Drishti: An AI-Based Navigation and Assistance System for Visually Impaired People

Mrs. Shraddha Raut[1], Shruti Kawale[2], Trupti Seware[3], Kasturi Awachat[4], Kritika Pandey[5], Vedant Wandhare[6]

[1,2,3,4,5,6]*Department of Computer Science and Engineering, G H Raisoni University, Amravati, India*

***Abstract*—People with visual impairments often face challenges while travelling to unfamiliar places, recognizing objects around them, reading printed information, and avoiding obstacles. Although several assistive technologies are available, many are designed to perform only one specific task, such as navigation or text recognition. To address these challenges, this paper presents Drishti, an AI-based navigation and assistance system designed to bring multiple assistive features together in a single mobile application. The proposed system aims to combine voice-based interaction, location-based navigation, object detection, optical character recognition (OCR), AI-based scene description, and emergency assistance. It is designed to provide useful information through audio feedback, allowing users to interact with the application without depending heavily on the smartphone's visual interface. Drishti aims to make everyday navigation and access to environmental information more convenient for visually impaired users. The performance of the system will be assessed through testing of its implemented features.**

***Index Terms*—Artificial Intelligence, Computer Vision, Assistive Technology, Visually Impaired People, Navigation, Object Detection, OCR, Voice Assistance.**

I. INTRODUCTION

For people with visual impairments, everyday activities such as walking in unfamiliar areas, identifying nearby objects, reading signboards, and finding the correct route can be challenging. A white cane is a valuable mobility aid, but it may not provide information about written text, object identity, or a destination located farther away.

With recent developments in artificial intelligence, computer vision, GPS, and mobile technology, it has become possible to develop applications that can provide additional support to visually impaired users. For example, computer vision can help identify objects in a camera image, while GPS and navigation services can provide directions to a selected destination. Voice-based interaction can make these features easier to access.

This paper introduces Drishti, a proposed mobile-based assistance system that aims to combine navigation and visual information assistance in one application. Users can interact with the system through voice commands and receive relevant information through audio feedback.

The main idea behind Drishti is to make useful assistance more accessible by bringing different functions together in a single platform.

II. LITERATURE SURVEY

TABLE I. LITERATURE SURVEY

Ref.	Authors	Research Paper
[1]	B. Kuriakose, R. Shrestha, F. E. Sandnes	DeepNAVI: A Deep Learning Based Smartphone Navigation Assistant for People with Visual Impairments
[2]	B. Li, J. P. Muñoz, X. Rong, Q. Chen, J. Xiao, Y. Tian, A. Arditi, M. Yousuf	Vision-Based Mobile Indoor Assistive Navigation Aid for Blind People
[3]	R. Tapu, B. Mocanu, T. Zaharia	DEEP-SEE: Joint Object Detection, Tracking and Recognition with Application to Visually Impaired Navigational Assistance
[4]	F. A. Rahman, F. A. Pusparani, W. L. Yeoh, O. Fukuda	Smartphone-Based Seamless Scene and Object Recognition for Visually Impaired Persons

Ref.	Authors	Research Paper
[5]	Q. Gao et al.	VI-OCR: "Visually Impaired" Optical Character Recognition Pipeline for Text Accessibility Assessment
[6]	H. Zhang, N. J. Falletta, J. Xie, R. Yu, S. Lee, S. M. Billah, J. M. Carroll	Enhancing the Travel Experience for People with Visual Impairments through Multimodal Interaction: NaviGPT, A Real-Time AI-Driven Mobile Navigation System
[7]	N. Bayat, J.-H. Kim, R. Choudhury, I. F. Kadhim, Z. Al-Mashhadani, M. A. Dela Virgen, R. Latorre, R. De La Paz, J.-H. Park	Vision Transformer Customized for Environment Detection and Collision Prediction to Assist the Visually Impaired
[8]	J. Bai, Z. Liu, Y. Lin, Y. Li, S. Lian, D. Liu	Wearable Travel Aid for Environment Perception and Navigation of Visually Impaired People
[9]	R. B. Islam, S. Akhter, F. Iqbal, M. S. U. Rahman, R. Khan	Deep Learning Based Object Detection and Surrounding Environment Description for Visually Impaired People
[10]	Author list to be confirmed from the full IEEE record	Safer Navigation: The AI-Powered Smart Cane for the Visually Impaired

III. PROBLEM STATEMENT

Visually impaired individuals may require different types of assistance while travelling or performing daily activities. These may include finding a destination, identifying objects, reading text, and understanding their surroundings.

However, when these functions are provided through separate applications or devices, users may need to switch between different tools to access the required information.

This creates a need for a mobile-based assistive system that can bring navigation, environmental information, and voice interaction together in one platform. Drishti is proposed to address this need by integrating multiple assistance features into a single application.

IV. OBJECTIVES

The main objective of the Drishti system is to provide navigation assistance through voice-based commands. It aims to help users identify objects in their surroundings and to provide information about potential obstacles.

The system also reads visible text using Optical Character Recognition (OCR), generates descriptions of the surrounding environment using AI, and provides an SOS feature for emergency assistance. All useful information is delivered through audio feedback, which makes the application easier to access without visual interaction.

V. PROPOSED SYSTEM

A. Voice Interaction

Drishti allows the user to give commands using their voice. The system understands the command and performs the required action. This reduces the need to use the smartphone screen.

B. Navigation Assistance

The user can tell Drishti the destination using a voice command. The system uses location and navigation services to provide route information.

C. Object Detection

The camera captures the user's surroundings. The AI model identifies objects in the image and provides the information to the user through audio.

D. Obstacle Assistance

The obstacle assistance component aims to provide information about potential obstacles detected in the user's surroundings. The purpose is to give users additional environmental awareness while moving. The accuracy of this feature depends on the detection method, camera conditions, and the types of obstacles the system can recognize. It is intended to supplement, not replace, established mobility aids.

E. Optical Character Recognition (OCR)

The camera captures text such as a signboard or printed document. OCR recognizes the text, and the system reads it aloud to the user.

F. AI-Based Scene Description

The scene description component analyses a camera image using AI and computer vision processing and generates a description of the surroundings. This information is then given to the user as audio feedback.

G. SOS Assistance

The SOS feature allows the user to request help during an emergency. Depending on the implemented feature, it can send an alert or share the user's location.

VI. SYSTEM ARCHITECTURE / FLOWCHART

The proposed architecture of Drishti consists of user input, processing modules, and audio output.

The user interacts with the application through voice commands or camera input. The input is processed by the relevant module, such as speech processing, navigation services, computer vision, or OCR. The resulting information is then prepared for audio feedback.

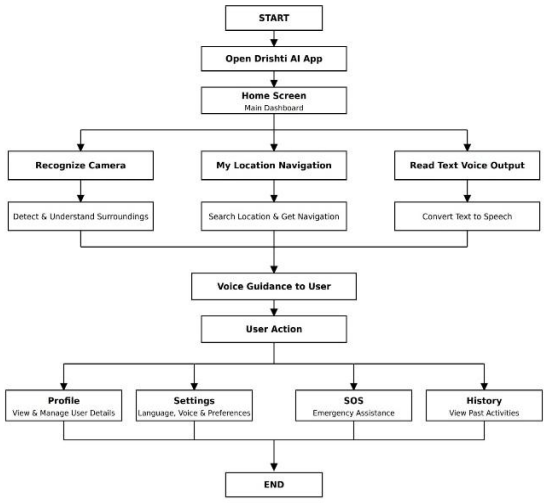

Fig. 1. Proposed System Architecture of Drishti

As shown in Fig. 1, the flow begins when the user opens the Drishti application and reaches the home screen. From there, the user can use the camera to detect and understand the surroundings, search for a location and get navigation, or convert text to speech to have it read aloud. The result of each function is given as voice guidance, after which the user can open the profile, settings, SOS, or history sections before the process ends.

The overall workflow can be represented as: User Input (Voice / Camera) → Input Processing → Relevant Module (AI / Computer Vision / Navigation / OCR) → Information Processing → Text-to-Speech → Audio Feedback to User. The navigation module and camera-based assistance modules perform different functions but are intended to work within the same application.

VII. METHODOLOGY

Step 1: User Input

The user provides input through voice commands or the smartphone camera.

Step 2: Input Processing

Voice input is processed using speech recognition, while camera input is processed using computer vision and OCR techniques.

Step 3: Functional Processing

Depending on the user command, the system performs the required operation. For navigation, the voice command is used for destination recognition, location retrieval, and route generation, followed by audio guidance. For object detection, the camera input goes through image processing and object detection, after which information is generated and given as audio feedback.

For OCR, the camera input goes through text detection and OCR, and the recognized text is converted to speech. For scene description, the camera input is processed using AI and computer vision to obtain scene information, which is delivered as audio feedback. For SOS, the SOS command triggers emergency processing and a predefined emergency action.

Step 4: Information Generation

The processed information is converted into a form that can be understood by the user.

Step 5: Audio Feedback

The final information is converted into speech and delivered through audio feedback.

Step 6: Real-Time Processing
The system processes voice and camera inputs in real time to provide quick and relevant feedback. This helps the user receive information without performing complex manual operations.

Step 7: Emergency Assistance
When an SOS command is detected, the system initiates the predefined emergency procedure. This feature is designed to provide quick assistance during critical situations.

Step 8: System Output
The final processed information is converted into clear and understandable audio output. This allows the user to access the required information without depending on visual interaction.

VIII. TECHNOLOGIES USED

TABLE II. TECHNOLOGIES USED

Technology	Purpose
Flutter	Mobile application development
GPS and Geolocation	Obtaining the user's location
Maps/Navigation APIs	Route guidance
Smartphone Camera API	Capturing images for visual assistance
AI and Computer Vision	Object detection and scene analysis
Optical Character Recognition (OCR)	Recognizing text from images
Speech-to-Text	Converting voice commands into text
Text-to-Speech	Providing spoken feedback to the user

IX. IMPLEMENTATION

Drishti integrates its assistance functions through a single mobile interface, and voice input is used to receive user commands. The camera input supports object detection, OCR, and scene understanding. The navigation module uses location and map services to provide route-related information, while the SOS function is intended to support emergency assistance. The information processed by the different modules is converted into clear audio feedback using Text-to-Speech, so that users can access the main features without depending heavily on the visual interface.

X. RESULTS AND DISCUSSION

The results of Drishti should be discussed based on observations collected during testing. The evaluation may consider the accuracy of voice command recognition, object identification, text/OCR recognition, navigation information, response time, and the consistency of audio feedback.
The collected results can help identify the features that perform as expected and the features that require further improvement. System performance should be supported by actual testing results rather than assumptions.

XI. FUTURE SCOPE

The proposed Drishti system can be extended in several ways. Real-time obstacle detection, scene understanding, and indoor navigation can be made more accurate, and support for multiple Indian languages can be added. The system can also be integrated with smart glasses or wearable devices and with additional accessibility services. Emergency location sharing can be improved, along with conversational AI assistance and personalized assistance based on user requirements.

XII. CONCLUSION

This paper presents Drishti, a proposed AI-based navigation and assistance system designed to support visually impaired individuals in their daily activities. The system aims to combine voice interaction, location-based navigation, object detection, OCR, AI-based scene description, and emergency assistance within a single mobile application.
By providing information through audio feedback, Drishti aims to make navigation and access to environmental information more convenient for users who may find visual interaction with a smartphone difficult.
The actual effectiveness of the system will depend on the successful implementation and testing of its individual modules. Further evaluation can help identify its strengths, limitations, and opportunities for improvement.

ACKNOWLEDGMENT

Acknowledgment text can be added here if required by the college, guide, or publication venue.

REFERENCES

[1] B. Kuriakose, R. Shrestha, and F. E. Sandnes, "DeepNAVI: A deep learning based smartphone navigation assistant for people with visual impairments," *Expert Systems with Applications*, vol. 212, Art. no. 118720, 2023, doi: 10.1016/j.eswa.2022.118720.

[2] "Vision-based mobile indoor assistive navigation aid for blind people," *Sensors*, 2019, doi: 10.3390/s19040828.

[3] F. Ceccarelli *et al*., "DEEP-SEE: Joint object detection, tracking and recognition with application to visually impaired navigational assistance," *Sensors*, 2017.

[4] "Smartphone-based seamless scene and object recognition for visually impaired persons," *Information*, vol. 16, no. 9, Art. no. 808, 2025, doi: 10.3390/info16090808.

[5] Q. Gao, R. Manduchi, P. Y. Ramulu, G. E. Legge, *et al*., "VI-OCR: 'Visually Impaired' optical character recognition pipeline for text accessibility assessment," *Scientific Reports*, vol. 16, Art. no. 1269, 2026, doi: 10.1038/s41598-025-30982-7.

[6] "Enhancing the travel experience for people with visual impairments through multimodal interaction: NaviGPT, a real-time AI-driven mobile navigation system," 2024.

[7] "Vision transformer customized for environment detection and collision prediction to assist the visually impaired," *Sensors*, 2023.

[8] "Wearable travel aid for environment perception and navigation of visually impaired people," *Electronics*, vol. 8, no. 6, Art. no. 697, 2019.

[9] "Deep learning-based object detection and surrounding environment description for visually impaired people," *Heliyon*, vol. 9, Art. no. e16924, 2023, doi: 10.1016/j.heliyon. 2023.e16924.

[10] "Safer navigation: The AI-powered smart cane for the visually impaired," in *Proc. 16th Int. Conf. Developments in eSystems Engineering (DeSE)*, 2023, doi: 10.1109/DeSE60595.2023.10469180.