The final phase consisted of testing and evaluation. System performance was assessed using metrics including object detection accuracy, response latency, emergency alert reliability, and user satisfaction. Practical testing was conducted under different environmental conditions to evaluate real-world usability.

System Architecture

The proposed system consists of five major components:

A. Voice Interface Module

This module serves as the primary interaction mechanism between the user and the application. Voice commands are captured using Android Speech Recognition services and translated into executable actions.

B. Object Detection Module

The object detection component employs the YOLOv8 model optimized through TensorFlow Lite. The module continuously processes camera frames and identifies surrounding objects and obstacles in real time.

C. Scene Description Module

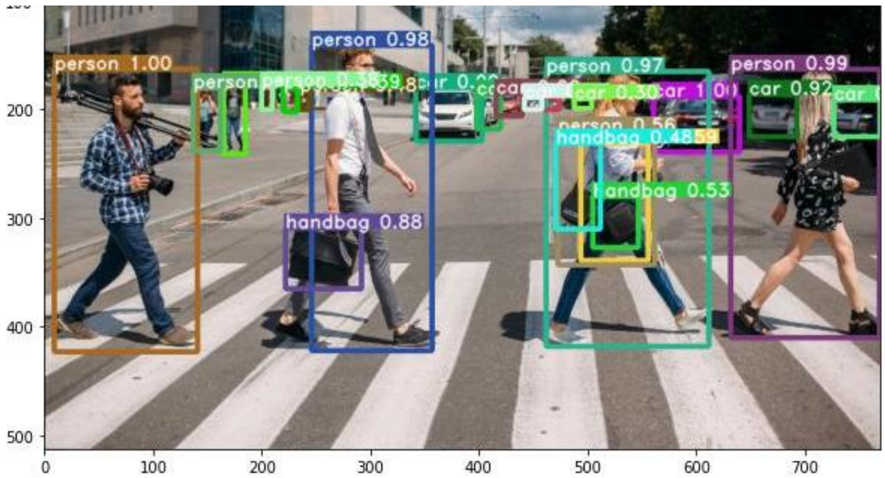
The scene description module utilizes the Google Gemini API to analyze captured images and generate meaningful descriptions of the surrounding environment. The generated information is delivered through synthesized speech.

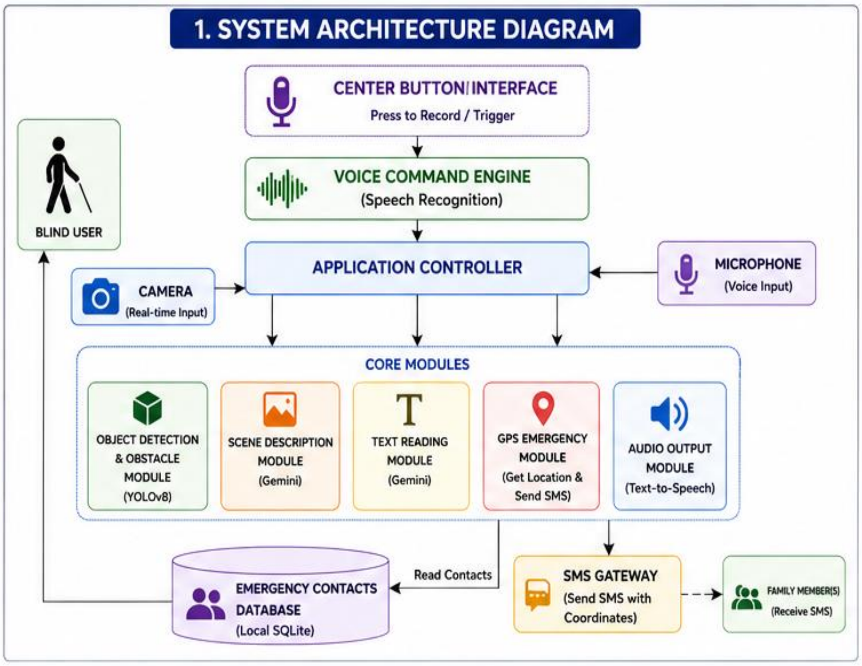
D. Text Reading Module

The text reading component identifies and extracts textual information from signs, documents, labels, and product packaging. Extracted text is converted into speech to facilitate accessibility.

E. Emergency Alert Module

The emergency alert subsystem retrieves the user's GPS coordinates and sends location information through SMS to predefined emergency contacts. This feature enhances personal safety during emergencies.





Results and Discussion

The developed system was evaluated on Android smartphones equipped with mid-range processors. Experimental results indicate that the YOLOv8 model successfully performs real-time object detection with high accuracy while maintaining low computational overhead.

The object detection module accurately recognized common environmental objects including people, vehicles, doors, chairs, tables, and obstacles. TensorFlow Lite optimization significantly reduced inference latency, enabling smooth operation on resource-constrained devices.

The Google Gemini integration demonstrated effective scene interpretation capabilities by generating concise and contextually relevant descriptions. Users were able to receive meaningful environmental information through voice feedback, enhancing situational awareness.

The text recognition module successfully extracted textual content from printed materials and provided accurate speech output. Emergency alerts were transmitted reliably, and GPS coordinates were correctly delivered to registered contacts.

User feedback highlighted improved confidence, independence, and ease of navigation. The voice-driven interface was particularly appreciated due to its accessibility and ease of use. Overall, the integration of multiple assistive functionalities into a single application significantly improved usability compared to existing standalone solutions.

Conclusion

This research presented the design and implementation of a lightweight real-time object detection and navigation system for visually impaired individuals using deep learning technologies. The proposed solution integrates YOLOv8 object detection, Google Gemini-powered scene understanding, text recognition, GPS-based emergency alerting, and voice interaction within a single Android application.

The results demonstrate that the system achieves high detection accuracy, low response latency, and strong usability while operating on affordable mobile devices. The offline-first architecture ensures functionality even in low-connectivity environments, making the system particularly suitable for deployment in developing countries.

Future enhancements will focus on multilingual voice support, indoor navigation capabilities, public transportation integration, and advanced contextual assistance features. The proposed system contributes to improving accessibility, independence, and quality of life for visually impaired individuals through the practical application of artificial intelligence.

Reference

World Health Organization (WHO), *World Report on Vision*, Geneva, Switzerland, 2023.

J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You Only Look Once: Unified, Real-Time Object Detection," *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 779–788, 2016.

G. Jocher et al., "YOLOv8 Documentation and Technical Report," Ultralytics, 2024.

Google LLC, "TensorFlow Lite Documentation," 2024.

Google DeepMind, "Gemini API Developer Guide," 2025.

Microsoft Corporation, "Seeing AI: Artificial Intelligence for Accessibility," 2024.

Google LLC, "Lookout App for Android," 2024.

Danquah, Nkrumah, and Asare, "Assistive Technology Adoption in Sub-Saharan Africa," *Disability and Rehabilitation: Assistive Technology*, vol. 19, no. 2, pp. 115–126, 2024.

Mbunge, Sigwejo, and Fashoto, "Digital Health and Assistive Technologies in Africa," *Telematics and Informatics*, vol. 58, pp. 101–114, 2021.

Mwale and Zulu, "Accessibility Challenges among Visually Impaired Individuals in Zambia," *African Disability Rights Yearbook*, vol. 11, pp. 45–59, 2023.