Indian Journal of Computer Science and Technology
https://www.doi.org/10.59256/indjcst.20260502045
Volume 5, Issue2 (May-August 2026), PP: 400-409.
www.indjcst.com
ISSN No: 2583-5300

Voice Based Assistant for Visually Impaired Using Machine Learning Techniques

Mangala S Biradar[1], Shilpa Rawale[2], Manjiri Dongare[3], Bhovati Rathod[4], Akanksha Choure[5] Sakshi Raut[6]

[1]*Professor, SRCOE, Department of Computer Engineering, Pune, Maharashtra, India.*
[2,3,4,5,6]*Student, SRCOE, Department of Computer Engineering, Pune, Maharashtra, India.*

To Cite this Article: *Mangala S Biradar [1], Shilpa Rawale[2], Manjiri Dongare[3], Bhovati Rathod[4], Akanksha Choure[5] Sakshi Raut[6], "Voice Based Assistant for Visually Impaired Using Machine Learning Techniques", Indian Journal of Computer Science and Technology, Volume 05, Issue 02 (May-August 2026), PP: 400-409*

Abstract *In the field of artificial intelligence and computer vision, object detection has become a fundamental technology for enabling machines to understand and interpret visual data. Object detection involves identifying and locating objects within images or video streams by assigning labels and bounding boxes. It plays a critical role in applications such as autonomous driving, surveillance, healthcare, and assistive technologies for visually impaired individuals. Various algorithms have been developed to improve the accuracy, speed, and efficiency of object detection systems. Among the most prominent are Region-Based Convolutional Neural Networks (R-CNN), Faster R-CNN, and You Only Look Once (YOLO), each offering unique approaches to object localization and classification. This paper explores and compares these algorithms in terms of architecture, working principles, and real-world performance. The study highlights how modern deep learning techniques, especially YOLO, enable real-time object detection with high efficiency and scalability.*

Key Words: *Object Detection, R-CNN, Faster R-CNN, YOLO, Computer Vision, Deep Learning.*

I. INTRODUCTION

The Blind Assistance Mobile Application is an AI-powered mobile application developed using the Flutter framework to assist visually impaired individuals in their daily activities. The main aim of this application is to improve safety, independence, and accessibility by helping users understand and interact with their surroundings more effectively. The system uses advanced technologies such as computer vision, machine learning, and speech processing to provide real-time assistance. By using the smartphone camera, the application detects nearby objects such as people, vehicles, doors, obstacles, and other important items through object detection algorithms like YOLO. After detecting objects, the application converts visual information into voice feedback using Text-to-Speech (TTS) technology, enabling users to hear clear instructions and descriptions of their surroundings. The application also supports voice command s and speech recognition, making it easy to use in a hands-free manner. Since the application works on mobile devices, it is portable, user- friendly, and accessible anytime and anywhere. It supports both indoor and outdoor navigation, helping visually impaired users move safely and independently in different environments. The application is designed with a simple and easy-to-use interface so that users can operate it without technical difficulties. It can also help users identify daily-use objects such as chairs, stairs, doors, and personal belongings. The system reduces dependency on others and improves confidence during travel and routine activities. In the future, the application can be enhanced with features like GPS navigation, facial recognition, text reading, and smart wearable integration for better assistance and accuracy. The application also includes speech recognition and voice command functionality, enabling users to interact with the system in a hands-free manner. Users can give voice instructions to activate features, request information, or control the application without needing visual interaction. Since the system is developed for smartphones, it is portable, cost-effective, and easy to access anytime and anywhere.

II.PROBLEM STATEMENT

To develop a voice-based assistant that helps visually impaired people identify objects, read text and navigate safely.

III.OBJECTIVES

The main objectives of the project are:

- To help blind people detect surrounding objects.
- To provide real-time voice guidance.
- To improve independent navigation.
- To use Artificial Intelligence for object recognition.
- To create a simple and user-friendly mobile application.

IV.LITERATURE REVIEW

Deshpande et al. Rohan (2024): Deshpande et al. Rohan's 2024 work, titled Vision Assist: Intelligent Navigation System for the Visually Impaired, presents an AI-based assistive solution for mobility support. The system integrates computer vision and deep learning to detect obstacles and recognize objects in real time. It uses a wearable camera connected to a mobile device for continuous environmental monitoring. Unlike traditional tools, it focuses on contextual understanding of surroundings. GPS integration enables navigation and route guidance for outdoor movement. The system provides audio instructions such as "Move forward" and "Obstacle ahead." It works in both indoor and outdoor environments with good adaptability. However, it requires high computational power and stable internet connectivity. Performance may reduce in low-light conditions. Privacy concerns arise due to continuous image capture and processing.

Sharma et al. Ananya (2023): Sharma et al. Ananya's 2023 research, titled Smart Guide: IoT-Based Assistive System for Blind Navigation, introduces a system combining IoT and machine learning. It uses ultrasonic sensors to detect obstacles and provide real-time alerts. The system delivers feedback through both vibration and voice outputs. A mobile application with GPS integration enables location tracking and navigation. Machine learning improves obstacle classification accuracy over time. It also includes an emergency alert feature for safety. The design is cost-effective and user-friendly for daily use. However, battery life and hardware maintenance can be challenging. Environmental noise may affect voice guidance clarity. Overall, it enhances safety and independence for visually impaired users.

Verma et al. Karan (2022): Verma et al. Karan's 2022 study, titled DeepNav: AI-Powered Visual Recognition for Blind Assistance, focuses on deep learning for object detection. The system uses CNN models to identify objects like stairs, doors, and vehicles. It operates through a smartphone, making it affordable and accessible. It also includes text-to-speech for reading printed content. The system provides high accuracy due to trained datasets. It adapts to user behavior for better personalization. However, it requires high processing power for smooth functioning. Continuous camera use leads to battery drain issues. Performance may drop in complex environments. Despite this, it improves interaction with surroundings significantly.

Joshi et al. Meera (2021): Joshi et al. Meera's 2021 research, titled NavEye: Real-Time Assistance System Using Computer Vision, presents a wearable navigation device. The system uses a camera mounted on smart glasses for capturing images. These images are processed to detect objects and provide audio feedback. It helps users identify obstacles and navigate safely. The system also includes facial recognition features. It can identify known individuals and inform the user. Real-time processing ensures quick response with minimal delay. However, it requires high computational resources. Performance depends on lighting and environmental conditions. Cost and battery life are additional limitations.

Kumar et al. Suresh (2020): Kumar et al. Suresh's 2020 work, titled Audino: Voice-Based Navigation System for the Blind, focuses on voice-assisted navigation. The system uses GPS for route guidance and location tracking. Users can input destinations using voice commands. It provides step-by-step audio instructions during travel. Basic sensors help detect nearby obstacles for safety. The system is simple, lightweight, and easy to use. It is suitable for users with minimal technical knowledge. However, GPS accuracy may reduce in indoor environments. It lacks advanced AI-based object recognition features. Still, it improves mobility and independence effectively.

V.SOFTWARE/HARDWARE REQUIREMENTS

Software Requirements
- **Flutter SDK:** The latest Flutter SDK is required for cross-platform mobile application development.
- **Android Studio:** The latest Android Studio is required for coding, debugging, and testing the application.
- **Dart SDK:** The latest Dart SDK is required to support Flutter application development.
- **Tensor Flow Lite:** The latest TensorFlow Lite is required for implementing machine learning features.

Hardware Requirements
- **RAM:** A minimum of 4 GB RAM is required for efficient multitasking and performance.
- **Storage:** At least 10 GB of free storage space is required for software installation and project files.
- **Processor:** An Intel i3 processor or higher is required for smooth application development.
- **Android Phone Camera:** A camera-supported Android phone is required for image and video capturing features.

VI.SYSTEM ARCHITECTURE/DATA FLOW DIAGRAM

1. **Input Layer**
- Mobile camera captures live video
- The input layer captures real-time surroundings using the mobile camera.
- It continuously records video and converts it into frames (images).
- These frames act as raw data and are sent to the processing layer for object detection.

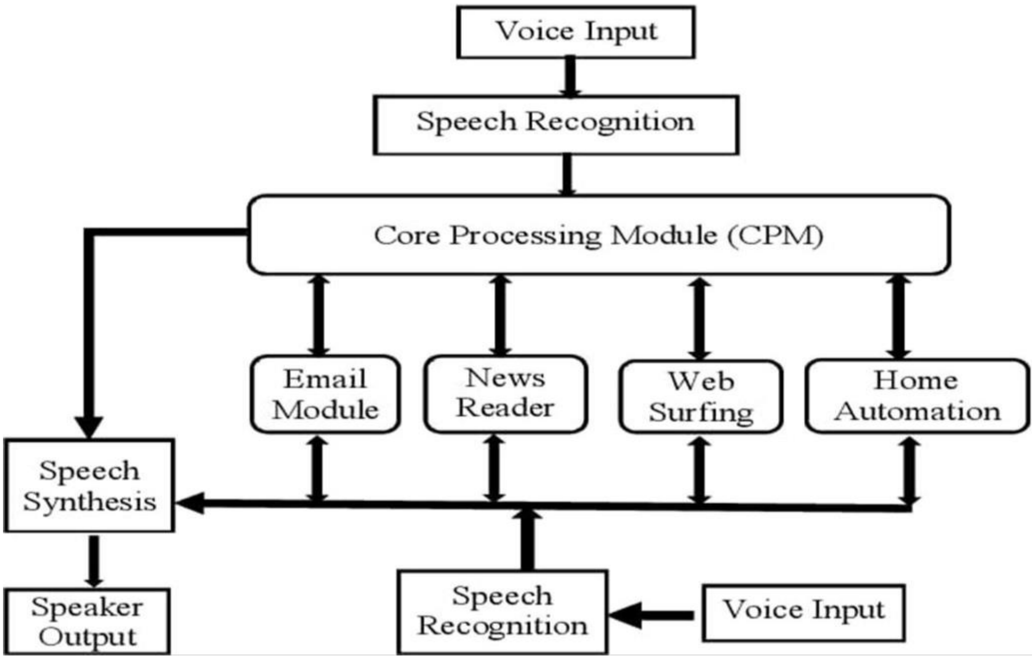

Fig1.1: System Architecture

2. Processing Layer

- Frames are converted into image data.
- TensorFlow Lite model processes images.
- YOLOv5 detects objects.

3. Output Layer

- The input layer captures real-time surroundings using the mobile camera.
- It continuously records video and converts it into frames (images).
- These frames act as raw data and are sent to the processing layer for object detection.

VII. UML DIAGRAM

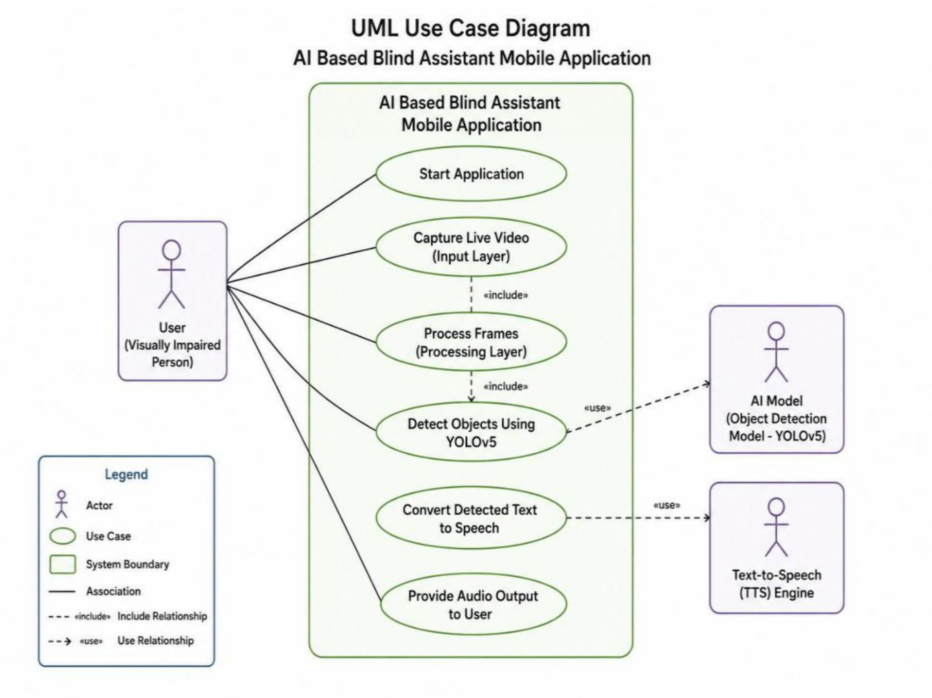

Fig1.2: UML Diagram

UML Diagram Description

The UML diagram represents the structure and workflow of the Voice Assistant System for Visually Impaired Users. It includes use case, class, sequence, and activity diagrams to explain system functionalities such as object detection, voice guidance, and audio feedback. The diagrams show communication between the user, mobile application, AI model, camera, and voice assistant modules.

- Works best on Android devices with camera and audio support.
- Helps visually impaired users in navigation and object recognition.

Types of UML Diagrams Use Case Diagram

- The Use Case Diagram represents the interaction between the visually impaired user and the application features such as object detection, voice guidance, speech recognition, and navigation assistance.
- It helps in understanding the functional requirements of the system and shows how users communicate with different services provided by the application.

Class Diagram

- The Class Diagram describes the structure of the system by showing different classes, attributes, methods, and relationships between modules like the Flutter application, camera module, TensorFlow Lite model, database, and voice assistant.
- It helps developers understand how system components are connected and how data flows between different modules

Sequence Diagram

- The Sequence Diagram illustrates the step-by-step communication between the user, mobile application, AI model, camera, and audio feedback system during system operation.
- It shows the order of message exchanges from capturing an image to detecting objects and generating voice output in real time.

Activity Diagram

- The Activity Diagram represents the complete workflow of the application, including image capture, object processing, object recognition, and voice feedback generation.
- It helps visualize the sequence of activities and decision-making processes involved in assisting visually impaired users during navigation and object identification

1. Use case Diagram

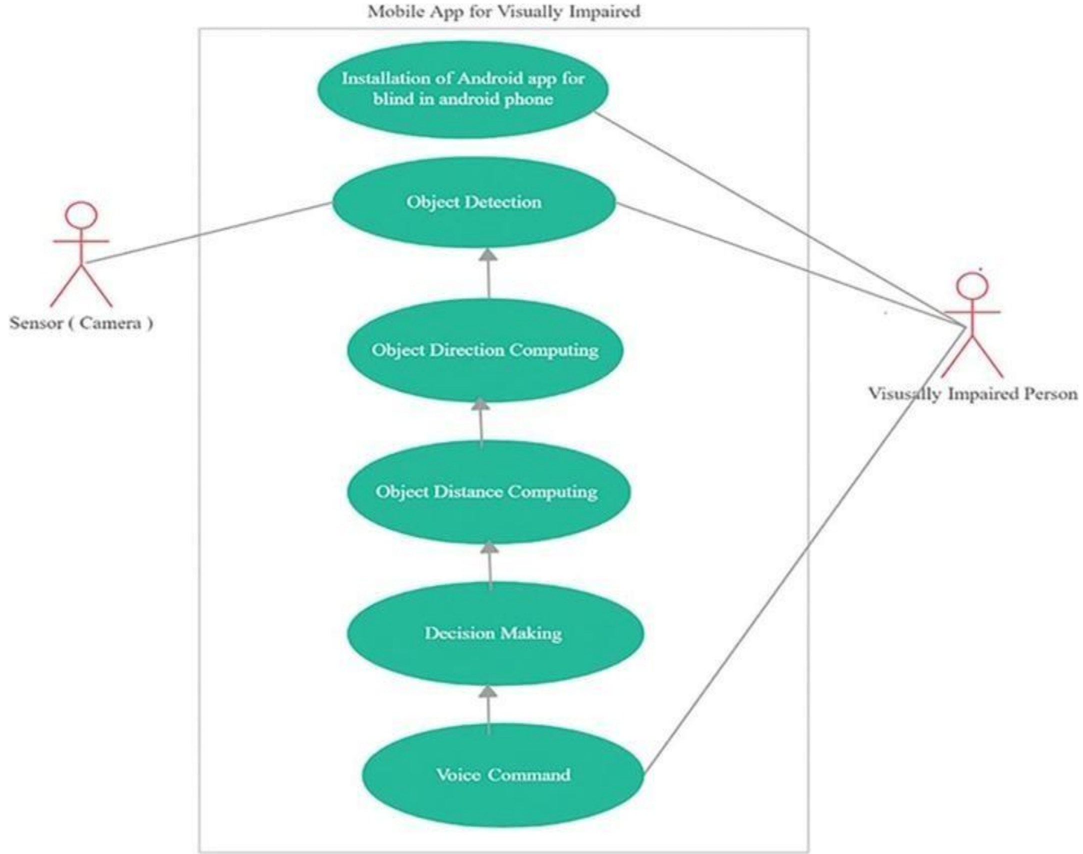

Fig1.3: Use case Diagram

2. Class Diagram

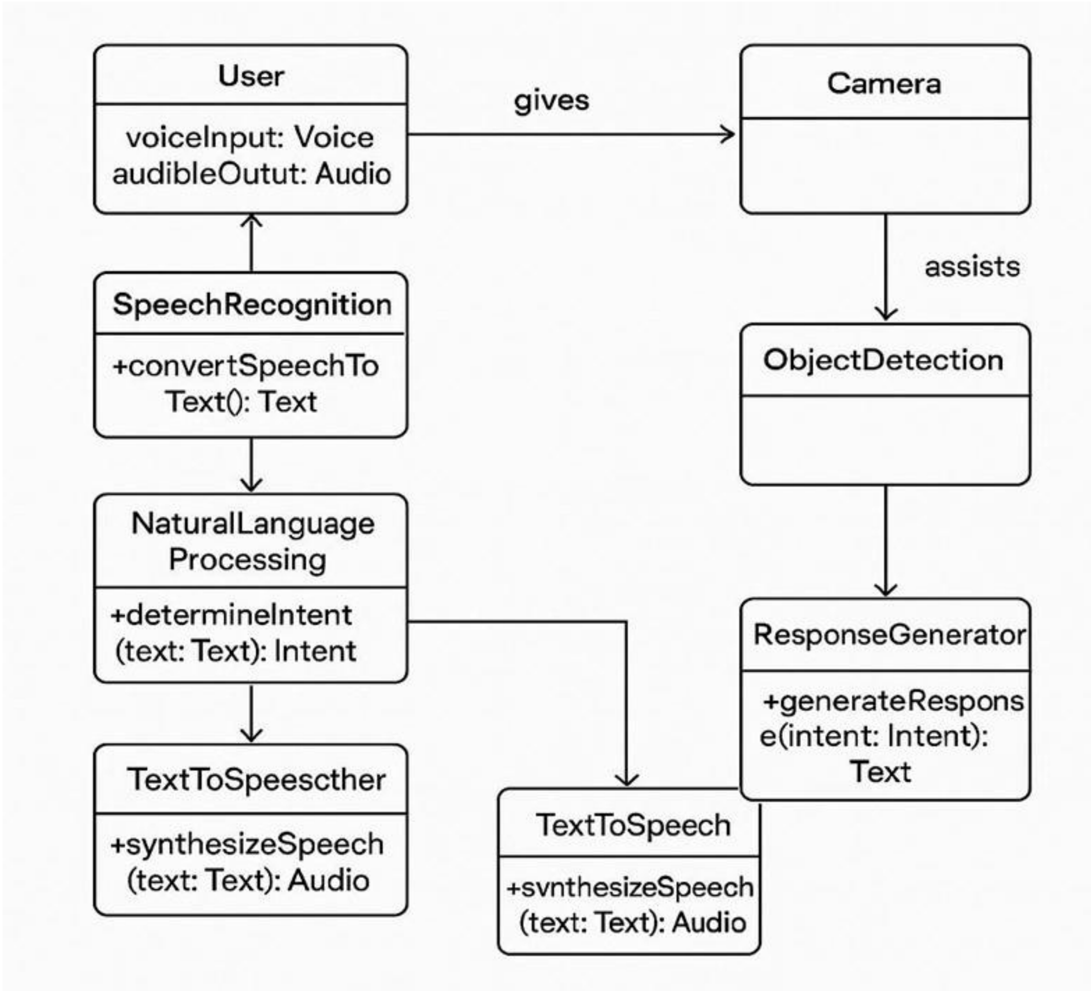

Fig1.4: Class Diagram

3. Sequence Diagram

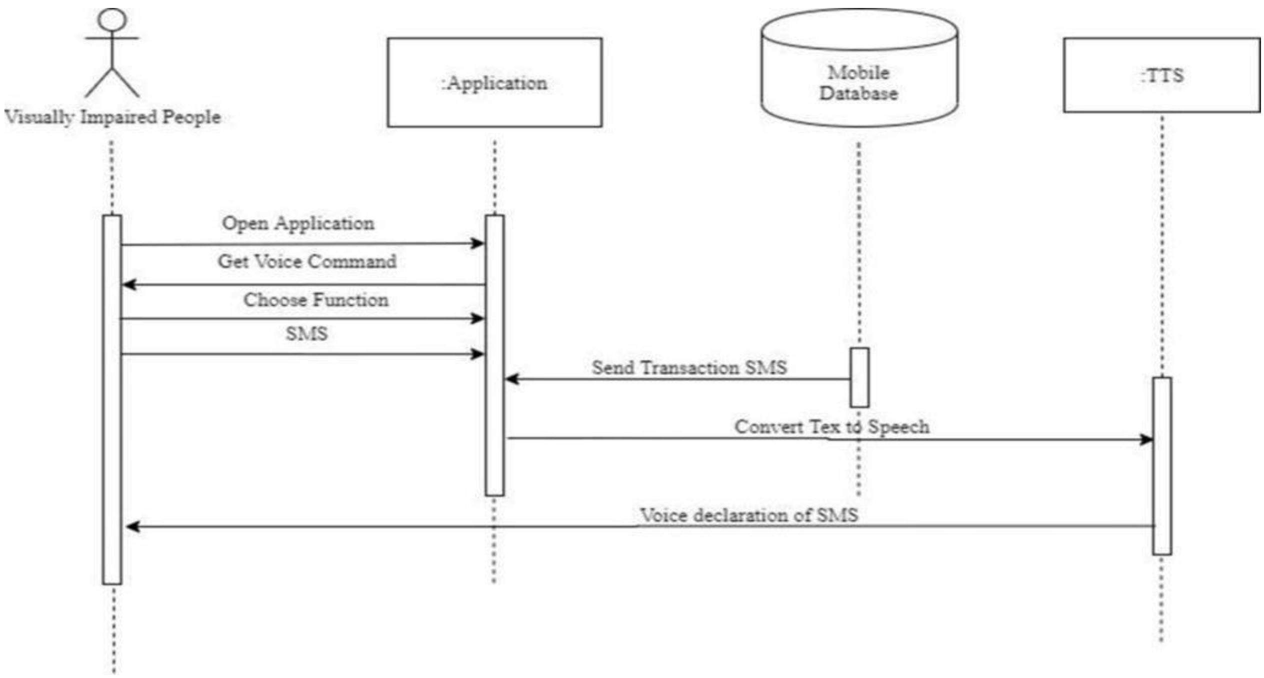

Fig1.5: Sequence Diagram

3. Activity Diagram

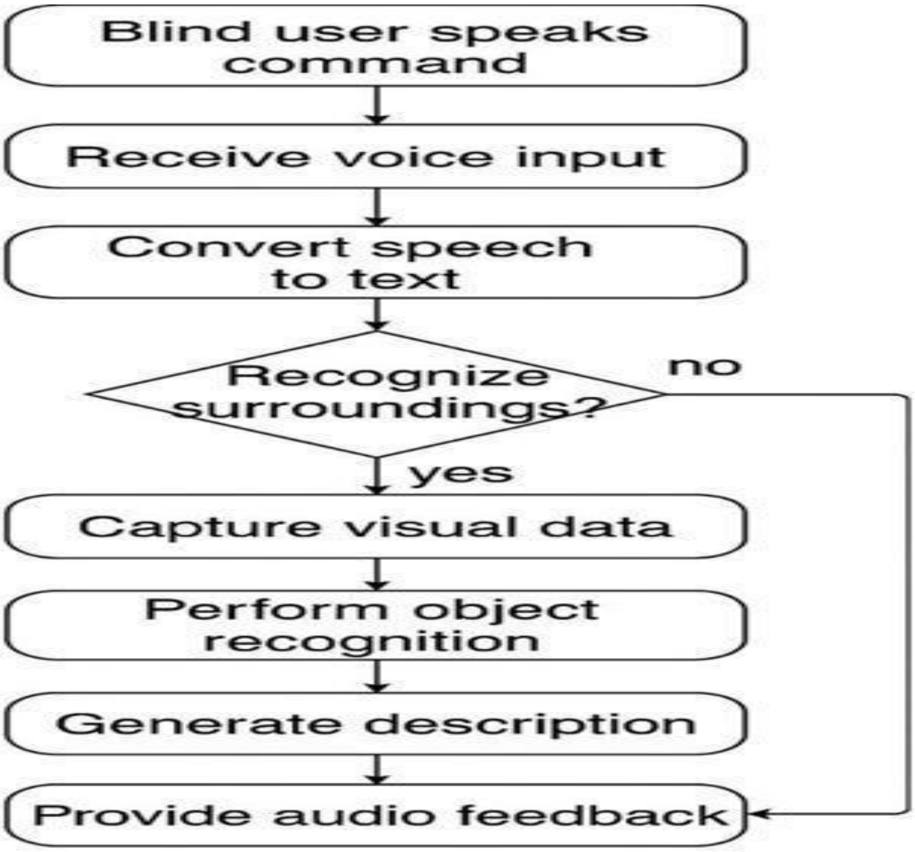

Fig1.6: Activity Diagram

VIII.ENTITY RELATIONSHIP DIAGRAM

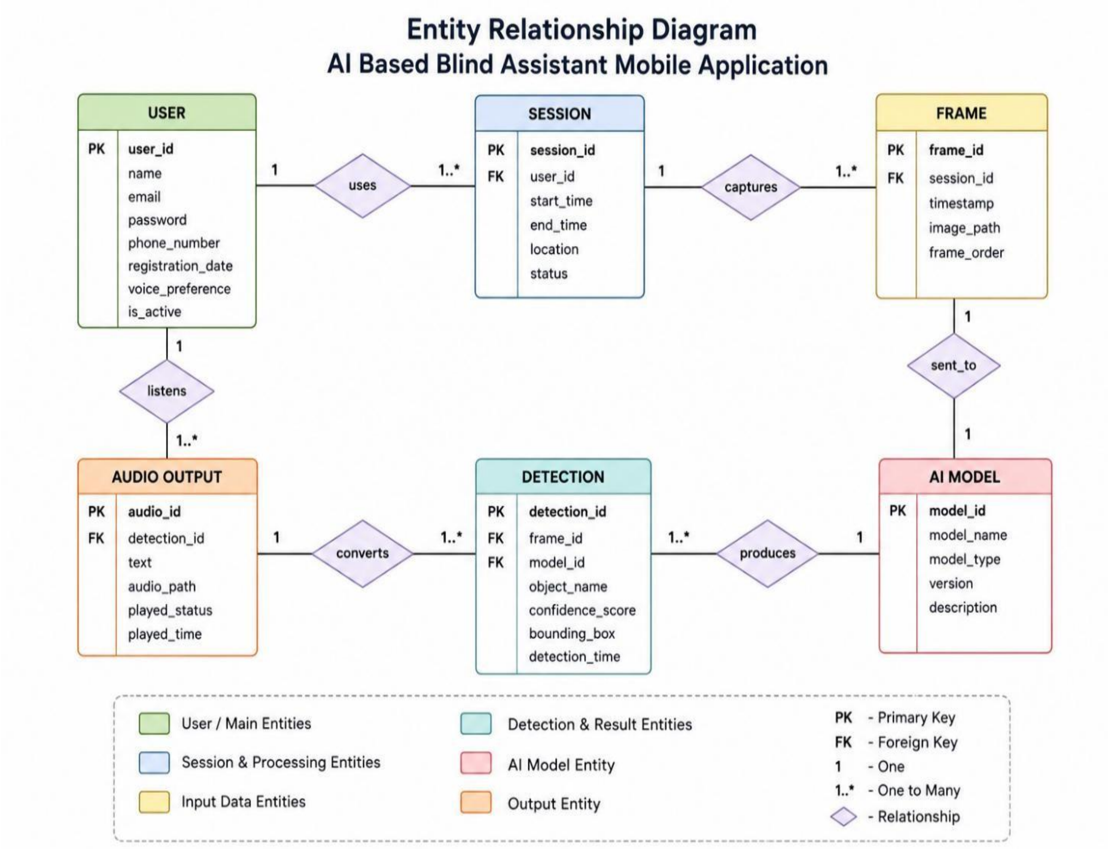

Fig1.7: Entity Diagram

- **User**

Visually impaired person who interacts with the system using voice commands.

- **Voice Input Module**

Converts user speech into text commands using speech recognition.

- **Camera Module**

Captures real-time images or video from the surroundings.

- **AI Processing Unit**

Processes voice commands and analyzes captured images.

- **AI Model**

Detects objects and recognizes surroundings using machine learning.

- **Object Detection Module**

Identifies objects, obstacles, and important items around the user.

- **Audio Output Module**

Converts system responses into speech for the user.

- **Database/Storage**

Stores user data, detected results, and system information.

- **Mobile Application**

Provides the interface for system control and communication.

IX.RESULTS

Fig 1.8 Read Text Page

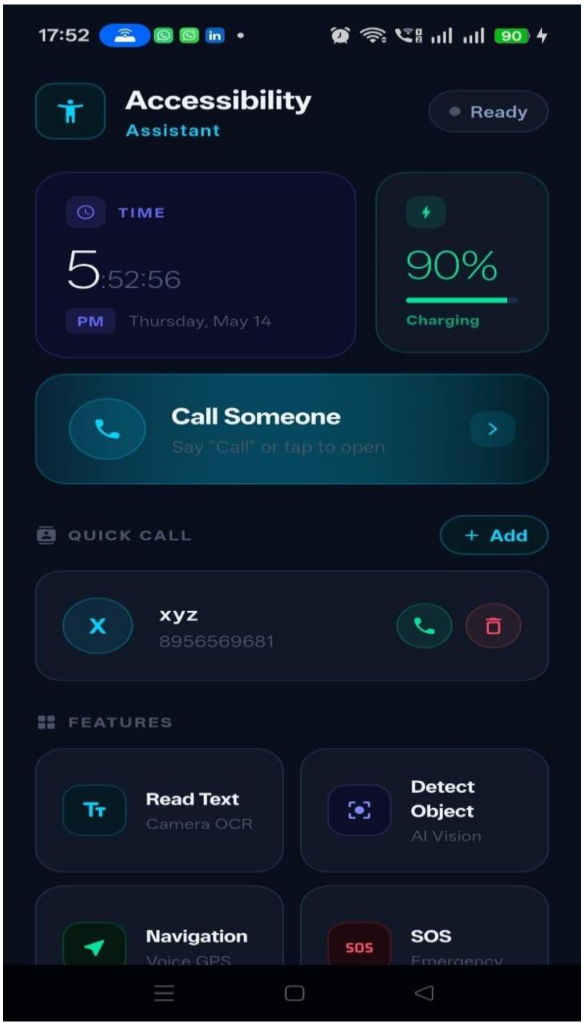

Fig 1.9. Home Page

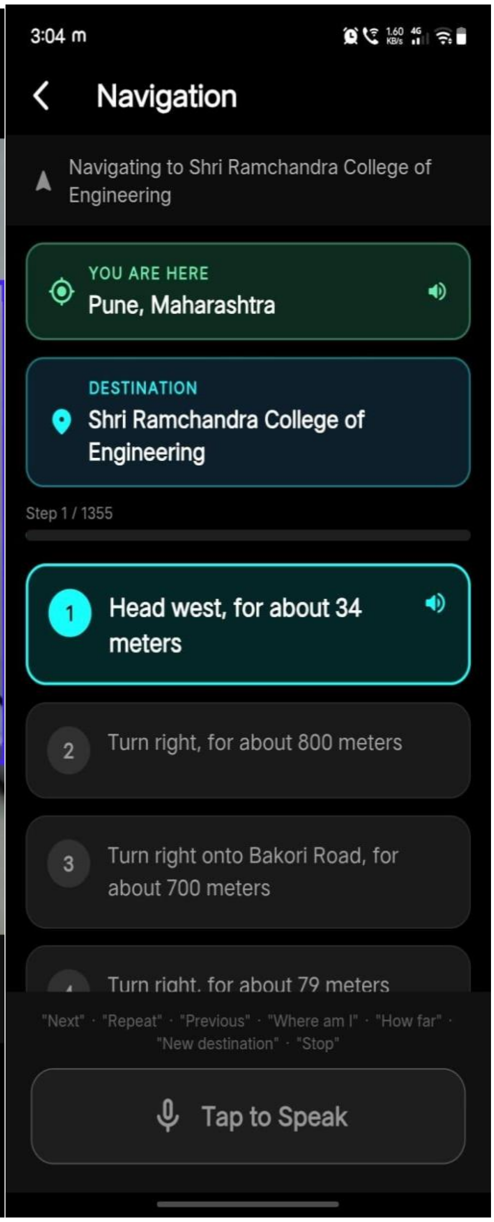

Fig2.0: Object Detection Page Fig2.1: Navigation Page

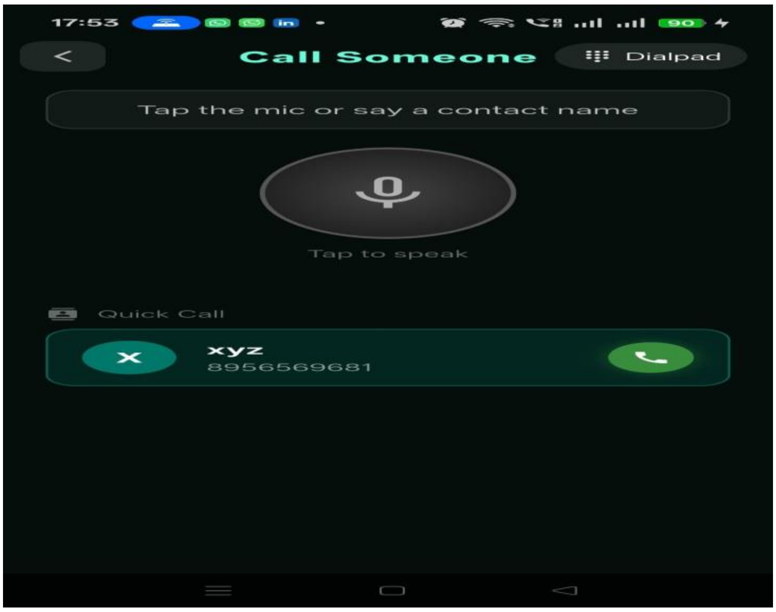

Fig2.1: Speak and Call Page

X.ADVANTAGES

- **Real-Time Object Detection:** The system provides real-time object detection using artificial intelligence and computer vision techniques. It continuously scans the surroundings and quickly identifies nearby objects, helping visually impaired users understand their environment more effectively.
- **Instant Audio Feedback Using TTS:** The application uses Text-to-Speech (TTS) technology to provide immediate voice feedback about detected objects. This enables users to receive quick audio instructions without needing to see the screen, making the system more accessible and user-friendly.
- **Safe Navigation Assistance:** The system helps visually impaired users move safely by detecting obstacles, pathways, and surrounding objects in real time. This reduces the risk of accidents and improves mobility in both indoor and outdoor environments.
- **Increases Independence:** By offering automated guidance and object recognition, the system reduces the dependency of visually impaired individuals on others for navigation and daily activities. Users can travel and perform tasks with greater confidenc e and independence.
- **Easy-to-Use Interface:** The application is designed with a simple and user-friendly interface so that users can operate it easily without technical knowledge. The voice-based interaction further improves accessibility and convenience.

XI.DISADVANTAGES

- **Performance in Low-Light Conditions:** The system's performance may decrease in low-light or dark environments because the camera may not capture objects clearly, reducing detection accuracy.
- **Reduced Accuracy in Crowded Environments:** In crowded or complex surroundings, the system may face difficulty identifying multiple objects simultaneously, which can lead to incorrect or delayed detection results.
- **High Processing Power and Battery Usage:** Continuous camera operation, AI processing, and audio feedback consume significant processing power and battery life, which may reduce the overall performance of the mobile device.
- **Dependence on Camera Quality:** The accuracy of object detection depends heavily on the quality of the smartphone camera. Devices with low-resolution cameras may produce less accurate results.
- **Incomplete Object Detection:** The system may not always detect all objects correctly, especially very small, transparent, or fast-moving objects, which can affect navigation safety.

XII.APPLICATION

- **Daily Navigation Assistance:** The system helps visually impaired users in their daily navigation activities such as walking on roads, avoiding obstacles, crossing pathways, and moving independently. Real-time object detection and audio guidance improve safety and confidence during travel.
- **Indoor Assistance:** The application can be used inside homes, offices, schools, and buildings to identify objects such as chairs, doors, tables, stairs, and walls. This helps users move safely and understand indoor surroundings more effectively.
- **Outdoor Navigation Support:** The system supports outdoor navigation by detecting vehicles, pedestrians, traffic signals, pathways, and obstacles in public places. This improves user safety while traveling in crowded or unfamiliar environment
- **Object Recognition for Daily Tasks:** The application assists users in recognizing nearby objects required for daily activities, such as finding personal belongings, identifying household items, and understanding surrounding objects easily through audio feedback.
- **Educational Environment Assistance:** The system can be used in schools, colleges, and learning environments to help visually impaired students identify classrooms, furniture, and learning spaces. It supports independent movement and improves accessibility in educational institutions.
- **Emergency and Safety Support:** In emergency situations, the application can quickly detect obstacles, dangerous objects, or unexpected hazards and provide instant audio alerts. This helps users react safely, avoid accidents, and improve overall personal safety.

XIII. LIMITATIONS

- **Dependence on Device Hardware:** The performance of the system highly depends on the smartphone's hardware components such as camera quality, processor speed, RAM, and microphone. Devices with lower specifications may experience slower processing speed and reduced object detection accuracy.
- **High Battery Consumption:** Continuous use of the camera, machine learning model processing, and voice feedback functionality consumes a large amount of battery power. Prolonged usage of the application may quickly drain the smartphone battery.
- **Reduced Accuracy in Complex Environments:** In crowded, cluttered, or highly dynamic environments, the system may face difficulties in detecting and identifying multiple objects correctly. This can affect the reliability of real-time assistance provided to the user.
- **Low-Light Performance Issues:** The accuracy of object detection decreases in dark or poorly lit environments because the camera captures low-quality images. Poor lighting conditions can reduce the effectiveness of the AI model.

XIV. CONCLUSION

The Blind Assistance Mobile Application is a useful AI-based solution designed to support visually impaired users in their daily lives. By combining object detection, camera processing, and voice assistance, the application helps users understand their surroundings safely and independently. It uses real-time image processing to identify objects and converts this information into clear audio feedback through text-to-speech technology. This enables users to navigate both indoor and outdoor environments with greater confidence and reduced dependency on others. The application is user-friendly, portable, and accessible through smartphones, making it a practical solution for everyday use. Despite some limitations, it significantly improves safety, awareness, and mobility. Overall, the project demonstrates the effective use of Artificial Intelligence and mobile technology for social good and inclusive development.

XV.FUTURE SCOPE

Improve accuracy and speed of object detection using advanced AI models. Add face recognition to identify known people. Integrate GPS navigation for better outdoor guidance. Support multiple languages for wider usability. Enhance performance in low light and complex environments. Add features like text reading (OCR) for reading signs and documents. Connect with smart devices (IoT) for a more interactive and intelligent system.

REFERENCES

1. K. Dhara Vath, V. Tejaswi, E. S. P. Kumar, and V. S. Pratham, “Deep Learning based Intelligent Alert System for Visually Impaired People,” 2023 International Conference on Computer Communication and Informatics (ICCCI), Coimbatore, India, 2023, pp. 1-6, Doi: 10.1109/ICCCI56745.2023.10128277.
2. Xuhui Hu, Aiguo Song, Zaikai Wei and Hong Zeng, “Stereo Pilot: A Wearable Target Location System for Blind and Visually Impaired Using Spatial Audio Rendering,” in IEEE Transactions on Neural Systems and Rehabilitation Engineering, vol. 30, pp. 1621-1630, 2022, Doi: 10.1109/TNSRE.2022.3182661.
3. Aicha Gourbi, Haythem Ghazouani, and Walid Barhoumi, “Driver Drowsiness Detection Based on Joint Monitoring of Yawning, Blinking and Nodding,” in 2020 IEEE 16th International Conference on Intelligent Computer Communication and Processing (ICCP), 2020, DOI: 10.1109/ICCP51029.2020.9266160.
4. Jalendra Singh, “Deep Learning Based Driver Drowsiness Detection Model,” in 2020 3rd International Conference on Intelligent Sustainable Systems (ICISS), India, 2020, DOI: 10.1109/ICISS49785.2020.9316131.
5. Ajzen Joshi, Survi Kyal, Sandipan Banerjee, Taniya Mishra, “In-the-wild Drowsiness Detection from Facial Expressions,” Paper from HSIM Workshop at IEEE Intelligent Vehicles Symposium 2020 (IV2020), https://doi.org/10.48550/arXiv.2010.11162.
6. P. Murugesapillai, Rd. Manimegalai, “Complex Background and Foreground Extraction in Colour Document Images using Interval Type-2 Fuzzy,” International Journal of Computer Applications Foundation of Computer Science (FCS), NY, USA, Volume 25 – Number 2, 2011, Doi: 10.5120/3006-4054.
7. Neelam Labhade-Kumar, Mangala S. Biradar, Ashvini Narayan Pawale, “Reinforcement Learning-Based Deep FEFM for Blockchain Consensus Mechanism Optimization with Non-Linear Analysis,” Journal of Computational Analysis and Applications, vol. 33, no. 05, 2024.
8. Neelam Labhade-Kumar, “Shot Boundary Detection Using Artificial Neural Network,” Advances in Signal and Data Processing, Lecture Notes in Electrical Engineering, Springer, vol. 703, pp. 44–55, Jan. 2021.
9. Neelam Labhade-Kumar, “Optimizing Cluster Head Selection in Wireless Sensor Networks Using Mathematical Modeling and Statistical Analysis of the Hybrid Energy-Efficient Distributed (HEED) Algorithm,” Communications on Applied Nonlinear Analysis, ISSN: 1074-133X, vol. 31, no. 6s, pp. 602–617, Aug. 2024.
10. Neelam Labhade-Kumar, “Experimental Design of Electricity Theft Detection and Alert System Using Arduino Assisted Controller and Smart Sensors,” 7th International Conference on Inventive Computation Technologies (ICICT), IEEE, 2024, pp. 1961–1968.
11. Neelam Labhade-Kumar, “Novel Management Trends Using IoT in Indian Automotive Spares Manufacturing Industries,” Journal of Pharmaceutical Negative Results, vol. 13, no. 09, pp. 4887–4899, Nov. 2022.
12. Neelam Labhade-Kumar, “Adaptive Hybrid Bird Swarm Optimization Based Efficient Transmission in WSN,” Journal of Pharmaceutical Negative Results, vol. 14, no. 02, pp. 480–484, Jan. 2023.
13. Neelam Labhade-Kumar, “Combining Hand-crafted Features and Deep Learning for Automatic Classification of Lung Cancer on CT Scans,” Journal of Artificial Intelligence and Technology, 2023.
14. Neelam Labhade-Kumar, “Enhancing Crop Yield Prediction in Precision Agriculture through Sustainable Big Data Analytics and Deep Learning Techniques,” Carpathian Journal of Food Science and Technology, Special Issue, pp. 1–18, 2023.
15. Neelam Labhade-Kumar, “Accident Prevention and Management System in Urban VANET for Improving Slippery Roads Ride After Rain,” Journal of Environmental Protection and Ecology, ISSN: 1311-5065, vol. 25, no. 2, pp. 586–599, 2024.
16. Neelam Labhade-Kumar, “An Image Processing Method for Kidney Stone Segmentation in CT Scan Images Based on CNN-Regularized Extreme Learning Machine Approach,” Hybrid and Advanced Technologies, 2023.