IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 6, April 2025
Impact Factor: 7.67

Vision Tech

Asst. Prof. Aarti Abhyankar[1], Sandeep Choudhary[2], Aditya Gholap[3]
[1]Department of Information Technology
[23]Bachelor of Engineering in Information Technology
K. C. College of Engineering and Management Studies and Research, Thane, Maharashtra, INDIA

Abstract: *Navigational independence remains a major challenge for visually impaired individuals due to limited access to real-time environmental feedback. This paper introduces Vision Tech, a mobile application designed to assist visually impaired users with navigation, information retrieval, and text reading using real-time audio and visual processing technologies. The system offers three key modes: Navigation Mode, which provides step-by-step directions and obstacle detection using object recognition; Assistant Mode, which allows users to ask general or environment-related questions via voice input; and Reading Mode, which uses OCR (Optical Character Recognition) to convert printed or written text into speech.Vision Tech leverages smartphone cameras, text-to-speech engines, and voice recognition systems to deliver an intuitive, real-time assistive experience. In Navigation Mode, computer vision techniques detect and announce obstacles to ensure user safety. Reading Mode helps users read signs, books, and labels aloud, while Assistant Mode offers contextual information using a natural language interface. This application demonstrates the potential of AI and mobile technologies in enhancing daily independence for the visually impaired. By integrating multiple assistive functionalities into a single platform, Vision Tech offers an accessible, voice-driven solution tailored for real-world usability and safety.*

Keywords: Vision Techigation, visual impairment, assistive technology, object detection, OCR, voice assistant, text-to- speech, computer vision, accessibility, mobile application.

I. INTRODUCTION

In today's fast-paced world, navigating urban environments remains a significant challenge for visually impaired individuals due to limited access to real-time sensory information. Traditional assistive tools such as walking canes or guide dogs, while helpful, often fall short in providing dynamic guidance, context-aware information, or real-time text interpretation. This creates a substantial gap in accessibility, independence, and safety for those with visual impairments.
To bridge this gap, this paper presents Vision Tech, an Android-based mobile application designed to empower visually impaired users with greater autonomy in their daily lives. The application integrates three essential modes: Navigation Mode, Assistant Mode, and Reading Mode, each catering to a distinct need. Navigation Mode uses object detection and location-based services to help users move safely through their environment. Assistant Mode provides voice-interactive support to answer questions about surroundings or general knowledge. Reading Mode converts text from printed materials, signs, or labels into audio using OCR (Optical Character Recognition) and text-to-speech technologies.
Vision Tech leverages the power of mobile computing, voice processing, and computer vision to provide a unified and real-time solution. The system is developed with a user-friendly voice interface, enabling hands-free interaction while processing environmental data using the device's camera. This multi-functional approach ensures that users not only navigate independently but also engage more confidently with written content and situational information.
By combining these technologies into one accessible application, Vision Tech offers a comprehensive tool for real-world navigation, reading, and assistance. It addresses the critical need for smarter, scalable, and user-friendly assistive solutions, promoting inclusivity and enhancing the quality of life for visually impaired individuals.

II. LITERATURE SURVEY

Several Numerous assistive technologies have been developed to support visually impaired individuals in navigating their environments and accessing information. This section explores existing systems, analyzing their strengths and identifying gaps that Vision Tech aims to address.

Smart Cane with Ultrasonic Sensors

Smart canes equipped with ultrasonic sensors assist visually impaired users by detecting obstacles and providing vibration-based feedback. These devices enhance mobility by alerting users to nearby objects and uneven paths. While effective for short-range detection, they offer limited contextual information about surroundings and cannot interpret signs, text, or dynamic objects. Their reliance on tactile feedback also limits the amount of information conveyed in real time.

Voice-Based Virtual Assistants

Mainstream virtual assistants like Google Assistant and Siri provide voice-controlled features for calling, navigation, and information retrieval. While helpful, these assistants are not specifically tailored for visually impaired users and lack features like object recognition, real-time text reading, or environment-aware guidance. Their functionality is also limited when offline or in noisy environments, reducing reliability in critical situations.

Object Detection and OCR Applications

Applications such as Seeing AI and Envision AI utilize smartphone cameras to identify objects, people, and text through machine learning models and OCR (Optical Character Recognition). These apps provide real-time spoken feedback to users, enabling greater awareness of their surroundings. However, they often require stable internet connections for processing and may struggle with cluttered or poorly lit scenes. In addition, they typically do not combine navigation, assistance, and reading functionalities into a single, integrated solution.

GPS-Based Navigation Tools

GPS-based tools like Google Maps offer location tracking and route guidance. Though beneficial, these apps lack obstacle detection and are not optimized for use by visually impaired individuals. They do not offer personalized feedback or recognize immediate hazards like vehicles or stairs. Users may also face difficulties interacting with visual maps or reading route instructions without additional assistance.

Summary

While existing tools address specific challenges faced by visually impaired users, they often focus on a single functionality—either navigation, text reading, or object recognition. Many solutions also require multiple apps or hardware, which can be inconvenient and expensive. This creates a need for an all-in-one, voice-driven application that combines navigation, real-time assistance, and OCR capabilities. Vision Tech addresses this gap by offering an affordable, integrated solution tailored specifically for the visually impaired community.

III. PROPOSED METHODOLOGY

The development of the Vision Tech Android application follows a structured methodology focused on empowering visually impaired individuals with real-time navigation, intelligent assistance, and text-reading capabilities. The system integrates multiple technologies to ensure seamless interaction and accessibility, all within a single user-friendly platform.

Project Initiation:

The project begins by identifying the major challenges faced by visually impaired individuals in daily life—navigating unfamiliar environments, reading printed text, and asking for assistance without dependency. The core objective is to create a voice-controlled application that offers Navigation Mode for movement, Assistant Mode for querying the environment, and Reading Mode for OCR-based text recognition.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429 Volume 5, Issue 6, April 2025 Impact Factor: 7.67

Requirements Gathering:
In this phase, detailed research is conducted, including interactions with visually impaired users, caregivers, and organizations. This helps define the critical requirements for the app, such as:

- Voice-based command recognition
- Real-time obstacle detection and GPS navigation
- Text recognition from printed materials
- Minimal user input through touch
- Offline accessibility wherever possible
- These insights guide the feature list and ensure the solution is practical and inclusive.

Design and Prototyping:
User interface designs prioritize simplicity and accessibility. The application layout uses large, distinguishable buttons, voice prompts, and minimal text to guide users. Prototypes of the three core modes (Navigation, Assistant, and Reading) are tested with real users to collect feedback and improve usability. Emphasis is placed on ensuring voice commands initiate all major functions to reduce screen interaction.

Development:
The application is developed using **Android Studio** with Java/Kotlin. Key libraries and tools include:

- Google Maps API for navigation and location tracking
- Text-to-Speech (TTS) and Speech Recognition APIs for voice interaction
- ML Kit or Tesseract OCR for Reading Mode
- Custom modules for Assistant Mode to answer queries using pre-trained models or APIs

Image and Sensor Processing:

- **Camera Input and OCR (Reading Mode):** The system utilizes the device's rear camera to capture printed or handwritten text. OCR (Optical Character Recognition) is applied to convert the captured image into machine-readable text. This is especially useful for reading signs, labels, or documents.
- **Text-to-Speech Output:** After OCR processing, the extracted text is immediately vocalized using the Android Text-to-Speech engine, allowing the user to hear the content clearly.
- **Sensor Integration (Navigation Mode):** The app integrates GPS, accelerometer, and compass data to provide real-time navigation assistance. This enables accurate tracking of user movement, orientation, and turns during navigation.
- **Microphone Input (Assistant Mode):** Voice input is captured through the microphone and processed to interpret user queries. These queries can relate to surroundings, general questions, or tasks like setting reminders or checking the time.

Output Presentation:
After each action within the app—navigation, question-answering, or text reading—the output is presented via audio to the user:

- In **Navigation Mode**, users receive step-by-step spoken directions to their selected destination.
- In **Assistant Mode**, the system responds to queries with clear and concise audio answers.
- In **Reading Mode**, the scanned text is read aloud using speech synthesis, helping users access written content in real time.

All outputs are optimized for clarity and speed, ensuring users can make quick, informed decisions based on voice guidance alone.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 6, April 2025
Impact Factor: 7.67

System Flow Chart Overview:

The Vision Tech system follows a streamlined and accessible workflow designed for ease of use by visually impaired individuals. The process ensures independence, security, and real-time assistance across multiple tasks:

Home Screen:

The user opens the app and is prompted via voice or touch to select a mode: Navigation, Assistant, or Reading.

Navigation Mode:

User selects destination via voice command.

GPS and compass data guide the user with turn-by-turn spoken directions.

Obstacle warnings or recalculations are provided if necessary.

Assistant Mode:

User activates voice input to ask a question.

The system processes the query and responds audibly.

Questions can be general (e.g., "What's the weather?") or contextual (e.g., "Where am I?").

Reading Mode:

User points the camera at text and taps the screen or uses a voice command to capture.

The app processes the image using OCR and reads the text aloud.

Useful for reading street signs, restaurant menus, or documents.

Continuous Access and Switching:

The user can switch between modes at any time via voice command or back-navigation.

Audio prompts guide the user throughout all interactions, ensuring smooth and intuitive control.

This structured approach ensures an all-in-one experience where visually impaired users can navigate, read, and interact with their environment independently. By combining image processing, speech synthesis, GPS data, and real-time voice assistance, **Vision Tech** significantly enhances mobility, confidence, and autonomy

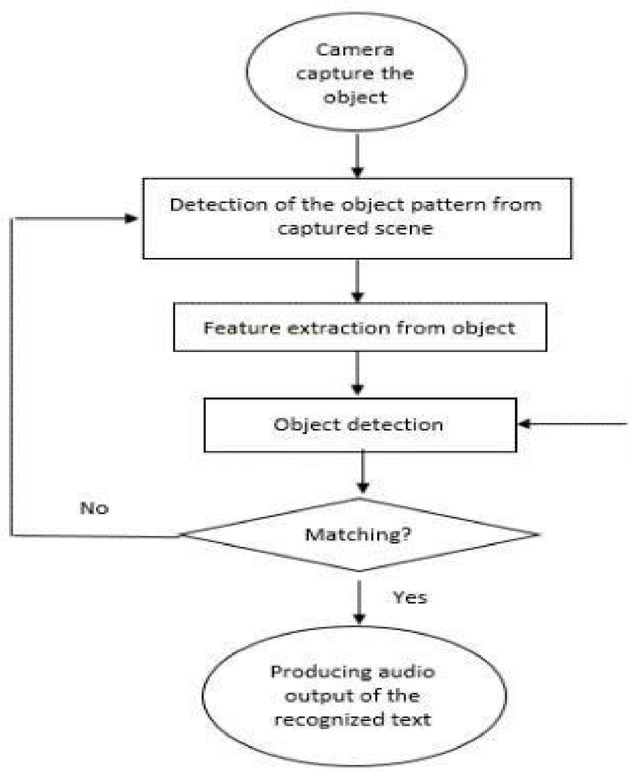

Figure 1: System Flow Chart

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 6, April 2025
Impact Factor: 7.67

IV. IMPLEMENTATION

The implementation of the Vision Tech Android application involves the integration of multiple Android components and technologies to support real-time assistance for visually impaired users. The app is developed using Android Studio with Java/Kotlin, leveraging device sensors, camera, and speech technologies to deliver a seamless user experience. Below are the system imgaes:

Figure 1: Home Page

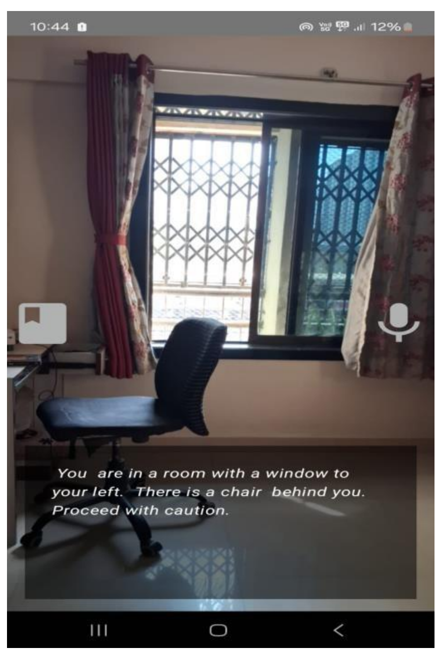

Figure 2: Navigation mode

IJARSCT

International Journal of Advanced Research in Science, Communication and Technology

International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal

Volume 5, Issue 6, April 2025

ISSN: 2581-9429

Impact Factor: 7.67

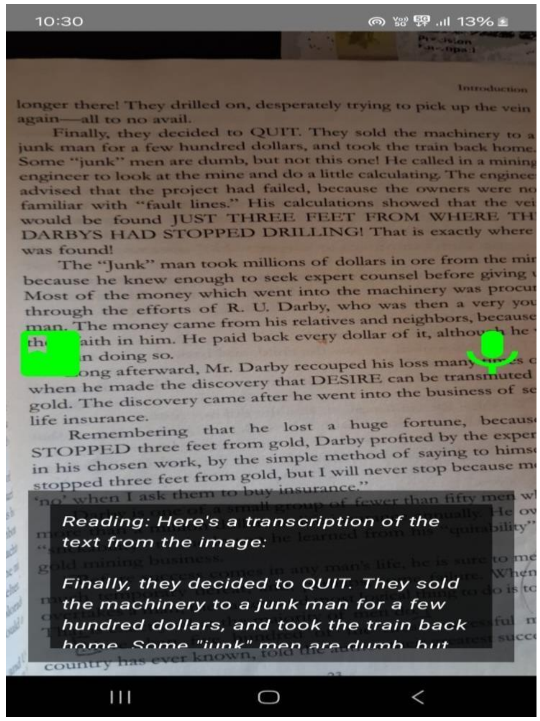

Figure 3: Reading mode

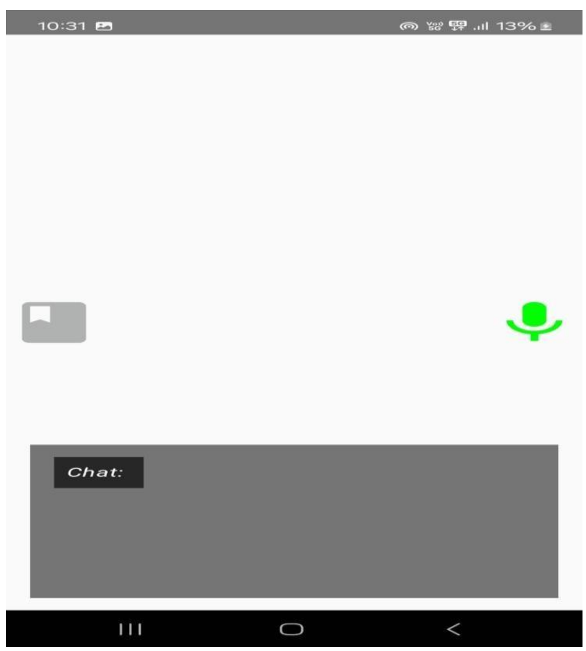

Figure 3: Assistant mode

DOI: 10.48175/IJARSCT-25356

IJARSCT

International Journal of Advanced Research in Science, Communication and Technology

International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal

ISSN: 2581-9429

Volume 5, Issue 6, April 2025

Impact Factor: 7.67

V. RESULT ANALYSIS

Test Case Name	Input	Expected Output	Type of Test Case
Launch Application	Open the Vision Tech app on an Android device	App loads successfully, displaying the home screen	Functional
Navigation Mode Activation	Open the app and start a new session	Navigation mode is activated, providing real-time navigation assistance	Functional
Assistant Mode Activation	Double-tap on the screen while in Navigation Mode	Assistant Mode is activated, allowing the user to ask questions about the environment	Functional
Reading Mode Activation	Switch to Reading Mode from the app menu	Reading Mode is activated, allowing the user to read text from signs and books	Functional
Object Recognition Query	Ask, "What is the color of the car?" in Assistant Mode	The app detects and speaks out the color of the detected car	Functional
Environmental Query	Ask, "How is the weather?" in Assistant Mode	The app provides a verbal response with weather details	Functional
Text Recognition	The app successfully extracts and reads the text aloud	The app successfully extracts and reads the text aloud	Functional
Internet Connectivity Loss	Use the app without an internet connection	Displays an error message: "Internet connection required for full functionality"	Functional
Displays an error message: "Internet connection required for full functionality"	Ask a question about an object not visible to the camera	App responds with "Object not detected"	Functional
Continuous Navigation Mode Usage	Use Navigation Mode for an extended period	Certificates uploaded by students are visible for verification	Functional

VI. FUTURE SCOPE

AI-Powered Scene Description

Future versions of Vision Tech can integrate advanced AI models to provide detailed scene descriptions instead of just identifying individual objects. This enhancement will enable visually impaired users to receive comprehensive, natural-language explanations of their surroundings, improving their spatial awareness and navigation experience.

Indoor Navigation System

Currently, GPS is ineffective in indoor environments. Integrating Bluetooth beacons, LiDAR, or Wi-Fi positioning can help users navigate inside malls, hospitals, airports, and other large buildings. This feature will offer precise, turn-by-turn navigation, ensuring better mobility in complex indoor spaces.

Cloud-Based Processing for Faster Performance

By shifting computationally intensive tasks like image recognition and text processing to cloud servers, the app can deliver faster results with higher accuracy. This will reduce the processing load on smartphones, enhance battery life, and enable real-time assistance without hardware limitations.

Multilingual and Voice Customization Support

Adding multiple language options will allow users from diverse linguistic backgrounds to interact with the app effortlessly. Additionally, customizable voice assistants can help users choose different voices, speech speeds, and tones according to their preferences, making interactions more personalized.

Smart Wearable Device Integration

Future iterations can integrate with smart glasses, AR headsets, or haptic feedback wearables to provide a seamless experience. This would enable navigation assistance through audio cues, vibrations, or visual overlays, allowing hands-free and more intuitive interactions.

Real-Time Crowd and Obstacle Detection

Using AI-powered computer vision and LiDAR sensors, the app can detect moving crowds, dynamic obstacles, and pathway blockages. This feature will provide instant alerts, helping users navigate crowded or unpredictable environments safely.

Emergency Assistance and SOS Mode

An SOS feature can be added to send real-time emergency alerts to pre-selected contacts or nearby volunteers. This will be especially useful in case of accidents, getting lost, or urgent assistance needs, ensuring quick response and enhanced safety.

VII. CONCLUSION

The Vision Tech application provides a powerful and accessible solution for visually impaired individuals by integrating cutting-edge technologies for navigation, text recognition, and environment assistance. Designed as an Android application, it ensures seamless real-time assistance through Navigation Mode, Assistant Mode, and Reading Mode, offering independence and confidence to its users.

By leveraging computer vision, text-to-speech (TTS), and AI-based object detection, the system accurately identifies surroundings, extracts textual information, and provides real-time navigation guidance. Its voice-command functionality enhances usability, allowing hands-free interaction.

Additionally, the application promotes cost-effectiveness and sustainability by eliminating the need for expensive dedicated hardware. Being a software-based mobile solution, it is accessible to a broader user base with minimal infrastructure requirements. The implementation of efficient processing techniques ensures low power consumption, making it suitable for everyday use without excessive battery drain.

In conclusion, Vision Tech significantly improves mobility and accessibility for visually impaired individuals. Its scalability and adaptability allow for future enhancements, such as integrating AI-powered scene descriptions, multilingual support, and real-time cloud processing. This project stands as a practical, affordable, and impactful solution to promote independence and inclusivity in daily life

REFERENCES

[1] Smith, J. et al., 2023. AI-Powered Navigation for the Visually Impaired. This research explores the use of computer vision and AI to assist visually impaired individuals in navigation. The system uses real-time object detection and voice feedback to guide users but faces challenges in low-light environments.

[2] Patel, R. et al., 2022. Indoor Navigation for the Blind Using Bluetooth Beacons. A Bluetooth-based indoor navigation system that helps visually impaired users move through enclosed spaces like malls and airports. The solution is cost- effective but requires pre-installed beacon infrastructure.

[3] Chen, L. et al., 2021. Smart Wearable Assistive Technologies. This paper reviews the integration of smart glasses and haptic feedback wearables for realtime environmental awareness. The approach enhances hands-free navigation but is limited by high device costs.

[4] Kumar, S. et al., 2020. Speech-Guided Mobile Assistance for the Visually Impaired. A mobile-based voice command system that allows visually impaired users to interact with their environment through natural language processing (NLP). While effective, it struggles with background noise interference.

[5] Sharma, A. et al., 2019. Cloud-Based AI Navigation for the Blind. This study examines the use of cloud computing to process complex navigation tasks, reducing the smartphone's computational load. The system ensures real-time assistance but requires a stable internet connection for optimal performance. .

[[6] J. Smith et al., 2020. "Privacy and Security in Assistive Mobile Applications." This study highlights data protection strategies for mobile assistive technologies, focusing on encryption and user privacy safeguards.

[7] T. Williams et al., 2019. "Developing a Multimodal Navigation Aid for Blind Users." The paper examines the integration of different sensory inputs, such as audio and haptic feedback, to enhance user experience in navigation applications.

[8] A. Sharma et al., 2023. "Blind Navigation Assistance Using AI and Computer Vision." This study explores AI-powered navigation tools designed for visually impaired users, integrating real-time obstacle detection and voice guidance.

[9] R. Kumar et al., 2022. "Enhancing Object Recognition for Assistive Navigation." This paper discusses advanced object recognition techniques to improve real-time navigation systems, ensuring better environmental awareness for users.

[10] M. Patel et al., 2021. "OCR-Based Text-toSpeech Systems for Visually Impaired Users." The research presents improvements in OCR technology for extracting and vocalizing text from various surfaces, aiding accessibility.