International Journal of Advances in Engineering and Management (IJAEM)
Volume 7, Issue 05 May 2025, pp: 165-172 www.ijaem.net ISSN: 2395-5252

Advanced Intelligent Navigation System: Redefining Accessibility and Mobility for Visually Impaired Individual

[1]Dr. B. Srinivasa Rao, [2]Bobbili Disshaa, [3]B. Sahithi, [4]Alampally Ramya

Teegala Krishna Reddy Engineering College, Hyderabad, Telangana.
Teegala Krishna Reddy Engineering College, Hyderabad, Telangana.
Teegala Krishna Reddy Engineering College, Hyderabad, Telangana.
Teegala Krishna Reddy Engineering College, Hyderabad, Telangana.

Date of Submission: 01-05-2025 Date of Acceptance: 10-05-2025

ABSTRACT

There are many people suffering from vision impairment or related issues. Although there are many assistive technologies to help these individuals, like Guide Dogs, White Canes, GPS enabled applications, etc. but they have few limitations like they are not flexible and are not helpful in real-time navigation. To overcome these drawbacks and help these individuals in their day-to-day life the Advanced Intelligent Navigation System, a smartphone application is designed. The application, developed with Flutter, combines several AI-powered features into an easy-to-use mobile application. The features of this app include object detection and crowd detection using Google API, nearby places discovery and real-time navigation assistance using Google Places and Maps API, using Gemini Vision API for reading text or image and the accelerometer sensors are monitored for detecting fall and alerts the emergency contacts on detecting fall. This application also supports multilingual voice assistance using Gemini API to identify intent and provide the required audio and haptic feedback to the user.

I. INTRODUCTION

1.1 Motivation

Individuals with visual impairments suffer a lot of problems in their daily life and there isn't much assistance available for them. Even though there are a few existing systems to assist them, they mostly rely on external hardware or are not accessible for real-time situations which makes these systems unfit for usage.

White Canes are still the most common aid. They are easy to access, in-expensive and reliable but are not really helpful as they only sense objects that are in close proximity and it needs direct contact with the objects. Depth estimation for dynamic obstacles like crowd, vehicle or even overhead or fast moving obstacles is not possible.

Guide Dogs are used in spatial navigation but they don't give detailed information about their surroundings. They are highly expensive and require continuous maintenance and training. And they are not easily available in low income or rural areas.

Alternatively, some electronic aids employ ultrasonic or infrared sensors to identify obstacles via electronic wearable technologies. They are expensive, dependent on hardware devices and made only for specific tasks. They also struggle to combine navigation, awareness of crowds or support for multiple languages.

Computer vision systems powered by mobile cameras and AI are becoming more popular, especially in object recognition and scene interpretation. But many of these systems are still under development, often require a lot of hardware and lack a user-friendly interface for those who are technically handicapped. They support only one language and don't work well in low-light settings or busy scenarios.

In these solutions only one specific problem like spotting obstacles, finding routes or reading text are addressed but they don't really come together to offer a seamless experience. There's definitely a growing need for more adaptive, context aware tools that can work in real time, adjust to different indoor and outdoor settings and serve users from various languages and region.

1.2 Objective

In essence, the project is concerned with building a comprehensive AI based navigation system to allow visually impaired people to be independent and safe during navigation in a contradiction of environments. This will incorporate multiple assistive features into a fully functional yet user friendly mobile app for smartphones which will not require any expensive or special device.

The project focuses on the use of real time computer vision and location-based technologies to resolve the disjointed nature of existing tools. The main features include detecting objects, crowd detection, fall detection, GPS navigation, nearby place discovery, text or image description and multilingual voice assistance for the users to have situational awareness and decision making.

This module is capable of detecting objects like walls, furniture, various vehicles and even recognizes people. The crowd detection may help deter users from walking in congested places. A sudden change in a gyroscopic orientation is detected by the fall detection system, and should such a change happen while wearing the device, the system immediately alerts caregivers of a fall.

However, the GPS navigation gives directions for traveling outside and also has a near-by places feature that connects the user to nearby services such as hospitals, shops, among others. Text recognition or image summariser makes it possible for the user to read signs, menus, describe images and other printed matters using visual or spatial means. This information gets transformed into audible narration feedback. Using Gemini API, a Voice Assistant that provides in-app functionalities in your personal language: better support for the non-English user created with Flutter framework, which provides cross-platform compatibility as well as seamless voice, vision, and motion integration.

In simple terms, the project aims to design a context-aware solution that helps the user's to understand their environments and act proactively using auditory output to promote safe, confident, and independent navigation of real-world environments by blind users.

II. LITERATURE SURVEY

Navigation aids for blind people have transitioned from mechanical navigation aids such as canes to smart navigation systems powered by computer vision, machine learning, artificial intelligence and mobile computing. However, despite many advancements for navigation aids for blind people there is still a lack of affordable, adaptable, and integrated solutions.

1. Asadpour et al. (2019). “Navigation Systems for the Blind Based on Artificial Intelligence: A Survey” – Journal of Rehabilitation Research and Development.

This survey looked at wearable and mobile navigation aids that can be designed using AI. The majority of the wearable systems provide very good detection, but are accompanied by the need for unique hardware or external processing units to obtain good accuracy, which raises the cost of the system and limits portability.

2. Tapu et al. (2020). “A Smartphone-based Obstacle Detection and Classification System for Visually Impaired Users” – Sensors Journal.

The study developed a smartphone system for classification based on semantic segmentation and edge detection. However, it did not excel in real-time co-crowded environments; did not offer support for more than 1 language; and did not have any voice interactions, thereby restricting usability to various contexts.

3. Mahmud et al. (2021). “Voice-Controlled Navigation Assistance Using Deep Learning on Mobile Platforms” – International Conference on Intelligent Systems.

They developed a voice controlled obstacle detection android app leveraging deep learning technology. Their work was certainly an interesting concept; however, they did not utilize GPS, depth sensing, and could not distinguish complex structures, such as stairs or if individuals were moving together in crowds.

4. Kim and Cho (2022). “Indoor and Outdoor Integrated Navigation for the Blind Using Wi-Fi, BLE, and GPS” – IEEE Access.

This hybrid system utilized a number of different location technologies to provide integrated navigation. The accuracy was superb, yet it needed a very particular pre-installed infrastructure (i.e. a location of Wi-Fi access points or BLE beacons) so it is not something which could be used in general terms.

5. Wang et al. (2023). “AI-Powered Fall Detection and Alert Systems for Elderly and Visually Impaired Populations” – ACM Transactions on Accessible Computing.

This study investigated safety monitoring with gyroscope-based fall detection, but while it scanned for dangers, and sent alerts to safety personnel, it was treated as a standalone service and did not explore support for navigation or object recognition, which admittedly, reduces its scope.

Together, the studies highlight the limitations of existing technologies: expensive, faulty, fragmented technologies that rely on either special hardware or installed infrastructure and a lack of systems that can engage with the various complexities of real world navigation. The widespread accessibility of smartphones opens up a window to optimize and bundle these apparently disparate features into a single and intelligent infrastructure that supports real-time, voice-assisted, multilingual communications in dynamic environments.

III. PROPOSED SYSTEM

The Advanced Intelligent Navigation System, is a mobile based assistive technology designed to allow visually impaired users to travel safely and increase spatial awareness. Developed using the Flutter framework, the system consists of multiple modules each addressing a unique aspects of navigation and real-world accessibility. These modules include: speech command recognition, GPS-based navigation, nearby places discovery, object and crowd detection, text reading or image summarisation, and emergency fall notifications. The system uses internet connectivity for performing most of the tasks. The voice assistant gets activated on tapping. The coordination of the various modules within the system allows for the system to provide communication to the user through audio and vibration in order to make quick decision processes in real-tone when navigating from one space to another.

It makes use of services such as Google Gemini API for intent recognition, Gemini Vision API to read text or summarise images and Google ML Kit to provide vision functionality. It monitors the accelerometer to detect a fall and alerts emergency contacts. Feedback to the user is provided in the form of speech or haptic vibration.

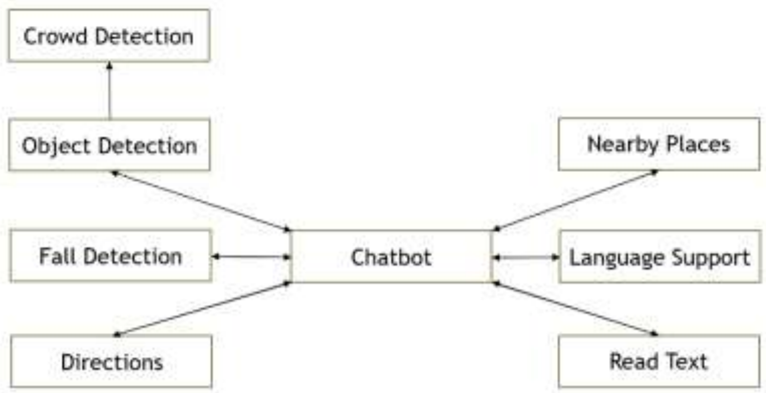

Fig 3.1: System Architecture of Proposed System

Voice Interaction Module:

The Voice Interaction Module is the interface to the system that allows the user to communicate with the system using their choice of language. Once the user taps the screen, it activates the Voice Interaction Module which begins to record the user's voice. The Flutter Speech-to-Text is used for converting user commands into text in real-time. When the spoken word is translated into text, it is sent to the Gemini API where the AI processing interprets what was said, and what the user wants to do with that command, as well as the referenced command modules (typically navigation, object detection, or reading text) that will be triggered and activated appropriately. Voice Interaction produces an audio output in the form of speech.

Navigation Module:

In the Navigation Module, the direction to the requested destination are given. After the user gives a command to the system for getting directions, the app will run a command through the Google Maps API to find a route. For example, the user may say, "Take me to XYZ Hospital." This voice command is translated to text using the Voice Interaction Module. The destination is sent to Google Maps, and the system will calculate the best route to use depending on live GPS and road conditions. The system will then give the user turn-by-turn route through Text-to-Speech (TTS), to prompt the user each turn, with distance and any other navigation details. The system monitors the route, and recalculate the directions as the user moves. The module can also use different

International Journal of Advances in Engineering and Management (IJAEM)
Volume 7, Issue 05 May 2025, pp: 165-172 www.ijaem.net ISSN: 2395-5252

languages so that directions will speak in the users's desirable language.

Nearby Places Module:

The Nearby Places Module allows users to find nearby services and facilities such as hospitals, shops, restaurants, and other necessary places. Once the user would ask for a places information for example, "Find a pharmacy near me", or "Where is the closest hospital?" the system will use the Google Places API to pull the relevant information based on the user's GPS coordinates. After the system gets the data for nearby places, the system will read out the 5 nearest places. The user receives this information through Text-to-Speech and the user can decide where to go next.

Object Detection Module:

As a key component of detecting obstacles in the user's environment, the Object Detection Module uses the phone's camera to actively capture a live feed of the user's environment. This live-feed is processed with the Google ML Kit's Object Detection API and OpenCV. OpenCV improves the live feed in terms of quality and object detection. The Google ML Kit is used to recognise objects like people, cars and furniture etc. It also uses ARCore Depth API to identify the distance between the user and the object detected. The depth API helps give more simulation of what's close or actually there so the system can provide timely help. When an object is detected, the application provides audio feedback and haptic feedback that will help guide the user. This module combines three different levels of information, object detection, image enhancement and depth information, to help provide accurate real-time feedback in order to make sure users navigate safely.

Crowd Detection Module:

The Crowd detection module, informs a user they are near a crowd. The method used is simple object detection, by performing analysis on the user's live camera feed and counting the people that are present. When there is a number of people detected in close proximity, it will output a sound alert and state, "Crowd detected, please proceed carefully." This feedback provides a chance for the user to make alternative choices if crowded or unsafe. The module was developed based on the ability of detecting people, as well as density, using Google ML Kit. Mostly, the system is meant to provide audio output, warning the user to cautiously proceed through areas dense with high numbers of individuals.

Fall Detection Module:

The Fall Detection Module recognizes when a user falls, using the accelerometer in the mobile device. The module considers the angle of the device, and how the device is moved, calculates if there was an uncontrolled sudden movement when a fall is detected. The moment the module identifies a fall, the system will enable a notification such as vibration and audio alert. The system will also automatically send an emergency short message service (SMS) to emergency contacts in addition to the GPS coordinates indicating the user's location, as set up by the user. This module system and alerts provide rapid response to accidents, and critical assistance during opportunities of great need.

Text Reader or Image Summariser Module:

The Text Reader or Image Summariser Module is intended to allow users to read printed text or summarise images in their environment. The system uses phone's cameras to capture an image of the text or image. This captured image is sent to the Gemini Vision API which applies optical character recognition (OCR) to extract the text from the image or describes that image. The extracted text or the description of the image is converted to Text-to-Speech, and the user can hear the output. The module also includes functionality for explaining the content of pictures, like a description of the content of a timetable or a menu. The system may also translate whatever text is extracted into the user's preferred language, using the Gemini API.

Multilingual Support Module:

The Multilingual Support Module provides multiple languages support for users who may speak in a different language. When a user sets up their preference for the system's language, the system will determine the user's desired setting using Gemini API along with Text-to-Speech capabilities so that spoken feedback is made available in the appropriate language. For example, if a user selected Hindi as their preferred language, then the system would respond in that same language reading all navigation instructions, with voice, such as nearby places, among other indications. The module will also facilitate translation of voice input from and to spoken language, so that communication across languages remains smooth and uninterrupted among users from diverse backgrounds.

Feedback Module:

The feedback module provides audio notifications through Text-to-Speech and tactile

International Journal of Advances in Engineering and Management (IJAEM)
Volume 7, Issue 05 May 2025, pp: 165-172 www.ijaem.net ISSN: 2395-5252

notifications through haptic notifications. For example, the system would provide audio feedback if an obstacle is detected in the walkway, or a fall is detected, that informs the user that an "Obstacle ahead" or "Fall detected." Furthermore, it vibrates the phone to provide the user with immediate feedback in situations where an audio notification would be difficult to hear (noisy environments). By using multi-sensory feedback, users can receive critical information regardless of their environment.

In the end, the Advanced Intelligent Navigation System employs smartphone sensors, and modular architecture to provide an available, real-time assistant for users with visual impairment. Its minimalist technology stack of robust and proven technologies such as Google ML Kit, Gemini API, and GPS services provide reliability and confidence to the user experience. Also, it is designed to improve user's spatial orientation through obtaining users preference of regional/native language and emergency contacts. The utility of the system in the real world with many of the assistive features integrated will provide the user's with intentional support and an improved ability to navigate in environments that can be complex.

IV. METHODOLOGY

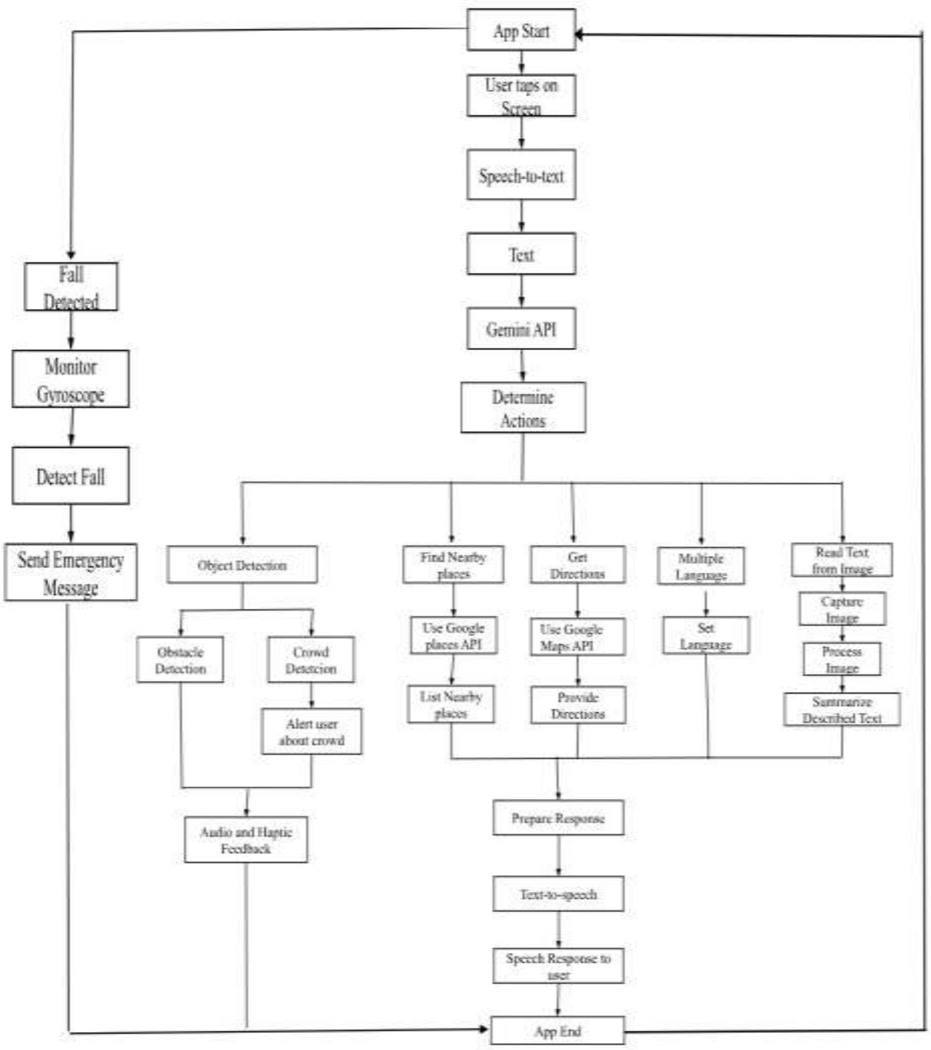

Fig 4.1: Flow Chart

STEP-1: On opening the app for the first time, the permissions screen appears and will be asked to set an emergency contact.
STEP-1.1: The app will alert the emergency contact with an SMS along with an audio feedback to the user whenever a fall is detected using mobile's accelerometer.
STEP-1.2: Tap-to-speak activates the speech recognition and takes the input in user's selected language.
STEP-2: The users voice command is passed to the Flutter speech-to-text engine, which converts the spoken command to text.
STEP-3: The text is sent to the Gemini API for intent detection, and based on which intent is identified, the app will perform one of the following actions:
STEP-3.1: Object Detection: Identifies obstacles and crowds using Google a ML Kit. If a crowd is detected app alerts the user's emergency contact and gives and audio feedback to the user.
STEP-3.2: Find Nearby Places: Uses Google Places API to provide the list of nearby places.
STEP-3.3: Get Directions: Google Maps API is used for step-by-step navigation.
STEP-3.4: Multilingual Support: Facilitates the users ability to select and set up a language preference.
STEP-3.5: Read Text Describe Image: Activates Gemini's Vision API to capture and process through their API and reads or summarizes the text or image based on the user's request.
STEP-4: For modules like nearby places, get directions, read or summarise images and multilingual support, the app uses Text-to-speech (TTS) to convert the responses into speech of user's selected language and give audio and haptic feedback.

V. IMPLEMENTATION

*******Hardware Information**

This system functions on devices with Android 10 or later and iOS 16 or later versions. It requires an Android device that supports ARCore from Google or an iPhone that supports ARKit from Apple. The rear camera of the device must have an 8 MP resolution or higher with autofocus in order to perform object detection with reliability and to use the depth information with accurate results. A multi-core CPU that runs at a clock speed of 2 GHz or higher, with at least 4 GB of RAM, will ensure that the object detection, depth estimation and navigation all occur without any delays. The devices must possess an accelerometer for incident detection, a GPS module for real-time tracking of the location of the user, and have a functioning microphone and speaker for the user to provide voice commands and audio feedback. Together all of this hardware will allow the system to provide effective real-time navigation and object detection to help users with visual impairments.

Software Information

The Advanced Intelligent Navigation System is developed with Flutter, a cross-platform framework that minimizes development time, provides a single codebase to deploy on Android and iOS devices, and gives the same features on each platform. The system activates speech recognition upon tapping the screen and input is taken in the form of speech. The voice assistant allows interaction in Telugu, Hindi, and English. The internet connectivity is essential for processing.

Google ML Kit is used for object detection; it uses camera frames in real time, recognizing and identifying images in each frame based on models that machine learning has trained for the system to use. Google Google ARCore on Android and Apple using ARKit to generate a depth map by capturing camera images from two camera frames, in which the application uses the disparity between camera frames to give an estimated distance to nearby objects. To navigate, the app uses the Google Maps API to get the user's real time location data and to provide turn-by-turn instructions while the Google Places API receives the user's location and subsequently identifies restaurants, hospitals, shops, etc. near the user's location.

Text reading or image summarisation is done with the Gemini Vision API, which takes images captured by the camera, extracts the text or analyses the images and produces speech output. Intent identification is handled by the Gemini API, which takes user speech or input to identify what the user wants to do, such as starting navigation or reading the text aloud. Crowd detection is a straightforward rule-based feature, when the camera sees multiple people in the field, using ML Kit's object detection, the system would warn the user with the message "Crowd detected, please proceed carefully". Fall detection monitors mobile's accelerometer, on sensing abnormal movement patterns it considers as a fall and notifies the user's emergency contacts.

International Journal of Advances in Engineering and Management (IJAEM)
Volume 7, Issue 05 May 2025, pp: 165-172 www.ijaem.net ISSN: 2395-5252

VI. RESULT

Fig 6.1:

Fig 6.2

Fig 6.3

Fig 6.4

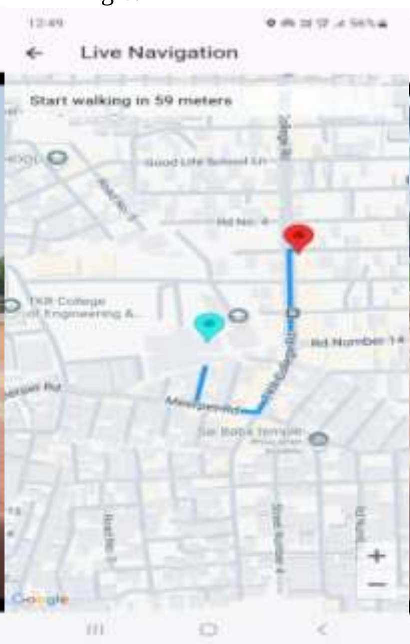

Fig 6.5

Fig 6.6

We tested on Android 10+ and iOS 16+. During testing the Advanced Intelligent Navigation System worked effectively in busy environments, it also alerted the user whenever the connectivity was weak. Object Detection used Google ML Kit and identified common obstacles like vehicles, doors, people and also crowds within realistic environments with sufficient lighting. ARCore on Android and ARKit on iOS provided expected distance measures for close objects, thereafter allowing timely audio and haptic feedback to the user. On identifying many people in the feed it alerted users stating "Crowd detected proceed carefully" which helps users to travel safely in crowded areas.

Along with the audio and haptic feedback the app also provided input and output visuals on the screen for every command and also gives the count-down before capturing images. The user commands and output responses also performed well in all the specified languages.

The touch activation feature was consistent and when Gemini API was implemented it processed user commands accurately. The Gemini Vision API properly recognises printed text and describes the images. Fall detection module accurately detects fall with very few false alarms and alerts the contacts quickly along with the location of the user. Google Maps and Places API worked properly and provided accurate navigation and local information.

VII. FUTURE SCOPE

Few features that can be added in the advanced versions of Advanced Intelligent Navigation System are adding offline functionality to the functions like real-time navigation, searching nearby places and reading text or images so that it improves the user's experience in areas where connectivity is weak, also adding many other languages so that the app can be accessible in any region. We can also possibly add features such as context-aware navigation which could help in identifying traffic signals, crosswalks, or even indoor environments increasing its usability and versatility. Or suggestive routes based on artificial intelligence and real-time crowd analysis that provides the user with information about under-construction or busy roads to ensure that they could proceed safely. In addition, the application can also initiate smart city infrastructure, such as traffic signals or public transport information to enable people with blindness or vision loss, to navigate independently.

VIII. CONCLUSION

The Advanced Intelligent Navigation System is developed on Flutter, combining with the present advanced technologies to enable the people with visual difficulties to navigate in their surroundings easily with more independence and safety, locate nearby places, know about the density of the crowd, converse and find assistance in emergency situations. It uses technologies like Google ML Kit and ARCore for identifying objects and estimating their distances, Gemini API for identifying intent and performing the action required, Places and Maps API to search for nearby places and travel through their surroundings.

It has a user-friendly interface i.e., the user can tap on the screen and voice assistant activates to listen to the user's commands and respond accordingly. The user can communicate with the app in natural language and the app supports multiple regional languages. It also provides realtime turn by turn navigation to users with voice feedback for directions. Other features include fall detection and alerting contacts on fall, crowd detection to enable the user to navigate safely. It uses a cross platform development framework Flutter to support both Android and iOS thus extending usability.

The application demonstrates how mobile technology can be a game-changer for many ease-of-use tools for the visually challenged and offers better navigation for independent and safe mobility throughout the user's premises.

REFERENCES

[1]. Anderson, R., & Green, S. (2019). Augmented Reality in Navigation Systems for the Visually Impaired. Journal of Assistive Technology, 10(3), 302-315.

[2]. Asadpour et al. (2019). “Navigation Systems for the Blind Based on Artificial Intelligence: A Survey” – Journal of Rehabilitation Research and Development.

[3]. Zhang, Z., & Jafari, R. (2020). Real-Time Object Detection in Mobile Applications. Journal of Computer Vision and Applications, 48(5), 457-467.

[4]. Kurniawan, S., & Whittaker, S. (2020). Assistive Technologies for Navigation: A Review. Journal of Disability Studies, 12(2), 120-135.

[5]. Tapu et al. (2020). “A Smartphone-based Obstacle Detection and Classification System for Visually Impaired Users” – Sensors Journal.

[6]. Smith, J., & Hutton, D. (2021). Depth Sensing for Navigation in Augmented Reality. IEEE Transactions on Mobile Computing, 20(1), 110-121.

[7]. Mahmud et al. (2021). “Voice-Controlled Navigation Assistance Using Deep Learning on Mobile Platforms” – International Conference on Intelligent Systems.

[8]. Kim and Cho (2022). “Indoor and Outdoor Integrated Navigation for the Blind Using Wi-Fi, BLE, and GPS” – IEEE Access.

[9]. Wang et al. (2023). “AI-Powered Fall Detection and Alert Systems for Elderly and Visually Impaired Populations” – ACM Transactions on Accessible Computing.

[10]. Dr. B. Srinivasa Rao, Bobbili Disshaa, B Sahithi, & Alampally Ramya (2024). Intelligent Navigation App for the Visually Impaired: A Smartphone based Assistive Technology. Journal of Emerging Technologies and Innovative Research.