

DRISHTI: AI-Powered Assistive Companion System for Visually Impaired Individuals for Daily Life Tasks

¹Shristi Tripathi, ²Dr. Pawan Kumar Pandey, ³Mudit Dubey

¹Digvijay Nath PG College, Gorakhpur, ^{2,3}Assistant Professor, Digvijay Nath PG College, Gorakhpur

Abstract:

This paper presents the design and implementation of Drishti, an AI-powered assistive mobile application developed to enhance the mobility and independence of visually impaired individuals. Traditional assistive methods, such as white canes, provide limited environmental awareness and lack intelligent interaction capabilities, making navigation and daily activities challenging.

To address these limitations, the proposed system is developed using the Flutter framework for cross-platform mobile application development and integrated with advanced technologies, including TensorFlow Lite for real-time object detection and speech processing, enabling voice-based interaction. Additionally, the system uses the Gemini API to deliver contextual, intelligent responses based on user input. The application uses a voice-driven interface, allowing users to interact with the system without relying on visual elements.

The system provides key functionalities including environment detection, text recognition, and obstacle-aware navigation assistance. By processing real-time camera input and delivering audio feedback, Drishti enhances situational awareness and supports safer navigation. The implementation results demonstrate improved accessibility, reliable performance, and user convenience under real-world conditions. The system achieves an object detection accuracy of approximately 80–85% and voice recognition accuracy of 85–90%, making it a practical and effective solution compared to traditional assistive approaches.

Keywords:

Assistive Technology, Artificial Intelligence, Computer Vision, Flutter, Object Detection, Voice Interaction, Accessibility, Navigation System, TensorFlow Lite, Mobile Application

1. INTRODUCTION

In today's digital era, advancements in artificial intelligence and mobile technologies are transforming the development of assistive systems for individuals with disabilities [6][7]. Visually impaired individuals often face significant challenges in navigating their surroundings and interacting with their environment due to the absence of visual perception. Traditional assistive

tools, such as white canes and guide dogs, provide limited support and cannot offer real-time contextual information [1][2].

To overcome these challenges, an AI-powered assistive mobile application, Drishti, has been developed to provide a centralised and intelligent platform for environmental awareness and navigation assistance [4][5].



Figure 1: DRISHTI AI App LOGO

The system integrates advanced technologies, including computer vision, speech processing, and artificial intelligence, to deliver real-time feedback and guidance. It not only detects surrounding objects but also enables voice-based interaction, allowing users to communicate with the system effectively [10][12].

The proposed system is a mobile-based assistive solution that enables users to interact via voice commands and receive audio responses. It simplifies daily navigation tasks by providing functionalities such as environment detection, text recognition, and obstacle-aware guidance [8][9]. By combining mobile computing with AI technologies, the system reduces dependency on traditional methods and enhances user independence and safety [13][19].

Furthermore, the integration of intelligent processing and real-time feedback ensures efficient decision-making and seamless interaction between the user and the environment. This makes the system a reliable and scalable solution for assistive technology applications, thereby improving accessibility and quality of life for visually impaired individuals [3][15].

2. OBJECTIVES

The main objectives of the Drishti Assistive Navigation System are explained in detail as follows:

2.1 Assist Independent Navigation

The primary objective of this system is to enable visually impaired individuals to navigate their surroundings independently. It reduces dependence on external assistance by providing real-time audio guidance tailored to environmental conditions.

2.2 Improve Environmental Awareness

Visually impaired users often face difficulty in identifying objects around them. The system aims to enhance awareness by detecting and recognising objects in real time using computer vision techniques.

2.3 Reduce Navigation Errors and Risks

Traditional assistive methods may not provide sufficient feedback to avoid obstacles. The system minimises the risk of accidents by identifying obstacles and providing directional instructions such as stop, move left, or move right.

2.4 Real-Time Data Processing

Another objective is to enable fast and efficient processing of visual and voice data. The system ensures rapid object detection and delivers immediate audio feedback to improve response time.

2.5 Voice-Based Interaction

The system is designed to operate through a voice-driven interface, allowing users to interact without relying on visual input. This improves accessibility and usability for visually impaired individuals.

2.6 Intelligent Assistance Using AI

The integration of artificial intelligence allows the system to provide contextual and meaningful responses. The AI module processes user queries and generates intelligent responses to improve interaction.

2.7 User-Friendly Interface

The system is designed with a simple, easy-to-use interface so users can operate it without advanced technical knowledge. Voice commands further simplify interaction.

2.8 Text Recognition Capability

The system enables users to read printed text using the camera. It converts visual text into audio output, helping users access written information such as books and labels.

2.9 Efficient System Operation

By automating tasks such as object detection, navigation guidance, and voice interaction, the system improves overall efficiency and reduces manual effort.

2.10 Scalability and Future Expansion

The system is designed to be scalable and flexible. It can be enhanced in the future by integrating features such as GPS-based navigation, offline AI processing, and compatibility with wearable devices.

3. METHODOLOGY

The Drishti Assistive Navigation System is a mobile app built with Flutter. It aims to provide real-time help to visually impaired users. The system uses a structured, modular approach to manage user interactions, visual processing, and intelligent response generation efficiently.

The process starts with the initial setup and user interaction. The app requests key permissions such as camera access, microphone access, and location services. Once users grant these permissions, the system begins a voice-based onboarding process. Here, users provide their name and basic information through speech input. The system confirms this information with voice feedback to ensure correctness and usability.

After setup, the system operates through a voice-driven interface, activated by a specific wake word. When activated, users can interact with the system via voice commands to access features such as environment detection, text reading, and navigation assistance.

The visual processing part of the system uses TensorFlow Lite to enable real-time object detection on mobile devices. The camera continuously captures frames, which the model processes to identify objects and obstacles around the user. The detected information is transformed into audio feedback, providing the user with situational awareness.

The speech interaction module manages communication between the user and the system. Speech-to-Text (STT) converts voice commands into text, while Text-to-Speech (TTS) creates audio responses. This ensures smooth, fully accessible interaction without requiring visual input.

To improve system intelligence, the app integrates the Gemini API. This processes user questions and generates context-aware responses. With this, the system can go beyond fixed commands and offer useful help based on user input and the surrounding environment.

The system's backend logic is implemented in the app using a service-based architecture. Different modules, such as camera handling, voice processing, and AI interaction, operate independently but communicate well. This setup ensures smooth coordination between components and allows real-time processing .

Overall, the system architecture ensures effective data flow among the input, processing, and output layers, enabling reliable performance in real-world scenarios. The modular design makes the system scalable and open to future improvements, making it suitable for deployment as an assistive technology solution.

4. SYSTEM ARCHITECTURE

The Drishti Assistive Navigation System uses a layered architecture to organise the application into distinct functional parts. This setup improves performance, scalability, and maintenance. The system has four main layers: Input Layer, Processing Layer, Application Layer, and Output Layer. Each layer has a specific role.

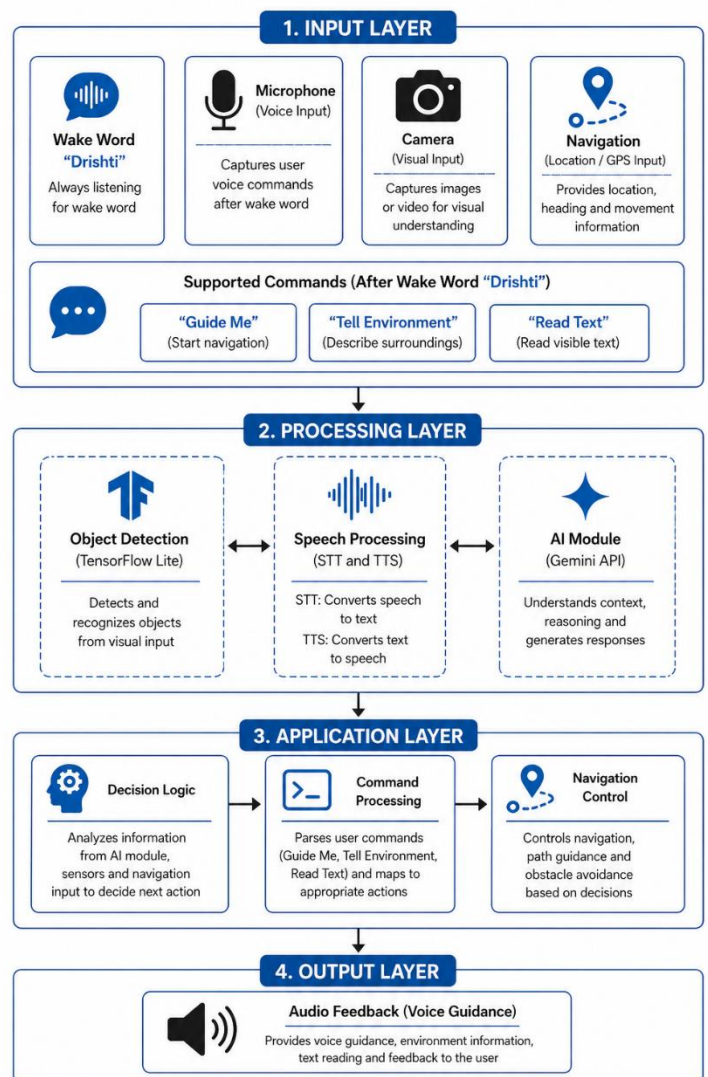


Figure 2: Dpicts all layers of DRISHTI AI Assistance

1. Input Layer

This layer includes the input sources that let the system connect with the user and the environment. It has a microphone to capture voice commands and a camera to

collect real-time visual data. This layer collects user input, such as voice instructions and environmental images, and sends it to the processing layer for analysis.

2. Processing Layer

This layer handles all key processing tasks of the system. It features several modules, such as object detection, speech processing, and AI-based analysis. The object detection module uses TensorFlow Lite to identify objects and obstacles from the camera input. The speech processing module turns voice input into text and produces audio responses. Moreover, the AI module (Gemini API) processes user queries and generates relevant replies. This layer serves as the central engine of the system.

3. Application Layer

The application layer oversees the system's overall logic and coordinates the different modules. It interprets user commands, controls the flow of operations, and applies logic to provide navigation assistance. For instance, based on the detected objects, the system decides whether to tell the user to move left, right, or stop. This layer ensures clear communication between the processing components and keeps the system running smoothly.

4. Output Layer

This layer presents the processed information to the user in an easy-to-understand way. The system primarily provides output via audio feedback, allowing users to receive real-time guidance without relying on visual cues. The output includes descriptions of objects, navigation instructions, and AI-generated responses, ensuring effective interaction and usability.

The layered architecture promotes efficient data flow between components, allowing real-time processing and reliable performance. It also makes it easier to modify and scale the system, which is helpful for future improvements and real-world use as an assistive technology solution.

5. FUNCTIONAL MODULES:

The Drishti Assistive Navigation System has the following modules:

1. User Interaction Module

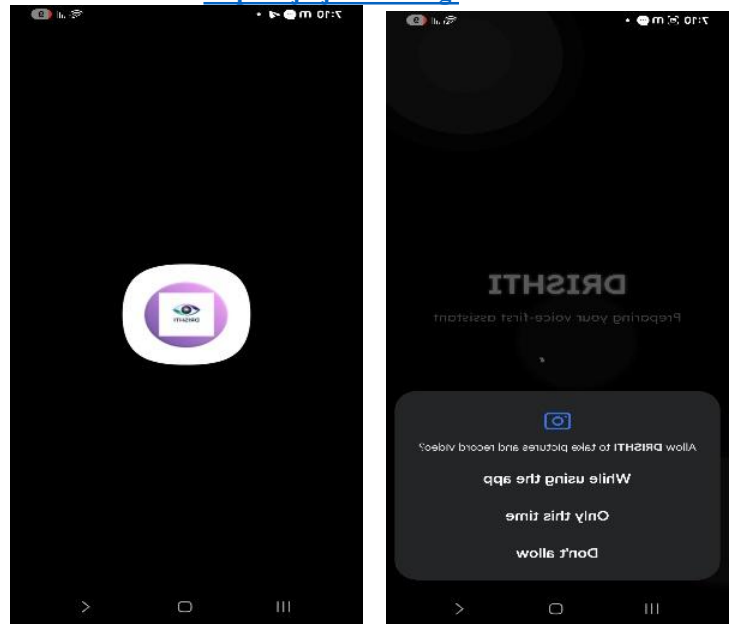


Figure 3.1: Drishti opening and asking for permissions

This module manages the first interaction between the user and the system. It handles permission requests for the camera, microphone, and location access. It also includes a voice-based onboarding process, where users provide their basic details through speech input and receive confirmation via audio feedback.

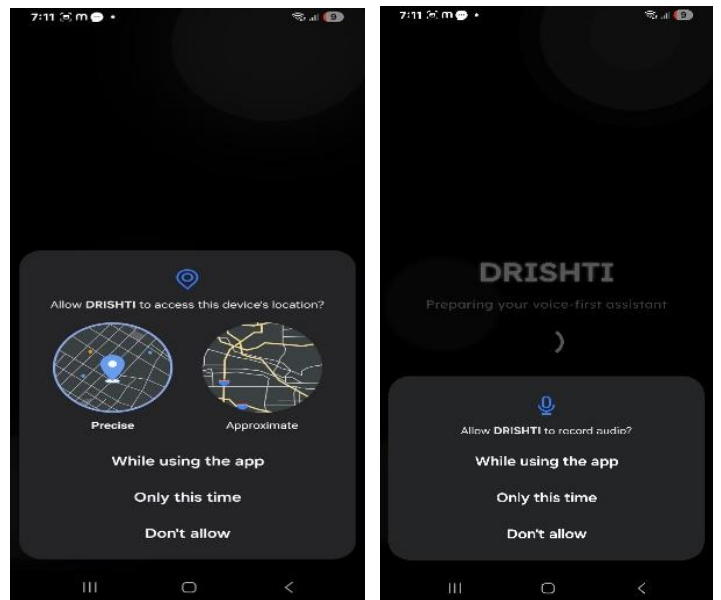


Figure 3.2: Drishti opening and asking for permissions

2. Voice Processing Module

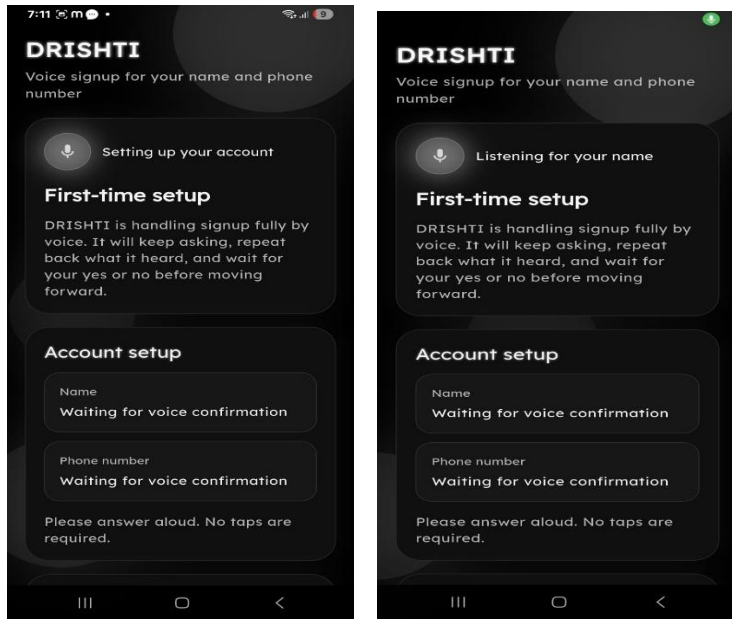


Figure 4.1: Account setup with Voice Input of Name and Phone Number from the user, and confirmation through voice output unless fully confirmed, and logs in

This module allows voice-based communication between the user and the system. It uses Speech-to-Text (STT) to convert user commands into text and Text-to-Speech (TTS) to generate audio responses. This allows the system to operate without visual interaction, ensuring accessibility for visually impaired users.

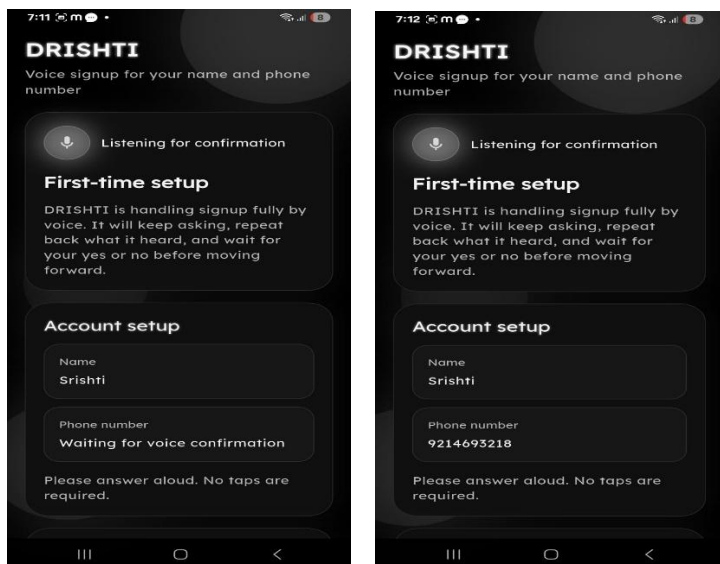


Figure 4.2: Account setup with Voice Input of Name and Phone Number from user and confirmation through voice output unless fully confirmed from user and logs in DRISHTI App

3. Object Detection Module

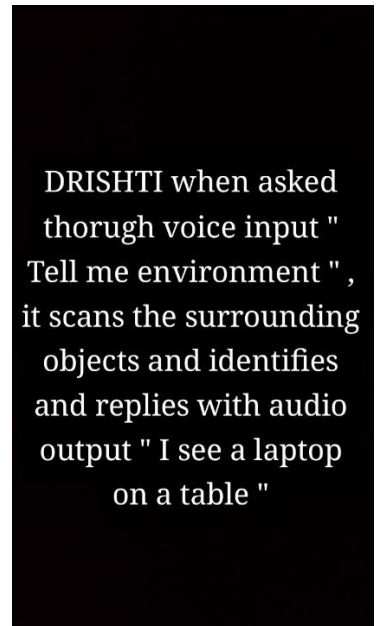
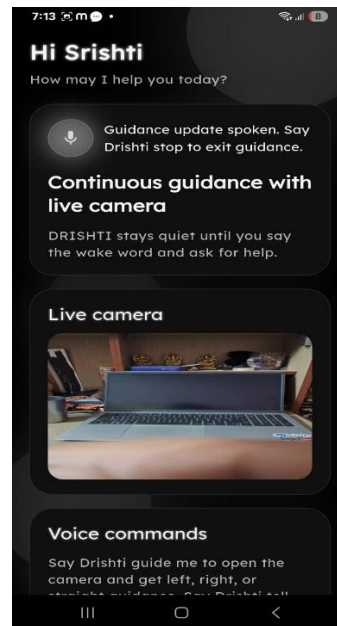


Figure 5: Object Identification and Environment Analysis with DRISHTI AI Assistance

This module identifies objects in the user's surroundings using the device camera. It uses TensorFlow Lite to process real-time visual input and detect objects such as obstacles, furniture, and other environmental elements. The detected information goes to the system for further processing.

4. Navigation Assistance Module

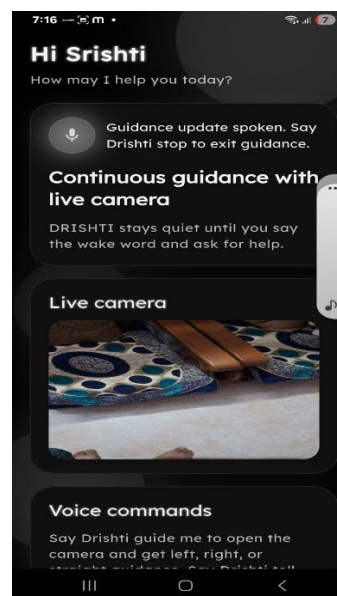


Figure 6: DRISHTI detects the environment and guides the user like a Virtual Cane.

This module gives real-time guidance to the user based on detected objects and obstacles. It applies decision-making logic to generate instructions such as “move left,”

“move right,” or “stop,” helping users navigate their environment safely.

5. Text Recognition Module

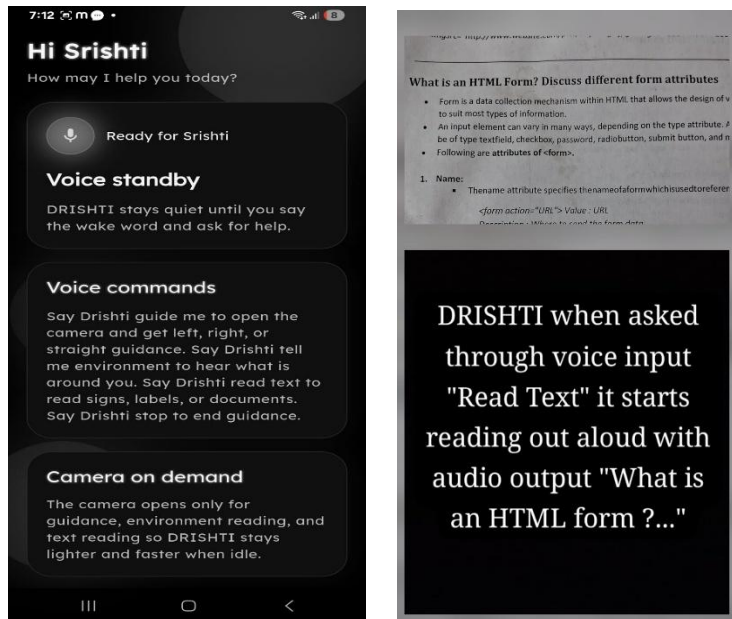


Figure 7: DRISHTI detects text and starts reading it out loud

This module allows the system to read printed text with the camera. It processes visual input to identify and extract written content, which is then converted into audio output. This helps users access information such as books, labels, and signs.

6. AI Interaction Module

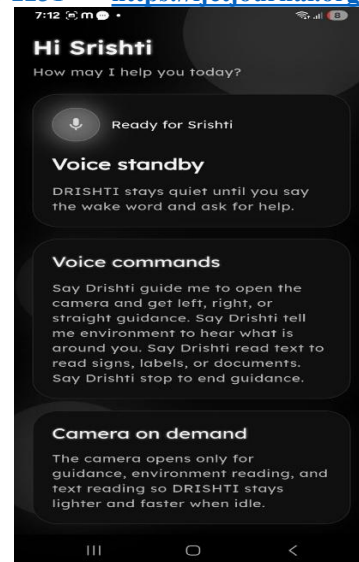


Figure 8: DRISHTI, when set up, wakes with the wake word “DRISHTI” and takes command.

This module integrates the Gemini API to deliver intelligent, context-aware responses. It processes user queries and generates meaningful outputs beyond preset commands, thereby improving the system's interaction and usability.

7. System Control Module

This module manages overall coordination between different parts of the system. It handles wake-word detection, controls the flow of operations, and ensures smooth communication among modules for efficient real-time performance.

The modular design of the system ensures flexibility, scalability, and efficient operation. Each module performs a specific function while working together to provide a seamless assistive experience for visually impaired users.

6. RESULTS AND DISCUSSION

The Drishti Assistive Navigation System shows how artificial intelligence, computer vision, and voice interaction can work together to help visually impaired users navigate and understand their environment in real time. This system reduces reliance on traditional assistive tools by providing smart guidance, thereby increasing user independence and safety.

The system captures real-time information about the surroundings by detecting objects and obstacles with the device's camera. Users can interact with the system via voice commands, enabling easy access to features such as environment detection, text reading, and navigation assistance. All processing occurs within a structured

framework, ensuring smooth operation and good coordination among system parts.

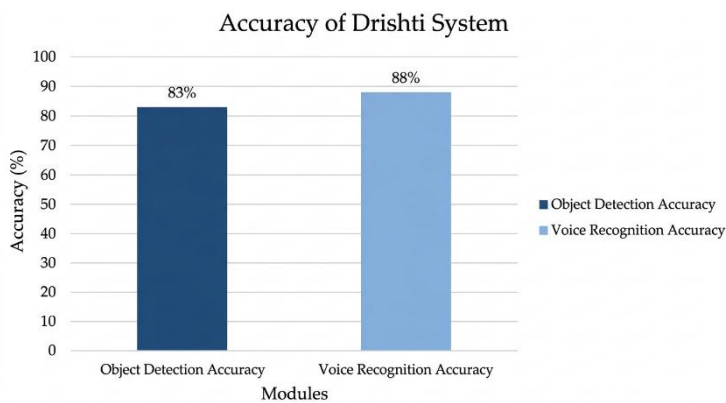


Figure 9: Accuracy Graph of DRISHTI AI Assistance

In experimental tests, the system was evaluated on a mid-range mobile device in various environments. The object detection module achieved an accuracy of about 80 to 85%, while the voice recognition module reached around 85 to 90%. The average response time for the system was between 2 and 5 seconds, depending on processing load and network conditions.

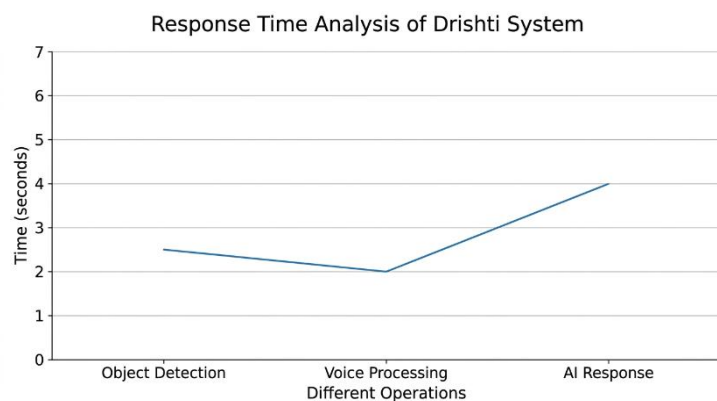


Figure 10: Response Time Analysis of DRISHTI Assistance

Parameter	Performance
Object Detection Accuracy	80–85%
Voice Recognition Accuracy	85–90%
Average Response Time	2–5 sec
Platform Tested	Samsung A14 5G With 4G SIM

Compared to traditional assistive tools, this system is faster, more interactive, and provides contextual information. It improves situational awareness by identifying objects and giving audio feedback instantly. The voice-driven interface makes it easy to use, allowing even those without technical knowledge to operate the system without visual input.

Furthermore, the system enhances overall accessibility by consolidating various functions into a single mobile app. It eliminates the need for separate assistive devices and offers a budget-friendly option that can be deployed on common smartphones. The use of AI enhances the system by enabling smart, context-aware responses.

However, the system has some limitations. Object detection performance may drop in low-light conditions or complex backgrounds. Additionally, the response time may fluctuate due to the simultaneous processing of camera input, voice commands, and API communication. The system also requires good device performance and stable network connectivity for the best operation.

Despite these drawbacks, the Drishti system delivers a reliable and efficient assistive solution. With further improvements, it has the potential to become a scalable and widely used technology for visually impaired individuals.

7. APPLICATIONS

The Drishti Assistive Navigation System can be used in a variety of real-world situations to help visually impaired individuals with their daily activities. It is especially beneficial in places such as homes, schools, public areas, transportation hubs, and workplaces, where navigating and identifying objects without visual cues can be difficult.

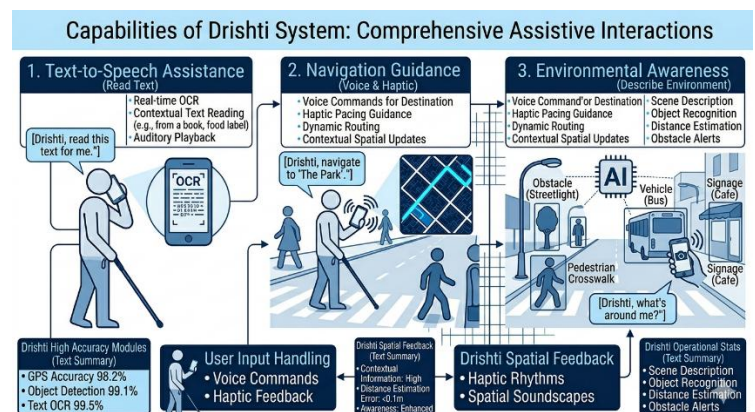


Figure 11: Illustration of a visually impaired individual using DRISHTI AI Assistance

The system offers a smart solution for environmental awareness by combining object detection, voice interaction, and navigation assistance in a single mobile application. This allows users to avoid using multiple assistive tools and enables them to perform tasks like identifying objects, reading text, and navigating safely in unfamiliar settings.

It boosts independence and confidence among visually impaired users by providing real-time audio feedback on their surroundings. The system lessens reliance on manual assistive methods and improves safety by detecting obstacles and guiding users accordingly.

Moreover, the application makes assistive technology more accessible by bringing it to commonly used smartphones. It reduces the cost and complexity associated with specialised hardware while maintaining efficiency and ease of use.

The system also enables efficient interaction via voice commands, allowing users to access information quickly and easily without visual input. This improves the user experience and makes the system suitable for a wide range of assistive applications.

Overall, the Drishti system enhances quality of life by supporting safer navigation, better understanding of the environment, and greater independence for visually impaired individuals in real-world situations.

8. CONCLUSION

The Drishti Assistive Navigation System offers a smart solution to help visually impaired individuals in digital and mobile environments. It replaces traditional assistive methods with an AI-driven approach that improves accessibility, efficiency, and user independence. By automating tasks like object detection, navigation guidance, and voice interaction, the system reduces manual effort and improves real-time decision-making.

The system is user-friendly and can be easily set up on common smartphones, making it accessible to many users. Its voice-driven interface allows visually impaired individuals to interact with the system without needing visual input or advanced technical skills. The integration of various functions into a single platform improves usability and convenience.

Additionally, the system boosts performance by providing real-time environmental awareness and structured interaction through smart processing. It enhances safety by detecting obstacles and guiding users while maintaining steady, reliable operation across different conditions. The modular design keeps the system organised, scalable, and ready for future upgrades.

Overall, the Drishti system advances assistive technology by integrating artificial intelligence,

computer vision, and mobile computing into a single solution. It offers a reliable, scalable, and cost-effective platform to improve the quality of life for visually impaired individuals and supports the larger goal of creating inclusive and accessible technology.

9. FUTURE SCOPE

The Drishti Assistive Navigation System has significant potential for future upgrades. The system can be enhanced with advanced AI models to improve object detection accuracy and environmental understanding. This will boost the system's performance in challenging situations, such as low-light areas and crowded spaces, making it more reliable overall.

The system can also work offline. This would enable important features such as object detection and basic voice interaction to function without internet access. This change would make it easier to use, especially in places with poor or unstable networks. Additionally, adding GPS-based navigation can offer location-aware guidance, helping users find their way in outdoor settings more effectively.

Moreover, the system can be made more useful by integrating it with wearable devices such as smart glasses or smart earphones. This would promote a seamless, hands-free experience for users. It would allow for real-time interactions, making the system practical for everyday life.

Features such as context-aware decision-making and predictive assistance could also be added using improved AI models. This would enable the system to anticipate users' needs and offer proactive help. Furthermore, performance optimisation methods could reduce response times and improve efficiency on low-end devices.

Security and privacy can be strengthened by implementing data protection mechanisms and handling user interactions securely. Safely processing voice and visual data is crucial for maintaining user trust and system reliability.

Overall, the Drishti system has strong potential for growth and innovation. With ongoing improvements and

the integration of new technologies, it could become a widely accepted assistive solution for visually impaired individuals.

REFERENCES:

- [1] Redmon J, Divvala S, Girshick R, Farhadi A. You Only Look Once: Unified, Real-Time Object Detection. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016.
- [2] Howard AG, Zhu M, Chen B, et al. MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications. arXiv preprint arXiv:1704.04861, 2017.
- [3] Sandler M, Howard A, Zhu M, Zhmoginov A, Chen LC. MobileNetV2: Inverted Residuals and Linear Bottlenecks. In: Proceedings of CVPR, 2018.
- [4] Goodfellow I, Bengio Y, Courville A. Deep Learning. Cambridge: MIT Press, 2016.
- [5] Szeliski R. Computer Vision: Algorithms and Applications. Springer, 2010.
- [6] Pressman RS. Software Engineering: A Practitioner's Approach. McGraw-Hill, 2005.
- [7] Sommerville I. Software Engineering. 9th ed. Pearson Education, 2011.
- [8] Google. TensorFlow Lite Documentation. Available: <https://www.tensorflow.org/lite>
- [9] Flutter. Flutter Documentation. Available: <https://flutter.dev>
- [10] Jurafsky D, Martin JH. Speech and Language Processing. Pearson, 2019.
- [11] Google AI. Gemini API Documentation. Available: <https://ai.google.dev>
- [12] OpenCV. OpenCV Documentation. Available: <https://opencv.org>
- [13] Android Developers. Speech Recognition and Text-to-Speech APIs. Available: <https://developer.android.com>
- [14] Lane ND, Bhattacharya S, Mathur A, et al. DeepX: A Software Accelerator for Low-power Deep Learning Inference on Mobile Devices. In: IPSN, 2016.
- [15] He K, Zhang X, Ren S, Sun J. Deep Residual Learning for Image Recognition. In: CVPR, 2016.
- [16] Simonyan K, Zisserman A. Very Deep Convolutional Networks for Large-Scale Image Recognition. arXiv, 2014.
- [17] Kameswaran SK, et al. Assistive Technologies for Visually Impaired: A Survey. IEEE Access, 2020.
- [18] Aladrén A, López-Nicolás G, Puig L, Guerrero JJ. Navigation Assistance for the Visually Impaired Using RGB-D Sensor. IEEE Systems Journal, 2016.
- [19] Patel PG, Shah SM. Survey on Data Security in Cloud Computing. International Journal of Engineering Research & Technology, 2012.
- [20] World Health Organization. World Report on Vision. WHO, 2019.
- [21] Chen LC, Papandreou G, Kokkinos I, Murphy K, Yuille AL. DeepLab: Semantic Image Segmentation with Deep Convolutional Nets. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2018.
- [22] Zhang D, Li X, Chen H. AI-Based Smart Assistive Navigation Systems for Visually Impaired. IEEE Access, 2021.
- [23] Bousbia-Salah M, Redjimi A, Fezari M, Bettayeb M. An Indoor Navigation System for the Visually Impaired. IEEE Transactions, 2011.
- [24] Khan SS, et al. Mobile Vision-Based Assistive System for Blind Users. International Journal of Computer Applications, 2018.