

AI-Powered Real Time Navigation Aid For Visually Impaired Using Mobile Live Video

S Revathi¹, B. Samhitha², E. Sarika³, K. Usharani⁴

¹Assistant Professor; Bhoj Reddy Engineering College for Women Department of Information Technology
Hyderabad India.

^{2,3,4}B.Tech.Students ; Bhoj Reddy Engineering College for Women Department of Information Technology
Hyderabad India.

Mail Id; samhithareddy005@gmail.com , sarikareddy594@gmail.com , usharanikondrasi@gmail.com

Abstract

This project presents a real-time navigation aid that utilizes a mobile camera integrated with on-device Artificial Intelligence (AI), specifically leveraging YOLO and TensorFlow Lite models, to provide continuous guidance and environmental awareness. Visually impaired individuals often face significant challenges in detecting obstacles and recognizing objects in real time, particularly in unfamiliar or crowded environments. The proposed system captures live video streams, which are processed using neural network models to detect objects and estimate their positions. The detected information is transformed into simple navigational cues and delivered to the user through real-time audio feedback using Text-to-Speech (TTS) technology for immediate guidance.

Haptic feedback (vibration) is incorporated as a key feature, where the vibration intensity and duration are adjusted based on the object's proximity (area ratio), enabling users to perceive obstacle distance and direction effectively. The complete detection and feedback mechanism operates entirely offline, ensuring reliable functionality without requiring internet connectivity. However, for enhanced accessibility during startup, the application can be launched and closed using voice assistants such as Google Assistant, Gemini, and Bixby, which require an active internet connection. Once activated, all core functionalities operate fully offline.

The application is designed for continuous operation through an Android Foreground Service, allowing efficient performance even when the device screen is turned off. In addition, an auto-torch feature is integrated to support accurate obstacle detection under varying lighting conditions. The proposed system aims to provide an efficient, accessible, and dependable navigation solution for visually impaired users through real-time mobile AI-based assistance.

Keywords— Assistive Technology, Visual Impairment, Real-Time Object Detection, YOLO, TensorFlow Lite, Offline Navigation System, Audio Feedback, Haptic Feedback, Android Foreground Service, Mobile Artificial Intelligence.

Introduction

Independent mobility remains one of the major challenges faced by visually impaired individuals in

everyday life. Traditional mobility aids such as white canes and guide dogs assist users in navigating their surroundings; however, these solutions often provide limited awareness in dynamic, crowded, or unfamiliar environments where rapid obstacle detection is required. Recent advancements in Artificial Intelligence (AI), computer vision, and mobile computing technologies have created new opportunities for developing intelligent assistive systems capable of improving safety, autonomy, and environmental awareness.

This research proposes an AI-based real-time navigation assistance system that utilizes a smartphone camera integrated with lightweight on-device deep learning models, including YOLO and TensorFlow Lite, to support visually impaired users. The application continuously captures live visual data from the surrounding environment and performs object recognition, localization, and proximity estimation directly on the mobile device. Detected information is translated into user-friendly navigation guidance using a multimodal feedback mechanism.

The proposed system employs Text-to-Speech (TTS) technology to deliver spoken instructions regarding detected objects and their directional placement. In addition, a haptic feedback mechanism generates vibration patterns whose intensity varies according to obstacle proximity, enabling users to perceive spatial distance without depending exclusively on audio communication.

A significant feature of the system is its complete offline operational capability. All object detection, processing, and feedback modules execute locally on the smartphone, minimizing latency and removing dependence on network connectivity. For improved accessibility, users may launch or terminate the application using voice assistant platforms such as Google Assistant, Gemini, or Bixby, which require internet connectivity only during initialization. Once activated, the application continues functioning independently through an Android Foreground Service, ensuring uninterrupted performance even when the display is inactive.

To improve detection performance under challenging illumination conditions, an automatic torch control mechanism is incorporated. By combining offline AI inference, smartphone sensing capabilities, and multimodal feedback, the proposed

system aims to provide a portable, affordable, and intelligent mobility assistance solution that enhances independent navigation, situational awareness, and user safety across diverse real-world environments.

Existing System

Several mobility solutions currently support visually impaired individuals; however, each presents certain operational limitations. The white cane remains one of the most widely adopted traditional navigation tools. It enables users to detect nearby ground-level obstacles through tactile interaction, providing fundamental mobility assistance. Despite its effectiveness, the white cane offers limited environmental understanding, as it cannot reliably identify elevated obstacles, recognize object categories, or determine directional context.

Modern assistive technologies, including wearable devices such as smart glasses, sensor-enabled bands, and camera-based navigation systems, attempt to address these limitations by employing computer vision, sensors, and machine learning techniques. These systems often provide audio or haptic notifications regarding detected obstacles and surrounding objects. Although they improve environmental awareness, many wearable solutions involve high acquisition costs, increased hardware dependency, frequent charging requirements, and reduced long-term comfort for users.

These limitations indicate the need for a cost-effective, portable, and software-based assistive navigation solution capable of delivering real-time environmental understanding without requiring specialized external hardware.

Proposed System

The proposed system is an Android-based mobile application designed to assist visually impaired users through real-time object detection and navigation guidance using a smartphone camera. Unlike conventional mobility aids, the application integrates multimodal feedback mechanisms to improve environmental awareness and navigation efficiency.

The system employs deep learning-based object detection models running directly on the mobile device to analyze live video input and identify surrounding obstacles and objects. The detected information is communicated to the user through Text-to-Speech audio announcements that describe object identity and relative direction. Simultaneously, vibration-based haptic feedback conveys obstacle proximity, allowing users to interpret distance information intuitively.

By utilizing widely available smartphone hardware, including built-in cameras and sensors, the proposed application offers a practical, affordable, and portable alternative to expensive wearable assistive devices or guide animals. The system is intended to

provide comprehensive environmental awareness across multiple obstacle heights and distances, enabling visually impaired individuals to navigate complex environments with greater independence, confidence, and safety.

Related Work

User Survey and Requirement Study

To assess the practical demand for assistive navigation systems among visually impaired individuals, a structured survey was conducted involving visually impaired users, caregivers, rehabilitation professionals, and students specializing in assistive technologies. Participants were selected from educational institutions, disability support communities, and volunteer organizations involved in mobility assistance.

The survey findings revealed a substantial need for improved navigation support technologies. Approximately 72% of visually impaired respondents reported encountering considerable difficulty while navigating unfamiliar or densely populated environments, including public transportation terminals, commercial complexes, and educational campuses. Most participants indicated regular dependence on traditional mobility aids such as white canes, guide dogs, or assistance from nearby individuals. Although these methods provide basic mobility support, respondents highlighted limitations in environmental awareness and obstacle recognition.

The study further examined the use of digital accessibility tools. Around 65% of participants stated that they had experience using smartphone accessibility features such as screen readers, voice assistants, or GPS-based navigation applications. However, many users reported that existing solutions lacked real-time obstacle detection, environmental understanding, and immediate situational guidance.

Interest in AI-enabled navigation assistance was notably high. Nearly 70% of respondents expressed willingness to adopt a smartphone-based application capable of identifying obstacles and delivering instant voice guidance using the device camera. Frequently reported navigation challenges included low-level obstacle detection, stair recognition, crowded-area mobility, and safe movement through unfamiliar surroundings.

Participants also provided recommendations for desirable system features. The most frequently requested capabilities included real-time object detection, voice-based instructions, obstacle proximity alerts, offline functionality, vibration notifications, and interfaces compatible with screen-reader technologies. Additional suggestions included GPS navigation integration and emergency location-sharing mechanisms.

The survey outcomes demonstrate a clear requirement for affordable, portable, and intelligent

assistive technologies. These findings strongly support the development of the proposed AI-powered mobile navigation system that combines computer vision, deep learning, and multimodal feedback to improve independent mobility and environmental awareness for visually impaired users.

Requirement Analysis

Functional Requirements

The proposed navigation system is designed to provide continuous environmental assistance for visually impaired individuals through smartphone-based computer vision and AI processing. The major functional requirements are described below.

• Real-Time Object Detection

The application shall continuously capture live video input using the smartphone camera and perform real-time object detection using lightweight AI models such as YOLO and TensorFlow Lite.

• Direction and Distance Estimation

The system shall analyze detected objects to determine their approximate direction and relative proximity with respect to the user's position.

• Audio Guidance Mechanism

The application shall generate immediate spoken instructions using Text-to-Speech (TTS) technology to inform users about detected objects and navigational directions.

• Haptic Feedback Mechanism

The system shall provide vibration-based feedback, where vibration intensity and duration vary according to obstacle distance and environmental conditions.

• Offline Operation

All major functionalities, including detection, processing, and feedback generation, shall execute locally on the mobile device without requiring active internet connectivity.

• Voice Assistant Compatibility

The application shall support launching and termination through external voice assistant platforms to improve accessibility and hands-free operation.

Non-Functional Requirements

The proposed navigation application must satisfy several non-functional requirements to ensure usability, efficiency, and dependable performance in real-world scenarios.

Performance

The system must provide high-speed processing with minimal response latency. Object detection, distance estimation, and feedback delivery should occur in near real time to ensure safe navigation assistance during user movement. Efficient resource utilization is required to maintain stable performance on mobile devices.

Security

User privacy and application security must be maintained throughout system operation. Any stored preferences, accessibility settings, or user-related data should be protected using secure storage mechanisms and encrypted communication protocols where applicable.

Scalability

The system architecture should support future enhancements, including integration of additional AI models, navigation modules, cloud synchronization features, and expanded object recognition capabilities without requiring significant redesign.

Usability

The application interface must be simple, intuitive, and accessible to visually impaired users. Interaction mechanisms should minimize complexity and ensure compatibility with screen readers, voice commands, and accessibility guidelines.

Design

Architecture

System Architecture

The proposed AI-based navigation system is designed to provide real-time environmental awareness and mobility assistance for visually impaired individuals using a smartphone platform. The architecture integrates mobile sensing technologies, computer vision algorithms, and intelligent feedback mechanisms to enable safe and efficient navigation in different surroundings.

In the proposed framework, the user interacts with an Android mobile application installed on a smartphone device. The smartphone camera continuously acquires live video streams from the surrounding environment. Simultaneously, embedded device sensors such as GPS, accelerometer, and gyroscope may be utilized to obtain supplementary information related to user movement, orientation, and location.

The captured visual data are forwarded to a preprocessing stage, where operations including frame extraction, image normalization, noise reduction, and illumination adjustment are performed to improve input quality for AI inference. This preprocessing step enhances detection robustness under varying environmental and lighting conditions.

Following preprocessing, the refined image frames are transmitted to the AI processing layer. This module employs lightweight deep learning models, including YOLO and TensorFlow Lite-based object detection techniques, to recognize surrounding objects, estimate obstacle proximity, and interpret spatial relationships within the environment. The processing module performs continuous scene analysis to detect potential hazards and identify navigable regions.

The extracted environmental information is then delivered to the navigation decision module. Based

on detected object position, movement direction, and distance estimation, the system determines suitable navigation actions such as forward movement, directional adjustments, obstacle avoidance, or stop commands. Decision-making logic aims to provide context-aware guidance that supports safe user mobility.

The generated navigation decisions are communicated to the user through a multimodal feedback mechanism. Audio instructions are produced using a Text-to-Speech (TTS) engine and delivered through the smartphone speaker or

connected earphones. In addition to spoken guidance, vibration-based haptic feedback may be used to convey proximity information, enabling users to interpret environmental conditions through tactile cues.

To support future scalability and system maintenance, an optional cloud or database layer can be incorporated for AI model updates, performance monitoring, and configuration management. However, all primary navigation functionalities are designed to operate locally on the device, ensuring low latency and uninterrupted offline performance.

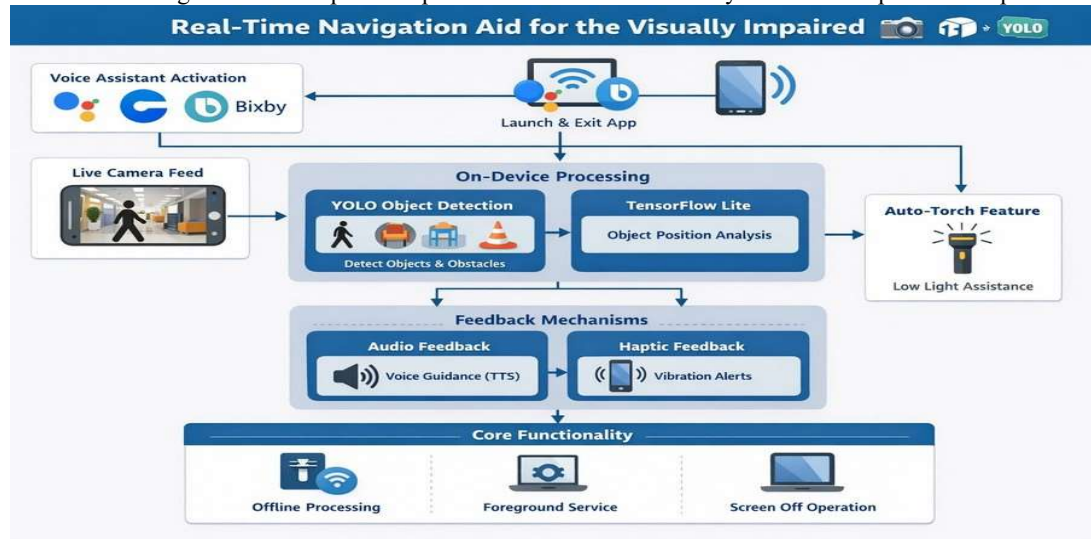


Fig.1; System Architecture

Technical Architecture

The technical architecture of the proposed AI-powered navigation system is designed to provide real-time environmental perception and navigation support for visually impaired individuals using a smartphone-based platform. The system combines mobile computing, computer vision, deep learning, and sensor integration to achieve efficient and reliable navigation assistance.

The smartphone camera functions as the primary sensing component by continuously acquiring live video streams from the surrounding environment. These captured video frames are supplied to a preprocessing layer responsible for preparing the visual input for efficient AI analysis. The preprocessing operations include image resizing, frame normalization, brightness adjustment, and noise reduction, which enhance image quality and improve model performance under varying environmental conditions.

After preprocessing, the refined image data are transmitted to the computer vision module. This module employs lightweight deep learning frameworks such as YOLO, MobileNet, or other convolutional neural network (CNN)-based object detection models to identify surrounding entities, obstacles, and important environmental features. The detection system recognizes objects including

pedestrians, vehicles, staircases, doors, pathways, and other navigation-relevant elements present in the user's surroundings.

Following object recognition, a distance estimation module evaluates the spatial relationship between detected objects and the user. Using parameters such as bounding-box dimensions, object position, and relative frame occupancy, the module estimates obstacle proximity and identifies potential navigation hazards.

The processed information is subsequently forwarded to the navigation decision engine. This component analyzes obstacle locations, estimated distances, and environmental context to determine appropriate movement recommendations. Depending on the detected scenario, the system generates navigation actions such as moving forward, turning left or right, slowing down, or stopping to avoid potential collisions.

The generated navigation outputs are converted into user-understandable guidance through a multimodal feedback mechanism. Spoken instructions are produced using a Text-to-Speech (TTS) engine and delivered via the smartphone speaker or connected earphones. Additionally, vibration-based haptic feedback may be incorporated to provide non-audio awareness of obstacle proximity and directional information.

To further improve navigation accuracy and contextual awareness, the architecture can integrate smartphone sensors such as GPS, accelerometer, and gyroscope. These components provide supplementary information regarding device orientation, user movement, and geographical positioning.

Although the proposed system primarily performs inference and decision-making directly on the mobile device to minimize response latency, an optional cloud or backend layer may be included for AI model updates, system maintenance, analytics, or dataset management. Lightweight mobile inference frameworks such as TensorFlow Lite or PyTorch Mobile enable efficient on-device execution while reducing computational overhead.

By integrating live video processing, deep learning-based object recognition, mobile sensor data, and intelligent feedback mechanisms, the proposed technical architecture delivers a portable, low-cost,

and real-time navigation solution that enhances environmental awareness and independent mobility for visually impaired users.

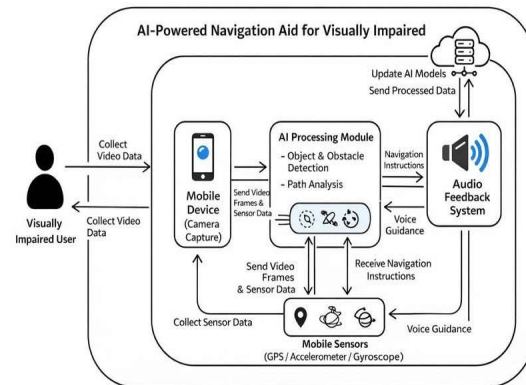
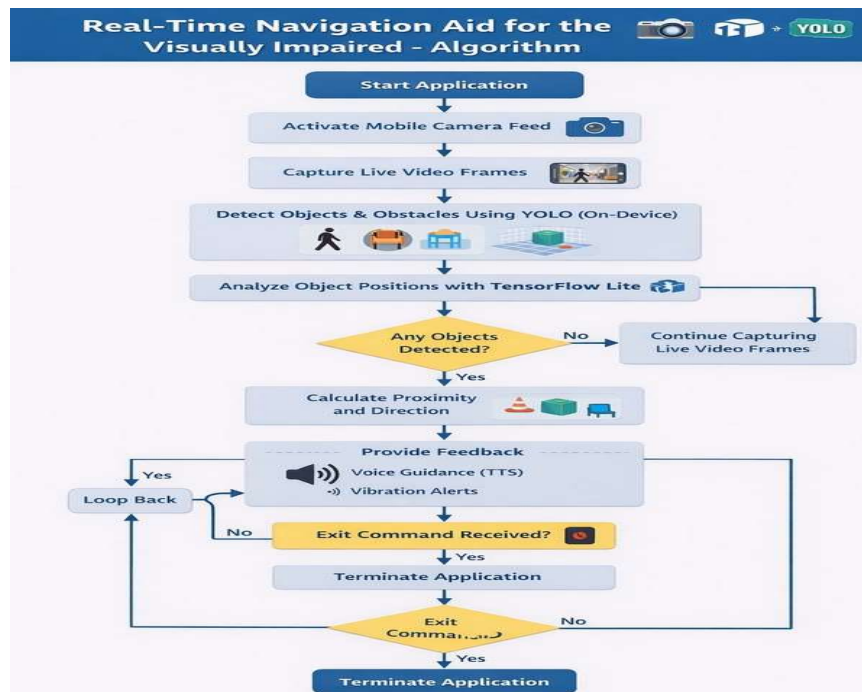


Fig. 2; Technical Architecture

Algorithm



System Modules

The proposed navigation application is organized into multiple functional modules, where each module is responsible for a specific operation within the overall system workflow. A modular design approach improves system maintainability, scalability, and independent component development. The primary modules of the proposed system are described below.

User Interaction Module

The User Interaction Module serves as the interface between the visually impaired user and the mobile application. This component manages basic user operations such as application activation,

termination, and accessibility control. To improve usability, the system supports simple interaction mechanisms, including voice-assistant-based launching and stopping functionality. The module ensures that users can access navigation services with minimal manual intervention.

Camera Processing Module

The Camera Processing Module is responsible for acquiring visual information from the smartphone environment. Using the device camera, the module continuously captures live video frames for further analysis.

To maintain reliable detection performance under different illumination conditions, the module

incorporates an adaptive lighting assistance mechanism. When low-light conditions are detected, the system automatically activates the smartphone torch or flash feature to improve image visibility. The processed camera frames are then forwarded to the object detection subsystem for AI-based analysis.

Object Detection Module

The Object Detection Module constitutes the core intelligence layer of the navigation system. This module employs lightweight deep learning models such as YOLO and TensorFlow Lite to perform real-time object recognition directly on the mobile device.

The AI model analyzes incoming camera frames and identifies surrounding obstacles, objects, and environmental elements relevant to navigation. For each detected object, the system generates bounding boxes and classification labels that represent object type and spatial position within the captured scene.

Direction and Proximity Analysis Module

The Direction and Proximity Analysis Module interprets detection results to estimate object location and relative distance from the user. Based on bounding-box coordinates and frame positioning, the module determines whether a detected object is located on the left, center, or right side of the user’s field of view.

In addition, obstacle proximity is estimated using scaling techniques derived from bounding-box

dimensions and relative object occupancy within the image frame. This analysis supports context-aware navigation decisions and enhances user understanding of environmental conditions.

Feedback Module

The Feedback Module converts processed environmental information into accessible guidance cues for the visually impaired user. The module employs a multimodal feedback mechanism combining both audio and haptic communication.

Audio alerts are generated using a Text-to-Speech (TTS) engine that communicates object identity, direction, and navigation instructions through spoken messages. Simultaneously, tactile awareness is provided through vibration patterns implemented using the Android Vibrator API. Vibration intensity and duration may vary according to obstacle proximity and navigational urgency.

Background Processing Module

The Background Processing Module enables continuous application execution independent of screen activity. The system operates as an Android Foreground Service, allowing essential functionalities such as object detection, analysis, and feedback generation to continue running even when the device display is turned off.

This module ensures uninterrupted navigation assistance, improves system reliability, and supports practical real-world deployment for visually impaired users during prolonged use.

Screenshots



Fig 3; The App

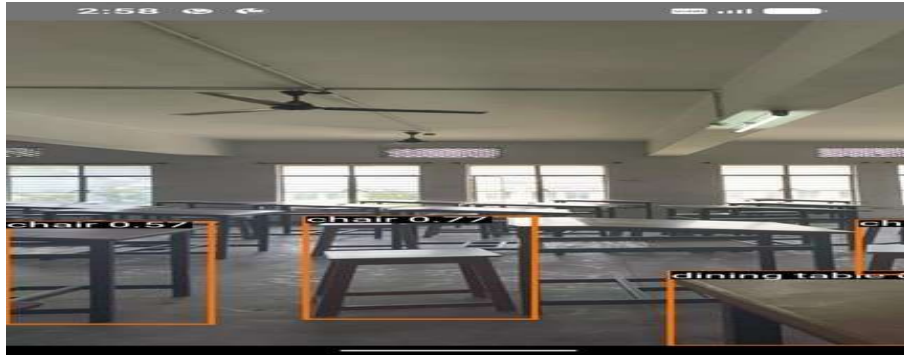


Fig 4; Objects Identification



Fig 5; Person Identification



Fig 6;Far off Object Identification



Fig 7;When lightning is unavailable, it ON torch automatically

Test Cases

The proposed navigation system was evaluated using multiple functional test cases to verify real-time object detection, feedback generation, processing stability, and distance estimation performance. The test results are summarized below.

Test Case ID	Test Scenario	Input / Action	Expected Result	Observed Result	Status
TC_06	Real-Time Object Detection	Present an object in front of the smartphone camera	System should detect the object successfully in real time	Object detected successfully	Pass
TC_07	Multiple Object Detection	Present multiple objects simultaneously to the camera	System should detect all visible objects with labels	Multiple objects detected correctly	Pass
TC_08	Bounding Box Generation	Show detectable objects to the camera	Bounding boxes should be displayed around detected objects	Bounding boxes displayed correctly	Pass
TC_09	Continuous Real-Time Processing	Execute the application continuously for an extended duration	Detection should continue without interruption	Stable detection and processing observed	Pass
TC_10	Audio Feedback Using TTS	Detect an object and analyze its position	System should announce object name and direction through voice output	Clear audio feedback generated	Pass
TC_11	Distance Estimation for Distant Objects	Place obstacles at larger distances from the camera	System should estimate object distance accurately	Reduced accuracy observed for far-away objects	Fail

Test Case Analysis

The experimental evaluation demonstrates that the proposed navigation application performs effectively in most core functional areas. Real-time object detection, multi-object recognition, bounding box visualization, and continuous background execution produced satisfactory results during testing. The Text-to-Speech feedback mechanism successfully generated understandable voice guidance regarding detected objects and their directional placement.

However, limitations were observed in distance estimation for distant obstacles. The current proximity estimation method, which relies primarily on bounding-box scaling, exhibited reduced accuracy when objects were positioned far from the camera. This limitation indicates an opportunity for future improvement through enhanced depth estimation methods, sensor fusion techniques, or advanced computer vision approaches.

The proposed AI-powered real-time navigation aid demonstrates an effective and practical solution for assisting visually impaired individuals in achieving safer and more independent mobility. By integrating smartphone-based computer vision with lightweight on-device deep learning frameworks such as YOLO and TensorFlow Lite, the system enables continuous object detection and environmental understanding without requiring persistent internet connectivity. This offline capability increases system reliability and allows operation across both indoor and outdoor environments, including locations with limited network access.

The combination of multimodal feedback mechanisms significantly enhances user interaction and situational awareness. Audio guidance generated through Text-to-Speech technology provides clear information regarding detected objects and navigation directions, while vibration-based haptic feedback conveys obstacle proximity and spatial awareness. These complementary feedback methods enable users to interpret

Conclusion

environmental conditions efficiently and make informed navigation decisions in real time.

Additional functionalities, including automatic torch activation under low-light conditions and Android Foreground Service execution, contribute to the robustness and stability of the application during continuous operation. Voice-assistant compatibility further improves accessibility by enabling hands-free application control.

Experimental evaluation indicates that the proposed system successfully fulfills its objective of delivering a portable, affordable, and intelligent navigation assistant for visually impaired users. The project demonstrates how mobile artificial intelligence, computer vision, and accessibility-centered system design can be combined to create practical assistive technologies that strengthen personal safety, environmental awareness, and independent daily navigation.

Future Scope

The future development of AI-based navigation systems for visually impaired individuals presents significant opportunities for technological enhancement and broader real-world adoption. Advances in computer vision, deep learning optimization, and edge AI processing are expected to improve detection accuracy, environmental understanding, and real-time responsiveness while maintaining efficient on-device execution.

Future versions of the system may incorporate advanced scene analysis, semantic environment recognition, and improved obstacle avoidance algorithms to support more intelligent navigation decisions. Integration with wearable technologies, including smart glasses, smart earphones, and enhanced haptic devices, could provide richer and more immersive guidance experiences.

The incorporation of emerging technologies such as 5G networks, Internet of Things (IoT) infrastructure, and smart-city communication systems may enable interaction with traffic signals, public transportation systems, and indoor navigation services. These developments could significantly expand mobility support in urban environments.

Further improvements may include multilingual voice interaction, personalized user adaptation using machine learning, and customized navigation preferences based on individual behavioral patterns. Research efforts may also focus on optimizing power consumption, reducing computational overhead, and improving affordability to increase accessibility for a larger user population.

Overall, future advancements in mobile AI and assistive computing have the potential to transform

navigation assistance systems into highly intelligent, adaptive, and universally accessible mobility solutions for visually impaired individuals.

References

- [1] M. Joshi, A. Shukla, J. Srivastava, and M. Rastogi, "DRISHTI: Visual Navigation Assistant for Visually Impaired," 2023. [Online]. Available: <https://arxiv.org/abs/2303.07451>
- [2] Y. Chen and J. Shen, "A Wearable Assistive System for the Visually Impaired Using Object Detection, Distance Measurement and Tactile Presentation," *Intelligent Robotics Journal*, 2023. [Online]. Available: <https://www.oaepublish.com/articles/ir.2023.24>
- [3] A. B. Atallah *et al.*, "An Effective Obstacle Detection System Using Deep Learning Advantages to Aid Blind and Visually Impaired Navigation," *Ain Shams Engineering Journal*, 2023. doi:10.1016/j.asej.2023.102387
- [4] M. A. Kamran *et al.*, "Visually: Assisting the Visually Impaired People Through AI-Assisted Mobility," *International Journal of Innovations in Science & Technology*, 2024. [Online]. Available: <https://journal.50sea.com/index.php/IJIST/article/view/793>
- [5] G. I. Okolo, T. Althobaiti, and N. Ramzan, "Assistive Systems for Visually Impaired Persons: Challenges and Opportunities for Navigation Assistance," *Sensors*, vol. 24, no. 11, p. 3572, 2024. doi:10.3390/s24113572
- [6] "Mobile Application with YOLO and Text-to-Speech Based Navigation Support," *International Research Journal of Advanced Engineering and Technology (IRJAET)*, 2025.
- [7] "AI-Based Navigation System for the Visually Impaired," *Scientific Reports*, 2025. [Online]. Available: <https://www.nature.com/articles/s41598-025-29947-7>
- [8] "AI-Based Navigation System Using YOLO and TensorFlow," Research Publication, 2025. [Online]. Available: <https://www.researchgate.net/publication/398981777>
- [9] "Edge AI-Based Navigation System for Visually Impaired Users," Elsevier / ScienceDirect, 2025. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S2590123025027963>
- [10] "TensorFlow Lite-Based Mobile Navigation System for Accessibility Applications," Research Publication, 2026. [Online]. Available: <https://www.researchgate.net/publication/401315259>