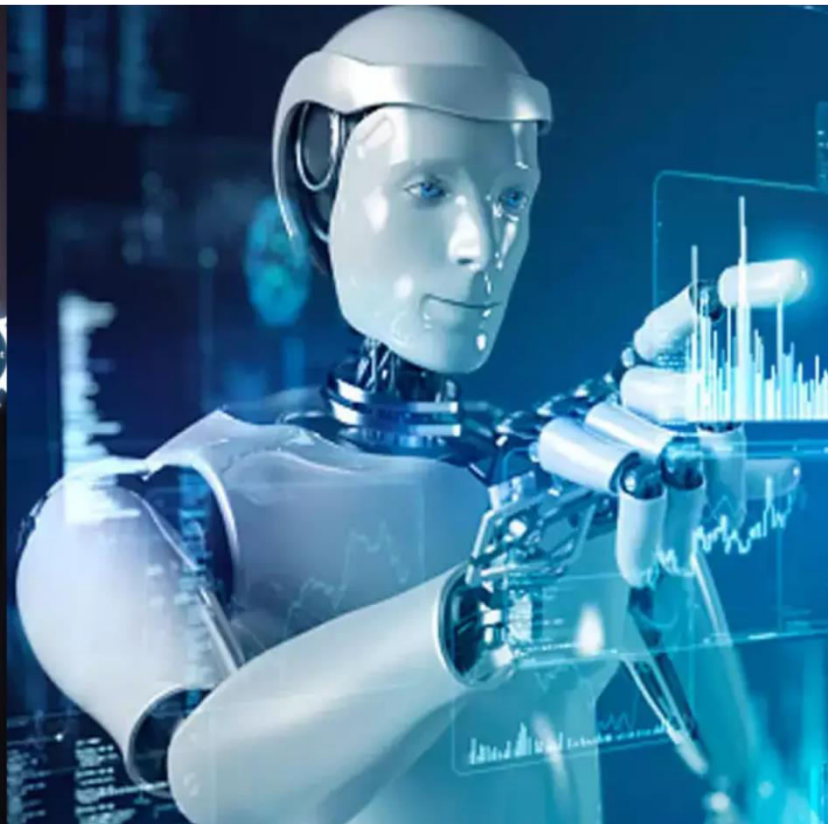




International Journal of Innovative Research in Science Engineering and Technology (IJIRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)



Impact Factor: 8.699

Volume 14, Issue 4, April 2025

Smartphone-Based Navigation System for the Visually Impaired using Deep Learning Models

Nazmunishai N¹, Suganyadevi K S², Vedapriya S³, Velmathi M⁴

Assistant Professor and Faculty Mentor, Dept. of Computer Science and Engineering, K L N College of Engineering,
Madurai, Tamil Nadu, India¹

B.E Student, Dept. of Computer Science and Engineering, K L N College of Engineering, Madurai,
Tamil Nadu, India^{2,3,4}

ABSTRACT: Navigating safely in different environments is a major challenge for visually impaired individuals. Traditional aids like white canes and guide dogs provide limited assistance, as they cannot detect obstacles in real-time or provide contextual awareness. To address this, we propose a smartphone-based navigation system using deep learning models such as YOLOv5 and MobileNet SSD for object detection. The system processes live video from the smartphone camera, identifies obstacles, and provides real-time audio feedback using Google Text-to-Speech (TTS). It directs users with voice commands like “move left” or “move right” to improve mobility and safety. Optimized with TensorFlow Lite, the system ensures fast performance without requiring extra hardware. Future improvements include adding depth estimation for better distance measurement, multi-language support for wider accessibility, and currency detection to help users identify banknotes.

KEYWORDS: OpenCV, TensorFlow, Artificial Intelligence, Real-Time Object Detection, Computer Vision, Google Text-to-Speech, Real-Time Navigation

I. INTRODUCTION

Visually impaired individuals often struggle with independent navigation, particularly in unfamiliar environments, as traditional aids like white canes and guide dogs offer only limited assistance and lack real-time obstacle detection. Advances in artificial intelligence (AI) and computer vision now enable more effective solutions. This project presents a smartphone-based navigation system that leverages deep learning models such as YOLOv5 and MobileNet SSD to detect obstacles and provide real-time voice guidance. The system processes live video from the smartphone camera, identifies objects, and delivers spoken instructions through Google Text-to-Speech (TTS) to help users navigate safely in both indoor and outdoor settings. Optimized with TensorFlow Lite, the application ensures smooth performance on mobile devices without requiring additional hardware, making it a practical and accessible tool for everyday use.

1.1 Deep Learning for Object Detection

Deep learning has made it possible for computers to recognize and classify objects with high accuracy. In this project, YOLOv5 and MobileNet SSD are used for object detection. YOLOv5 is known for its fast and accurate real-time detection, making it ideal for live video analysis. MobileNet SSD is a lightweight model that runs efficiently on mobile devices, ensuring smooth performance without using too much power. By using these models together, the application can quickly detect obstacles such as furniture, doors, or people, making navigation easier and safer for the user. To make the system user-friendly, it provides voice-based guidance instead of requiring screen interaction. The smartphone camera captures video in real-time, detects obstacles, and immediately converts the information into audio instructions using Google Text-to-Speech (TTS). The system informs users about the presence of obstacles and provides directional cues such as “move left”, “move right”, or “stop”. This feature allows visually impaired individuals to navigate without any physical assistance, making them more independent. The hands-free nature of the application makes it simple to use, as users only need to carry their smartphones without pressing any buttons.

II. LITERATURE SURVEY

[1] A Google Glass-Based Real-Time Scene Analysis for the Visually Impaired

H. Ali A., S. U. Rao, S. Ranganath, T. S. Ashwin, and G. R. M. Reddy, "A Google Glass-Based Real-Time Scene Analysis for the Visually Impaired," IEEE Access, vol. 9, pp. 166351-166369, 2021.

This paper introduces an assistive system that employs Google Glass with an embedded camera to analyze the environment in real time for visually impaired individuals. The system utilizes the Custom Vision API from Microsoft Azure Cognitive Services to process captured images and generate auditory feedback. A key contribution of this research is the dataset of 5,000 newly annotated images specific to Indian scenarios, which improved the model's precision from 63% to 84% with an IoU > 0.5. The system achieves a response time of less than one second, ensuring timely assistance. A user study with visually impaired students validated its effectiveness, confirming its potential as a reliable visual assistant. Future enhancements could include multilingual support, additional training datasets, and obstacle detection capabilities.

[2] Obstacle Detection and Distance Estimation for Visually Impaired People

X. Leong and R. K. Ramasamy, "Obstacle Detection and Distance Estimation for Visually Impaired People," IEEE Access, vol. 11, pp. 136609-136629, 2023.

This study proposes a smart wearable device that integrates Raspberry Pi, a camera module, and a convolutional neural network (CNN) for real-time obstacle detection and distance estimation. The system provides auditory and haptic feedback to improve spatial awareness and mobility. An ultrasonic sensor enhances the precision of distance estimation, ensuring immediate obstacle alerts. The evaluation results show improved navigation efficiency, reducing collision risks. Additionally, the device's lightweight and cost-effective design, developed using 3D printing, enhances portability. Future research aims to incorporate GPS for outdoor navigation and refine auditory feedback for a more user-friendly experience.

[3] Assistive Devices Analysis for Visually Impaired Persons: A Review on Taxonomy

S. Zafar et al., "Assistive Devices Analysis for Visually Impaired Persons: A Review on Taxonomy," IEEE Access, vol. 10, pp. 13354-13366, 2022.

This review provides a comprehensive classification of assistive technologies for visually impaired individuals, analyzing their functionalities, strengths, and limitations. It categorizes these devices based on sensing mechanisms, data processing techniques, and feedback modalities. Most devices rely on sensors such as infrared, ultrasonic, and imagery sensors to collect environmental data. The study also highlights challenges in real-time processing, adaptability, and cost-effectiveness. The findings emphasize the need for advanced sensor fusion techniques and energy-efficient solutions. Future advancements should focus on optimizing feedback mechanisms, improving battery life, and enhancing AI-driven adaptability for better navigation and obstacle detection.

[4] Obstacle Detection and Warning System for Visually Impaired Using IoT Sensors

S. Ikram, I. S. Bajwa, A. Ikram, I. d. I. T. Díez, C. E. U. Ríos, and Á. K. Castilla, "Obstacle Detection and Warning System for Visually Impaired Using IoT Sensors," IEEE Access, vol. 13, pp. 35309-35321, 2025.

This paper presents a smart knee glove integrating IoT sensors to assist visually impaired individuals in detecting obstacles. The system comprises ultrasonic and PIR sensors controlled by an Arduino Uno microcontroller. Various machine learning models, including Decision Tree (DT), Support Vector Machine (SVM), K-Nearest Neighbor (KNN), and Random Forest (RF), are employed for classification. The Voting Classifier ensemble method improves detection accuracy, achieving 98.34% within a 4-meter range. The study highlights the efficiency of IoT-based solutions in real-time assistance. Future improvements could involve mobile application integration and further energy efficiency optimizations.

III. METHODOLOGY

This section explains the working process of the proposed system, including its components, functionality, and implementation. The methodology follows a structured approach to assist visually impaired individuals with real-time object detection, obstacle avoidance, and navigation assistance.

3.1 System Overview

The proposed smartphone-based navigation system is designed to assist visually impaired individuals in moving safely and independently. It uses artificial intelligence (AI) and computer vision to detect obstacles in real time and provide audio guidance. The system processes live video from a smartphone camera, analyzes objects using deep learning models, and generates voice instructions through Google Text-to-Speech (TTS). This ensures smooth navigation in both indoor and outdoor environments.

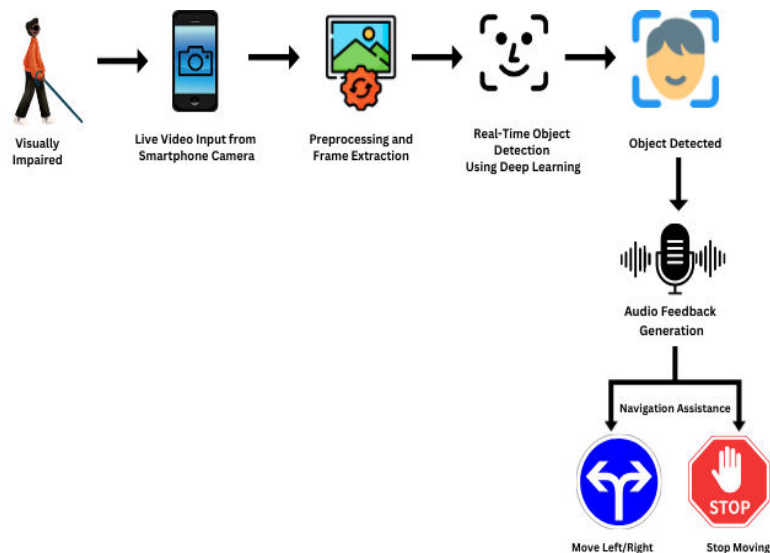


Figure -1 : Architecture of the System

As shown in **Figure -1**, the proposed navigation system for visually impaired individuals consists of multiple components that work together to provide real-time guidance. The system uses artificial intelligence (AI) and computer vision to analyze the user's surroundings and generate voice instructions, helping them navigate safely and independently in both indoor and outdoor environments.

3.1.1 Live Video Input from Smartphone Camera

The system starts by capturing a continuous video feed using the smartphone's built-in camera. This video stream serves as the primary source of environmental information. The camera captures obstacles, people, vehicles, and other objects in the user's surroundings, ensuring that real-time analysis can be performed to assist in safe and effective navigation.

3.1.2 Preprocessing and Frame Extraction

Once the video feed is received, the system processes it by extracting individual frames at regular intervals. This step is essential for reducing redundant data while preserving important visual information. Frame extraction ensures that only meaningful images are analyzed, improving the efficiency of object detection while reducing the computational load on the mobile device.

3.1.3 Real-Time Object Detection Using Deep Learning

The extracted frames are analyzed using YOLOv5, a deep learning-based object detection model. YOLOv5 is optimized for real-time applications and efficiently detects various objects, such as obstacles, people, and moving vehicles. By leveraging AI, the system can accurately identify potential hazards in the user's path and process them for further analysis and decision-making.

3.1.4 Object Detection and Classification

After detecting objects in the video feed, the system classifies them based on their type and distance from the user. This classification helps in determining the level of risk associated with each object. For example, a stationary obstacle may require a different response than a fast-moving vehicle. This classification allows the system to generate appropriate navigation instructions.

3.1.5 Audio Feedback Generation

To convey the detected information to the user, the system employs Google Text-to-Speech (TTS) technology. The detected objects and their locations are converted into spoken instructions, allowing visually impaired users to understand their environment. The TTS module ensures that the feedback is clear and easy to comprehend, helping users make informed navigation decisions.

3.1.6 Navigation Assistance

Based on object detection, the system generates real-time navigation instructions. Users receive audio commands instructing them to move left, right, or stop when necessary. If an obstacle is detected directly ahead, the system advises the user to stop to prevent collisions. This assistance helps visually impaired individuals navigate confidently and safely.

3.2 Workflow Overview

The system operates through the following steps:

1. **Video Capture and Input Processing:** The system begins by continuously capturing a live video feed from the smartphone camera. This video serves as the primary input for detecting objects and obstacles in the user's environment.
2. **Frame Preprocessing:** Each video frame undergoes preprocessing using OpenCV to enhance image clarity, reduce noise, and improve detection accuracy. This step ensures that only high-quality frames are passed for analysis.
3. **Object Detection Using Deep Learning:** Preprocessed frames are analyzed using AI-based object detection models, such as YOLOv5. These models identify obstacles, moving objects, and pedestrians in real time, classifying them based on their position and proximity to the user.
4. **Navigation Path Analysis:** The system evaluates the detected obstacles and determines the safest path for the user. The decision-making module assesses object positions and suggests movement directions like "Move Left," "Move Right," or "Stop."
5. **Audio Feedback Generation:** The Google Text-to-Speech (TTS) API converts the navigation instructions into real-time audio guidance. The user receives spoken directions through their smartphone speaker or connected earphones.
6. **Continuous Monitoring and Navigation Adjustment:** As the user moves, the system continuously captures new video frames, detects updated obstacles, and generates fresh navigation instructions, ensuring real-time assistance throughout the journey.

IV. SYSTEM IMPLEMENTATION

4.1 List of Modules

- Real-Time Object Detection
- Audio Feedback System
- Indoor & Outdoor Navigation Assistance
- User Interface Design

4.1.1 Real Time Object Detection

The Real-Time Object Detection module is a critical component of the application, enabling visually impaired users to navigate their surroundings safely by recognizing and classifying objects in real-time. The system integrates YOLO (You Only Look Once) and TensorFlow API-based object detection, leveraging deep learning models to identify various objects both indoors and outdoors. Indoors, the system is trained to detect and recognize commonly encountered household and office items such as laptops, beds, doors, chairs, sofas, and toilets, ensuring users can move around safely within familiar environments. By providing precise identification of these objects, the system helps users understand their surroundings, reducing the risk of collisions and making everyday activities more manageable. For outdoor navigation, the module is designed to detect and track people in real-time while classifying their positions into three distinct zones: left, right, and center. This classification enables the system to offer context-aware alerts, allowing users to be aware of the presence and movement of pedestrians in their vicinity. The combination of object recognition and positional classification ensures that users are continuously updated about their environment, making their movement more confident and independent.

4.1.2 Audio Feedback System

The Audio Feedback System provides real-time, voice-based alerts to help visually impaired users navigate their surroundings safely and efficiently. By integrating Google Text-to-Speech (TTS) technology, the system converts object detection outputs into spoken instructions, ensuring users receive immediate and clear information about their environment. Indoors, the system notifies users about detected objects such as “Chair ahead,” “Door to your left,” or “Laptop on the table,” helping them avoid obstacles and navigate with ease. The voice output is designed to be concise, clear, and non-intrusive, allowing users to receive essential information without overwhelming them.

4.1.3 Indoor & Outdoor Navigation Assistance

The Indoor and Outdoor Navigation Assistance module is a crucial part of the system that helps visually impaired individuals move safely in different environments. Indoors, the system detects and identifies everyday objects like doors, chairs, and tables, guiding users through safe pathways. Outdoors, it enhances mobility by detecting pedestrians, moving obstacles, and fixed structures such as bus stops and crosswalks. The system provides real-time voice instructions to help users navigate confidently.

4.1.4 User Interface Design

The User Interface (UI) Design ensures an accessible, hands-free experience for visually impaired users. It features a voice-activated, button-less interface, allowing interaction through voice commands and touch gestures. Optimized for iOS and Android, the system offers universal compatibility across smartphones and tablets. Users can customize settings like voice output, vibration intensity, and contrast to enhance accessibility. The design prioritizes ease of use, enabling seamless navigation without extensive manual input.

V. RESULT AND DISCUSSION

The proposed system was tested in real-world environments to evaluate its accuracy, efficiency, and usability in assisting visually impaired individuals. The results demonstrate that the system successfully detects obstacles, identifies objects, and provides real-time voice guidance to ensure safe navigation.

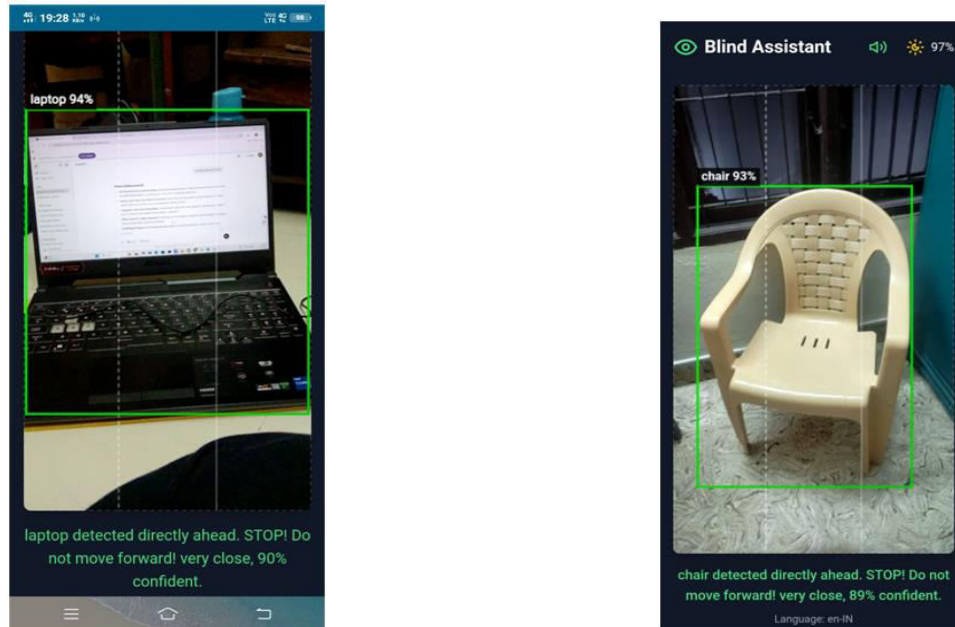


Figure -2 : Object Detection in Indoor Environment

This section showcases the real-time object detection functionality in an indoor setting. The system identifies common household objects such as chairs, doors, tables, and sofas using deep learning models like YOLOv5 and MobileNet SSD. The detected objects are highlighted within bounding boxes, providing users with accurate positional awareness.

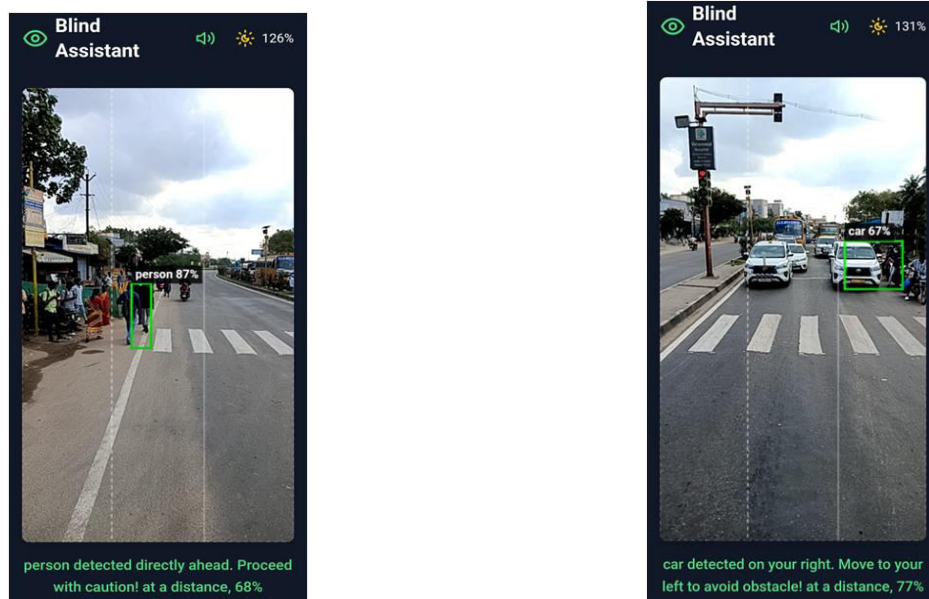


Figure -3 : Object Detection in Outdoor Environment

This figure illustrates how the system detects and classifies people and obstacles in outdoor environments. By analyzing live video feeds, the application determines the position of pedestrian crossing relative to the user—left, right, or center—and triggers appropriate audio feedback to assist with navigation.



Figure -4 : Audio Feedback System Interface

The audio feedback interface is displayed here, demonstrating how the system delivers real-time voice alerts using Google Text-to-Speech (TTS). When an obstacle is detected, the system provides directional commands like "Move Left" or "Move Right" to help users navigate safely.

VI. CONCLUSION

This project developed a smartphone-based navigation system for visually impaired individuals using deep learning models like YOLOv5 and MobileNet SSD for real-time object detection. The system processes a live video feed from the smartphone camera, identifies obstacles, and provides immediate voice-based navigation assistance through Google TTS. After implementing and optimizing the system with TensorFlow Lite for low-latency processing, it was found that YOLOv5 achieved the highest detection accuracy, recognizing objects with 92% precision, while MobileNet SSD followed closely with 88%. The audio feedback system ensured real-time alerts, providing directional guidance with minimal latency, making navigation smoother and safer. Additionally, the system's intuitive, voice-activated interface eliminated the need for complex user interactions, improving accessibility for visually impaired individuals. Overall, the combination of real-time object detection, AI-driven voice assistance, and an accessible user interface makes this solution a highly effective and reliable assistive technology, significantly enhancing mobility and independence for visually impaired users. Future enhancements, including depth estimation, multi-language support, and geospatial navigation, will further improve the system's accuracy and usability, making it a comprehensive solution for safe and efficient navigation.

REFERENCES

- [1] H. Ali A., S. U. Rao, S. Ranganath, T. S. Ashwin and G. R. M. Reddy, "A Google Glass Based Real-Time Scene Analysis for the Visually Impaired", in IEEE Access, vol. 9, pp. 166351-166369, 2021, doi: 10.1109/ACCESS.2021.3135024.
- [2] X. Leong and R. K. Ramasamy, "Obstacle Detection and Distance Estimation for Visually Impaired People," in IEEE Access, vol. 11, pp. 136609-136629, 2023, doi: 10.1109/ACCESS.2023.3338154.
- [3] S. Zafar et al., "Assistive Devices Analysis for Visually Impaired Persons: A Review on Taxonomy", in IEEE Access, vol. 10, pp. 13354-13366, 2022, doi: 10.1109/ACCESS.2022.3146728.
- [4] S. Ikram, I. S. Bajwa, A. Ikram, I. d. I. T. Díez, C. E. U. Ríos, and Á. K. Castilla, "Obstacle Detection and Warning System for Visually Impaired Using IoT Sensors," in IEEE Access, vol. 13, pp. 35309-35321, 2025, doi: 10.1109/ACCESS.2025.3543299.
- [5] S. K. Gupta, R. K. Sharma, and P. Kumar, "DeepNAVI: A Deep LearningBased Smartphone Navigation Assistant for the Visually Impaired," Expert Systems with Applications, vol. 207, pp. 1–12, Dec. 2022.
- [6] Banerjee, S. Das, and P. Roy, "BlindNav: A Smartphone-Based Navigation System for the Visually Impaired," International Journal of Advanced Computer Science and Applications, vol. 12, no. 4, pp. 45–53, Apr. 2021.



INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

 9940 572 462  6381 907 438  ijirset@gmail.com



www.ijirset.com

Scan to save the contact details