AI-Powered Voice Recognition Navigation Aid For Accessibility

- **Team Members**: Sarvesh K(2023PECCS479)
- Sri Hari T K(2023PECCS503)
- **Department**: Computer Science Engineering
- **Domain**: Artificial Intelligence
- **Guide**: MR.Sasikumar A N
- **Coordinator:** Dr.Subetha V

AGENDA

- Introduction
- Objectives
- Literature Survey
- Problem Statement
- Architecture Diagram
- URL Diagrams
- Tools and Technology
- System Architecture
- Implementation
- Test Cases and Validation Testing
- Performance Analysis
- Demo Screenshot
- Conclusion
- Future Enhancement
- References

Introduction

- The project aims to assist visually impaired individual in navigating safely and independently.
- Traditional navigation system rely on visual maps,making them unsuitable for blind users.
- It combines **AI-based voice recognition,object detection,and navigation guidance**.

Objectives

- Convert voice to text using Whisper.
- Find destination routes using OpenRouteService.
- Give step-by-step audio navigation.
- Use camera + AI for object detection.
- Display route on interactive map.

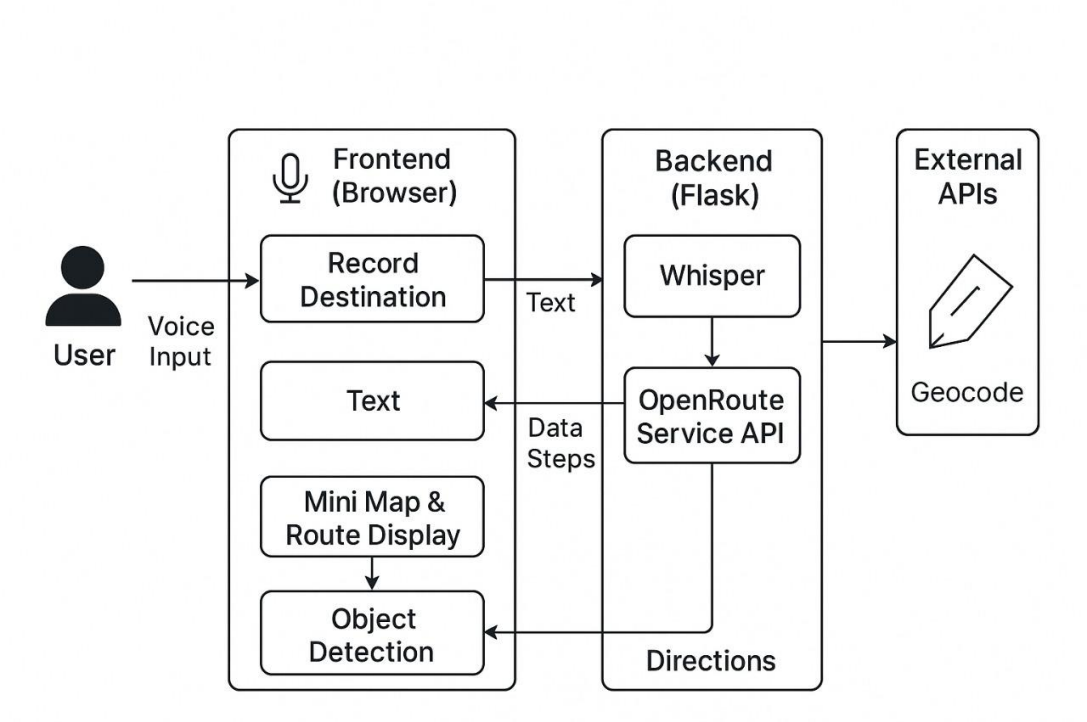
Literature Survey

S.No.	Author(s) & Year Title	Methodology / Approach	Key Findings / Contributions
1	Patel, R. et att. (2023) Voice-Actiated Navigation for the Visually Visualy Impaired Using AI	GPS-enabled voice navigation with AI-based speech recogrition	Improved independent moblity mobility in outdoor environments
2	Santos. P. ett. (2021) Ddoor Navigation Aid Speech Recogrition for Computer Vision and Voice Commands	Hybrit CNN-RNN speech recogritien model detection	Higher accuracy in real-ritme voice commanɑ indo rutes
4	Kumar S. eta. (2023)	AI-driven path optimization algorithm	Provided shortes, obstetle-fre routes
4	Rahman M. (2020 for Accessbility - ACM	IOT sensors + GPS + voice alerts	Speech Recogrition Accuracy in Envronments
5	IOT-Based GPS and Voice Assistance System System for the Blind	Accurate obstatle ate moeling Provided shottion and navigation	Noise reduction + acoueling Improved recogition in crouwed spaces
6	Gonzalelz. A. ett. (2021) Sensors Journal	Low-Power Embedded Voice Recogrition - IBEE	Wearble device with + voice + feedback
8	Ahmedz T, & Park J. (2020 Wearable Aid - Raigation Aid - Sensors Journal	Energy-efficient embedded systems	Increased obstatle portable for awareness
9	Fernandez L. ett. (2023) Multiungeal Voice Interfaces for Accessbility - HCIS	Suitable for deployed on AI model deployed on edge devices	Expanded usallity for Expanded usallity for diverse users
10	Sharma V. eta. (2021)	IEE Ede Computing	Redued in navigation feedback

Problem Statement

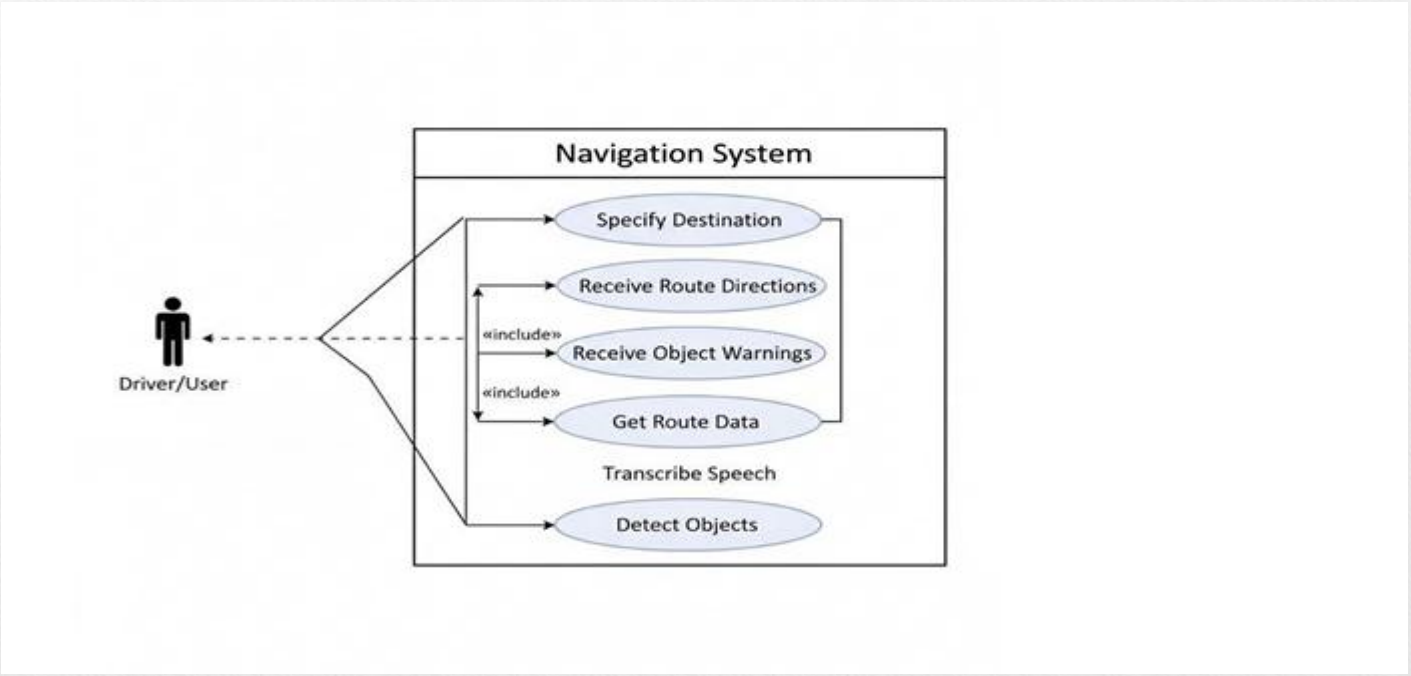
- **Problem**: Visually impaired people face difficulty using tradional GPS app because they require typing or visual interaction.
- **Solution**: An AI- powered voice navigation system with obstacle detection that provide hands-free,safe,and accessible travel.

Architecture Diagram

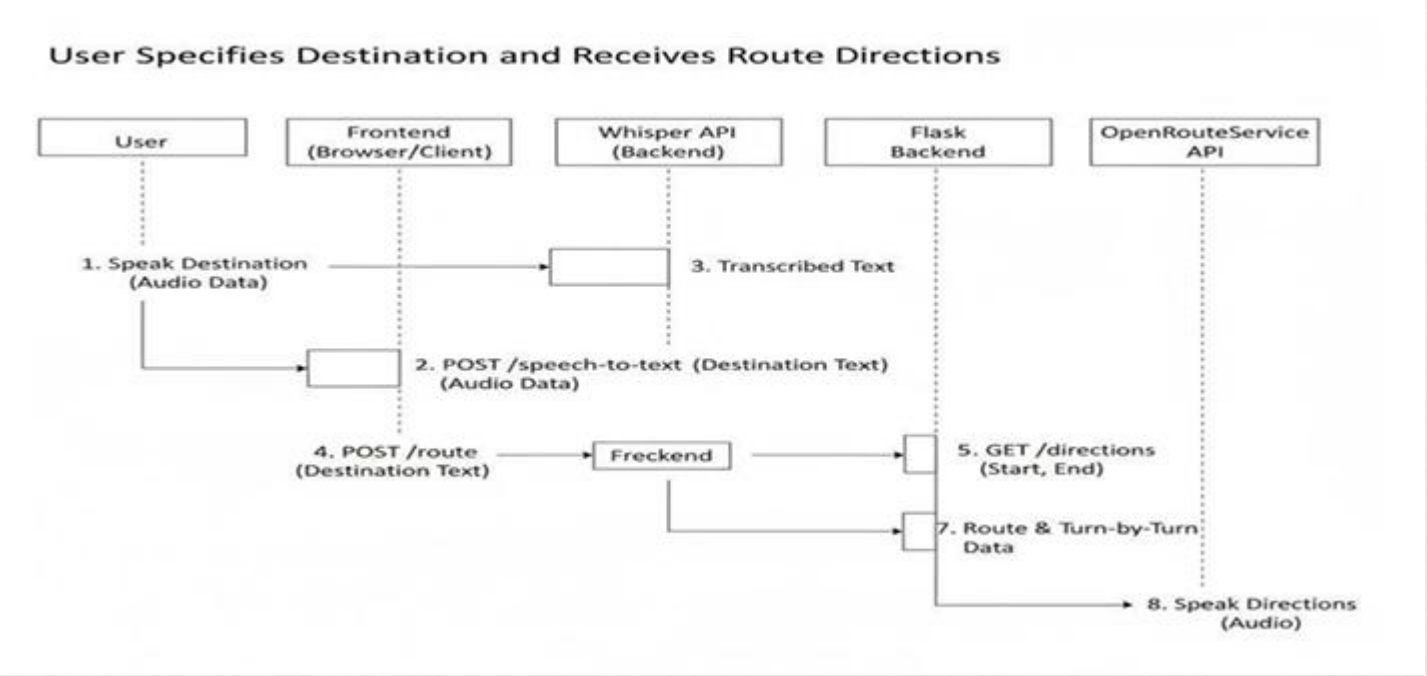


URL Diagram

Use Case Diagram:



Sequence Diagram:



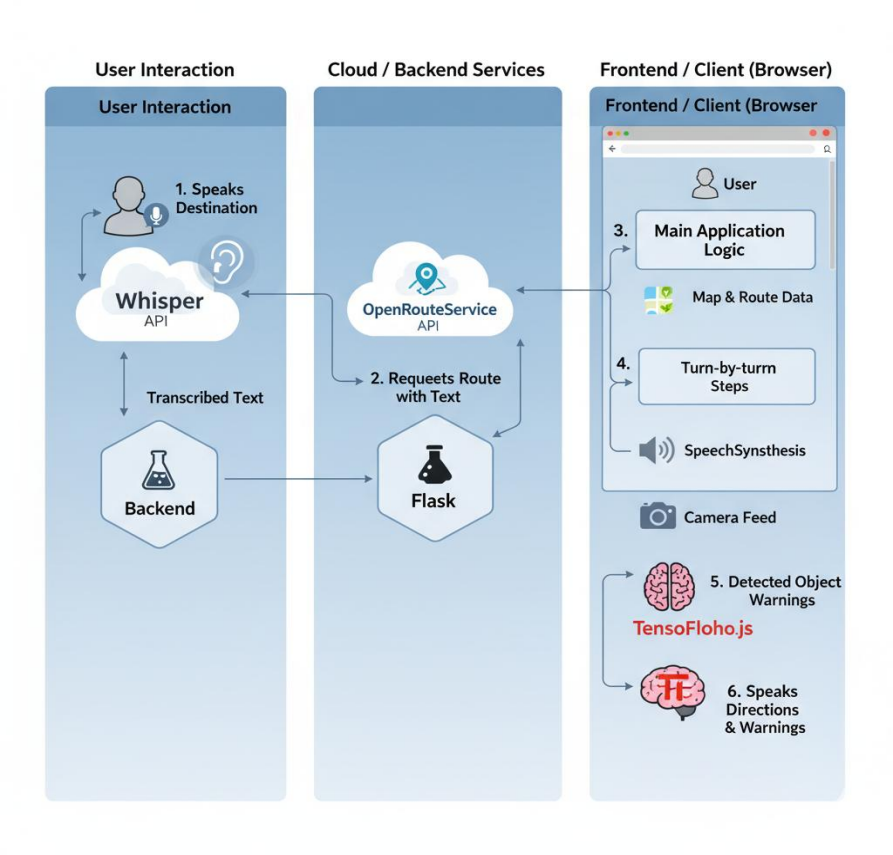
Tools & Technologies

- **AI Tools:**
- - OpenAI Whisper → Speech to Text
- - TensorFlow.js COCO-SSD → Object Detection

- **Supporting Tools:**
- - OpenRouteService → Geocoding & Directions
- - Leaflet.js → Interactive Maps
- - Flask (Python) → Backend Server
- - FFMPEG → Audio conversion

System Architecture

- **Flow**:
- **User Voice** → Whisper → Destination Text
- **OpenRouteService** → Coordinates & Steps
- **Flask** → Sends route → Frontend
- **Leaflet.js** → Shows Map
- **TensorFlow.js** (Camera) → Detects Objects
- **SpeechSynthesis** → Speaks Directions & Warnings



Implementation

- **Backend (Flask):**
- - Handles Whisper transcription
- - Gets routes from ORS API

- **Frontend (JS):**
- - Records audio & sends to Flask
- - Shows navigation steps & map
- - Runs TensorFlow.js for object detection

Test Cases & Validation Testing

Voice Recognition Testing:

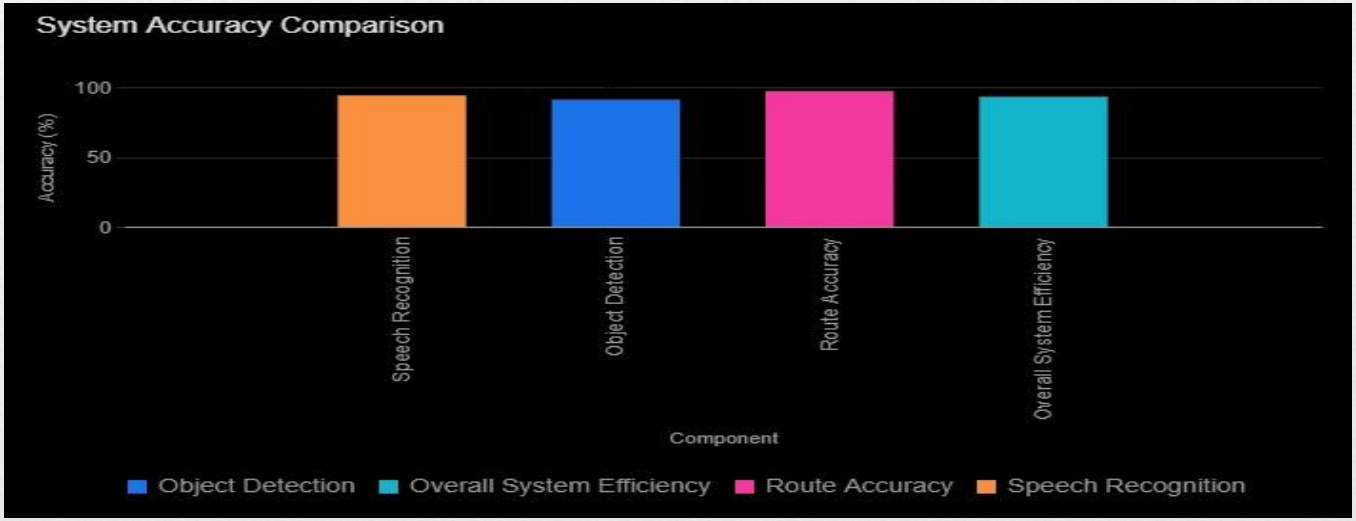
Test Case	Input	Expected output	Result
TC01	User says "Stanley Hospital Chennai"	Text recognized as "Stanley Hospital Chennai"	✅ Passed
TC02	User speaks softly or unclearly	System retries and improves recognition	✅ Passed
TC03	Background noise present	Minor delay but destination recognized correctly	⚠️ Partially Passed

Object Detection Validation:

Test Cases	Object	Expected Output	Result
TCO1	Person	“Person detected” voice alert	✅ Passed
TC02	Mobile Phone	"Mobile Phone ahead" voice alert	✅ Passed
TC03	No object	No alert generated	✅ Passed

Performance Analysis

- ✓ The **Overall System Efficiency is approximately 93%,** which shows that the entire system works together very reliably and effectively.
- ✓ The strongest individual component is **Route Accuracy at nearly 97%**. This means the system is extremely good at providing correct navigation and directions.
- ✓ **Speech Recognition is also very high (around 94%),** indicating the system accurately understands what the user is saying most of the time.
- ✓ **Object Detection (around 91%)** is the component with the most room for improvement.



Demo Screenshots

Blind Navigation & Object Detection Demo

Record Destination

Stanley Government Hospital, Chennai, Tamil Nadu

13.107023, 80.285266 (Government Stanley Hospital Blood Bank, Chennai, TN, India)

Start Navigation

Object detection model loaded!
Uploading audio for transcription...
Transcribed: Stanley Government Hospital, Chennai, Tamil Nadu
Destination resolved: Government Stanley Hospital Blood Bank, Chennai, TN, India
Starting navigation...
Step 1: Head southeast | Distance: 17117.2m
Step 1: Head southeast | Distance: 17117.2m

Front Camera Object Detection

Mini Map

Conclusion

- Combined AI speech + navigation + vision in one app.
- Useful for smart navigation systems.
- Demonstrates real-world AI integration.

Future Enhancements

- ✓ Expands the application's accessibility to a global audience by allowing users to navigate in their native language.
- ✓ Ensures the app remains reliable and functional in areas with poor or no internet connectivity by pre-downloading map data.
- ✓ Provides faster and more accurate real-time warnings for a wider range of hazards by using a state-of-the-art detection model.
- ✓ Delivers a smoother, more responsive user experience and enables background navigation by using the phone's native hardware directly.

References

- OpenAI. (2022). *Whisper: Robust Speech Recognition via Deep Learning*. Retrieved from https://github.com/openai/whisper
- TensorFlow.js. (2023). *COCO-SSD: Real-Time Object Detection in the Browser*. Retrieved from https://www.tensorflow.org/js/models
- Flask. (2023). *Flask Web Framework Documentation*. Retrieved from https://flask.palletsprojects.com
- OpenRouteService. (2023). *OpenRouteService API Documentation*. Retrieved from https://openrouteservice.org/dev/#/api-docs
- Mozilla Developer Network (MDN). (2024). *SpeechSynthesis API – Web Speech API Documentation*. Retrieved from https://developer.mozilla.org/
- Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). *You Only Look Once: Unified, Real-Time Object Detection*. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR).
- Kumar, R., et al. (2021). *Integration of Speech and Vision Systems for Assistive Navigation*. International Journal of Artificial Intelligence Research.