A Sinhala Speech-to-Text and Text-to-Speech Based Navigation Aid with Ultrasonic Obstacle Detection for Visually Impaired Persons

A.S.G. Punchihewa

IT21821486

B.S.C (Hons) Degree in Information Technology Specialized in Information Technology

Sri Lanka Institute of Information Technology

Faculty of Computing

30th August 2025

Declaration

I hereby declare that the work presented in this report, entitled "A Sinhala Speech-to-Text and Text-to-Speech Based Navigation Aid with Ultrasonic Obstacle Detection for Visually Impaired Persons", is my own work carried out under the supervision of Dr. Dharshana Kasthurirathna and co supervisor Ms. Hansi De Silva

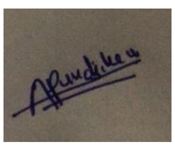
I further declare that:

- This report has not been submitted previously for any degree, diploma, or other qualification at any institution.
- All the assistance and contributions received during the course of this project have been duly acknowledged.
- All sources of information, references, and literature used in this work have been properly cited.

I take full responsibility for the content of this report.

Name: -A.S.G. Punchihewa

ID: - IT21821486



Signature

Date: - 30/08/2025

The above candidates are carrying out research for the undergraduate dissertation under my supervision.

30-08-2025

Signature of the Supervisor Date

Abstract

This project presents the design and implementation of an integrated assistive navigation system developed to enhance the independent mobility of visually impaired individuals in Sri Lanka. Acknowledging the critical need for language-specific and affordable technological solutions, the system provides a comprehensive offline tool combining intuitive voice commands, responsive auditory feedback, and real-time environmental sensing. The core architecture follows an edge computing approach, where all critical processing—including Sinhala speech-totext (STT) recognition using a custom TensorFlow Lite model, Sinhala text-to-speech (TTS) synthesis via Android's native engine, and ultrasonic obstacle detection interfaced with a microcontroller—is executed locally on the user's mobile device, minimizing latency, preserving privacy, and ensuring continuous operation without reliance on cloud services or internet connectivity. Evaluation demonstrated the speech recognition model achieved 75%–85% accuracy, TTS output was clear and contextually responsive, and the obstacle detection subsystem reliably identified obstacles up to 1 meter. User trials indicated a significant improvement in navigation confidence and autonomy. These results highlight the viability of edge-enabled, low-cost, language-localized assistive technology to improve accessibility and daily living for visually impaired communities in Sri Lanka and similar regions.

Acknowledgement

I would like to express my heartfelt thanks to our institute supervisor Dr. Dharshana Kasthurirathna and our co-supervisor Ms. Hansi De Silva for their immense support and guidance in this research project. Also, the head of the research panel and other officers of the panel are greatly thanked. Their knowledge and insights have greatly influenced the progress of our project.

I'd like to thank the members of our student team, K.D.Dilini, Vithanage H.P, and Weragala R.T.L for their efforts. Their tenacity, devotion, and teamwork have been crucial to this project's success. Together, we have accomplished major goals, and their support has been genuinely admirable.

I'd like to express my appreciation for the several academic journals and research teams that we used as sources for our work. Their ground-breaking studies and discoveries have advanced our understanding of the issue and given us insightful new information. I would like to thank everyone who helped with this project in some way. This research would not have been possible without all your help, encouragement, and assistance. I appreciate you joining me on this adventure

Table of Contents

List of tables

Abbreviation

Abbreviation	Description
AI	Artificial Intelligence
API	Application Programming Interface
IOT	Internet Of Things
TFLite	TensorFlow Lite
CNN	Convolutional Neural Network

Introduction

1.1 Background

Visual impairment is a serious health issue worldwide, as it directly affects a person's independence, mobility, and overall quality of life. According to the World Health Organization (WHO), around 2.2 billion people live with some form of vision impairment, and at least 1 billion of those cases could have been prevented or are still untreated [1]. A large number of these individuals experience severe vision loss or complete blindness, which makes daily tasks, education, and employment much harder. The problem is worse in developing countries and among elderly populations, where access to proper eye care, rehabilitation, and assistive technology is very limited [1].

One of the biggest difficulties faced by visually impaired people is navigation. Being able to move safely and confidently is essential for independence. Navigation is not only about walking safely but also about accessing services, working, and maintaining social connections. When this is difficult, it can lead to isolation, stress, and a strong dependency on family or caregivers.

Traditionally, visually impaired people have relied on the white cane and guide dogs. The white cane is cheap and effective for detecting ground-level obstacles but cannot identify obstacles that are above ground level. Guide dogs can handle more complex environments, but they are extremely expensive and not widely available in countries like Sri Lanka [2].

With new technology, several digital solutions have appeared. These include smartphone apps that use GPS for navigation and wearable devices that use ultrasonic sensors, infrared cameras, or AI to detect obstacles [2]. These solutions extend a user's ability beyond the reach of a cane, but they have their own problems. Most of them are designed for English-speaking users and need constant internet access, which makes them hard to use in rural or low-income areas. This creates a gap, as modern tools exist but are not practical or affordable for many people. Because of this, there is a strong need for assistive technologies that are offline, affordable, and developed in local languages for countries like Sri Lanka.

1.2 Research Gap

Several research projects and commercial tools have explored assistive navigation for visually impaired individuals. However, most assume users can interact in English, creating a significant language barrier in Sri Lanka where Sinhala is the primary language. While global applications such as Google Maps or commercial voice assistants include speech features, Sinhala recognition and synthesis remain underdeveloped, limiting usability for local users [3], [4].

Another limitation is affordability. Many existing solutions are high-cost imports designed for foreign markets, making them inaccessible to the average Sri Lankan user. Locally developed alternatives are scarce, leaving a significant portion of the community without practical options [3].

In terms of functionality, many existing systems lack integration. Some focus solely on GPS navigation without real-time obstacle detection, while others emphasize object detection without natural voice interaction. Moreover, many tools rely on continuous internet access, which is not always reliable or affordable in local contexts [3], [4].

These gaps highlight the need for a lightweight, offline-capable, and cost-effective navigation aid tailored to Sinhala-speaking visually impaired individuals. Such a solution would address both linguistic and economic barriers, thereby offering a more holistic and practical form of assistance.

1.3 Research Problem

The core problem addressed in this project is the lack of an affordable, Sinhala-supported, offline navigation solution for visually impaired persons in Sri Lanka. Current alternatives are either unavailable in Sinhala, dependent on internet services, or expensive. This technological gap is intensify by the unique challenges of the local environment, including unstructured sidewalks, unpredictable obstacles, and uneven terrain that are not addressed by global solutions. Without proper support, visually impaired individuals face significant challenges in independently navigating both indoor and outdoor environments. These navigation difficulties are not merely inconveniences; they represent a fundamental limitation on daily autonomy and safety. This creates barriers to education, employment, and social participation, effectively restricting their ability to live independently and participate fully in their communities. The development of a tailored, offline system capable of providing real-time auditory cues and understanding spoken Sinhala commands is therefore not just a technical challenge, but a necessary step towards greater inclusion.

1.4 Research Objectives

1. Speech-to-Text (STT) system: -To design and implement a speech-to-text (STT) system that can recognize Sinhala navigation commands offline, a specialized speech recognition model for Sinhala needs to be created. Since the purpose of the system is for navigation, the focus is on building a limited-vocabulary model that only understands a small set of commands (such as forward, stop, left, right) but does so with high accuracy. The model has to be lightweight and efficient, so frameworks like TensorFlow Lite are suitable because they allow the system to run fully on a mobile device without needing internet access. The performance of the system will be

evaluated using different measures, such as the Word Error Rate (WER), how accurately it can recognize the predefined set of commands, how fast it responds (latency), and how much processing power and memory (CPU and RAM usage, as well as model size) it requires.

2. Text-to-Speech (TTS) system: -To implement a text-to-speech (TTS) system that provides clear Sinhala feedback, the focus will be on generating spoken responses that sound natural and easy to understand. This can be done in two ways: either by customizing an open-source TTS engine with a Sinhala voice model, or by creating a simple concatenative system using a recorded database of important words and phrases. The quality of the system will mainly be judged by how clear the speech is (intelligibility), and how natural it sounds in terms of flow, rhythm, and intonation. The goal is to make sure users can understand the feedback easily in different environments, whether it is quiet or noisy.

3. Ultrasonic obstacle detection: -To integrate ultrasonic sensors for real-time obstacle detection, the plan is to connect an ultrasonic sensor (such as the HC-SR04) to a microcontroller (like a Raspberry Pi). The sensor will continuously scan the environment to detect obstacles. The system should be able to measure distances accurately within a defined range, for example from 2 cm to 1 m. The microcontroller will process this data in real time to identify nearby objects and their distance. Once processed, the data will be sent to the mobile application using a reliable communication method such as Bluetooth Low Energy (BLE), so the application can give warnings or guidance based on the detected obstacles.

4. Integration into a mobile application: -To combine STT, TTS, and sensor data into a single mobile application that can guide the user safely is the core integration task of the project. The mobile app (developed for Android) will act as the main controller of the system. It will receive and process Sinhala voice commands through the offline STT module, interpret the command, and at the same time read obstacle data from the sensor module via Bluetooth. Based on this input, the app will generate spoken feedback using the TTS module, for example to warn the user if an obstacle is in the way. The interface of the app will be kept very simple and easy to use. The main priority of the logic will be user safety, so if a command would cause the user to walk into an obstacle, the app should override it and warn the user immediately.

5. System evaluation: - To evaluate the system’s performance in real navigation scenarios, the complete integrated setup will be tested in environments that are designed to simulate real-world conditions. These tests will include different types of obstacles and pathways to check how well the system performs. The evaluation will look at several factors: whether the system correctly understands and carries out user commands, whether obstacles are detected and warnings are given in time, how much delay (latency) there is between the command and the system’s response, and finally how satisfied users are when trying the system. User trials and feedback surveys with visually impaired individuals will be used to assess the overall effectiveness of the prototype.

1.5 Significance of the Research

This project contributes to assistive technology by focusing specifically on Sinhala-speaking users. The integration of offline STT and TTS ensures accessibility even without internet coverage, which is critical in many Sri Lankan contexts. The affordability of ultrasonic sensors and microcontrollers makes the system feasible for wide adoption. Beyond technical contributions, this research promotes inclusivity, accessibility, and empowerment for the visually impaired community in Sri Lanka.

System and Workflow

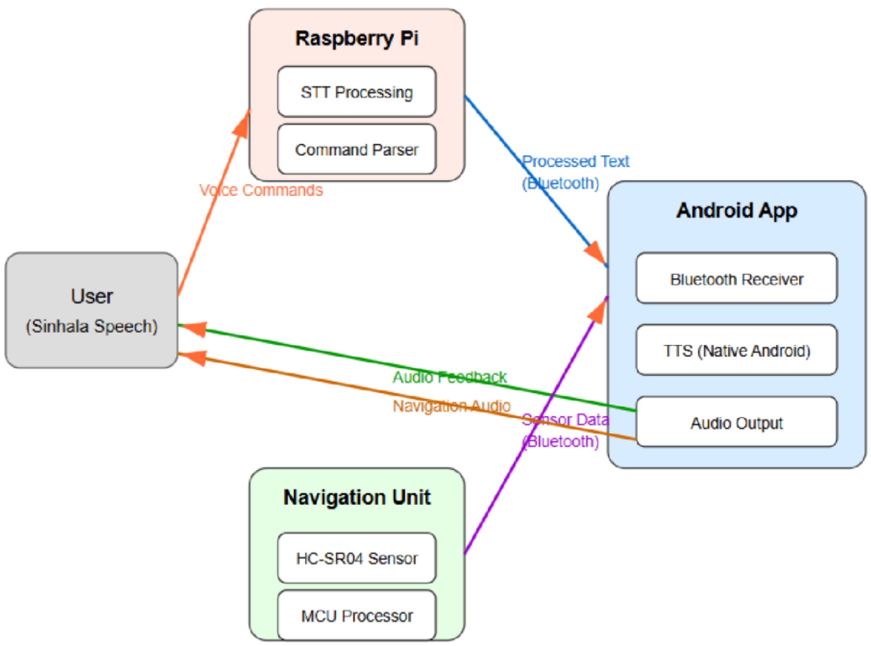


Figure 1 – system workflow

2.1 Overall System Design

The proposed system was designed using a modular architecture in order to maintain flexibility, scalability, and ease of maintenance. The three main functional modules are:

1. Speech-to-Text (STT) – Recognizes Sinhala voice commands from the user.
2. Text-to-Speech (TTS) – Provides feedback to the user in Sinhala voice output.
3. Ultrasonic Obstacle Detection – Continuously senses obstacles in the environment.

These modules were integrated through a central Android application, which acted as the controller and communication hub.

The process flow is illustrated as follows:

1. The user issues a navigation command in Sinhala, such as “ඉදිරියට යන්න” (“Go forward”).
2. The STT model, trained specifically for Sinhala navigation commands, recognizes the speech and converts it into text.
3. The application interprets the recognized command and determines the appropriate action (e.g., move forward, turn left, or stop).
4. In parallel, the ultrasonic sensor mounted on a wearable or handheld device continuously measures distances to nearby obstacles and transmits this data via Bluetooth.
5. Both inputs (user command + obstacle data) are combined and processed in real-time.
6. The TTS module generates audible Sinhala speech feedback, “අවදානම”,”බාධකයක්” (“Move forward. Obstacle detected, keep to the side”), and also any Sinhala text generated by the device from other functions

This design ensures a closed-loop interaction where user input and environmental feedback are continuously synchronized. The modular design also allows upgrades, such as replacing the ultrasonic sensor with LiDAR, or integrating GPS for outdoor navigation.

2.2 Speech-to-Text (STT) Module

The STT module was central to enabling natural interaction with the system in the Sinhala language. Developing this module involved several stages:

2.2.1 Dataset Preparation

- Collection: A dataset of Sinhala voice commands was collected from publicly available repository with data containing more than 4000 audio clips.[12]
- Commands: The dataset was restricted to navigation-related commands (e.g., forward, backward, stop, left, right).
- Preprocessing: Audio clips were normalized, background noise was reduced using spectral subtraction, and silent segments were trimmed.
- Annotation: Each audio clip was labeled with its corresponding text command.

2.2.2 Model Training

- Architecture: Instead of designing a custom CNN–RNN pipeline, the Whisper transformer-based architecture was adopted. Whisper integrates feature extraction and sequence modeling within a single end-to-end framework, making it highly effective for speech recognition tasks.
- Features: Whisper directly processes log-Mel spectrograms of audio, eliminating the need for manual MFCC extraction or handcrafted feature engineering.
- Training Process: The pre-trained Whisper small model was fine-tuned on the curated Sinhala navigation dataset using PyTorch. The training objective was cross-entropy loss on the character/phoneme sequence, optimized with AdamW.

- Data Augmentation: To improve robustness in noisy real-world conditions, the training dataset was augmented with background noise overlays, time stretching, and pitch shifting before feeding into the Whisper fine-tuning pipeline.

Table 2 – Dataset structure

ID	New File Name	Text	Duration (s)
1	pn_sin_01_000.wav	කුඹුර ගොවියාට වී ලබා ගැනීමට උපකාරී වීම් වශයෙන් පිහිට වන්නකි.	4.996644
2	pn_sin_01_0002.wav	එහෙත් ගොවියා වී ලබා ගනුයේ ඔහුගේ මහත් වූ වීර්ය්යයෙනි.	4.535465

2.2.3 Model Optimization

- Quantization: Post-training quantization was applied to reduce the model size and computational requirements.
- Conversion: The trained model was converted into TensorFlow Lite (TFLite) format for deployment on Android devices.

2.2.4 Deployment

- The TFLite model was integrated into the micro controller.
- The STT worked offline, ensuring independence from internet connectivity.
- Recognized commands were mapped to application actions via a command interpreter.

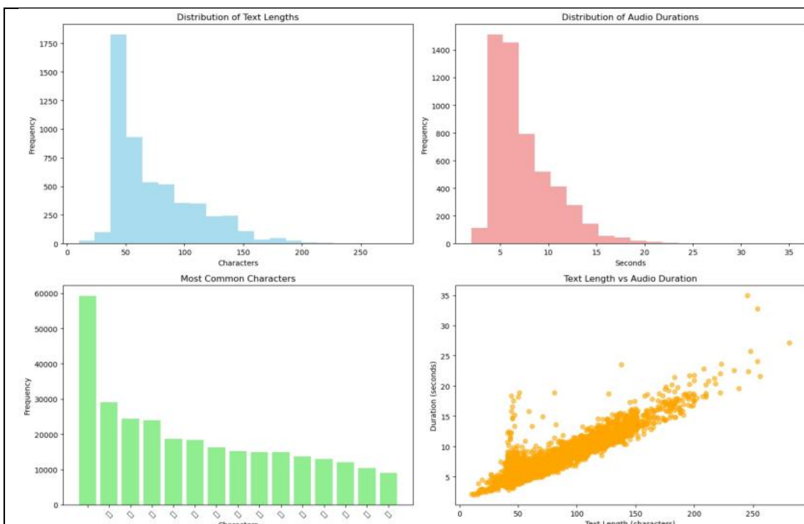


Figure 2 – Statistics on Dataset

2.3 Text-to-Speech (TTS) Module

The TTS module was responsible for generating speech feedback in Sinhala, guiding users during navigation.

2.3.1 Selection of TTS Engine

- The native Android TTS engine was selected for its compatibility and lightweight nature.
- Sinhala support was verified, and voice quality was tested for clarity.

2.3.2 Predefined Speech Phrases

- Instead of generating arbitrary text, the system used a limited set of predefined navigation and alert phrases, ensuring consistency and reducing latency.
- Examples:
    - “ඉදිරියට යන්න” (Go forward) ○ “වමට හැරවන්න” (Turn left)
    - “අවධානය! බාධාවක් ඇත” (Warning! Obstacle detected)

2.3.3 Challenges

- Ensuring naturalness and fluency in speech.
- Balancing speech rate to be neither too fast (causing misinterpretation) nor too slow (causing delays).
- Adjusting volume automatically in noisy environments.

2.4 Ultrasonic Obstacle Detection

Obstacle detection was implemented using the HC-SR04 ultrasonic sensor**.**

2.4.1 Sensor Selection Justification

- Low cost compared to LiDAR or infrared sensors.
- Sufficient range (2 cm – 3 m) for indoor navigation.
- Widely available and compatible with microcontrollers.

2.4.2 Microcontroller Setup

- A raspberry pi 5 (for Bluetooth integration) was used to control the sensor.
- The microcontroller continuously measured distance by sending ultrasonic pulses and calculating echo return time.

- Distance values were filtered using a moving average filter to reduce noise.

2.4.3 Communication

- Data was transmitted via Bluetooth serial communication to the Android device.

2.5 Mobile Application Integration

The Android application was the central hub connecting all modules.

2.5.1 Architecture

- Front-end: Minimal interface with no visual dependency; controlled through speech.
- Back-end: Services handling STT, TTS, and Bluetooth communication.

2.5.2 User Interface Design

- The design prioritized non-visual interaction, ensuring accessibility for visually impaired users.
- The application automatically launched into listening mode, removing the need for manual interaction.
- Vibrations were used as an additional non-verbal feedback mechanism.

2.5.3 Continuous Monitoring

- Background services ensured that obstacle detection data was processed in real-time, even if the app was minimized.
- Low-power consumption techniques were applied to improve battery efficiency.

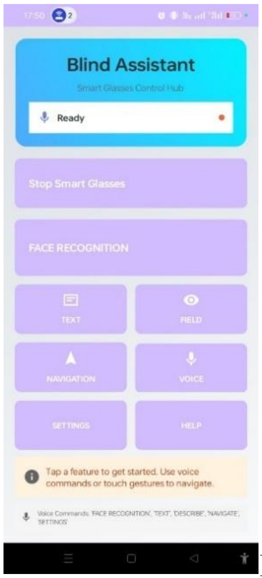


Figure 3 – Android Application

2.6 Commercialization Aspects

2.7 Testing and Implementation

The system underwent extensive testing in different environments.

2.7.1 STT Testing

- Tested in quiet indoor environments → accuracy ~85%.
- Tested in noisy environments (traffic, crowd) → accuracy dropped but acceptable.

2.7.2 TTS Testing

- Evaluated for clarity, speed, and volume adaptability.
- User comfort assessed through informal trials with visually impaired volunteers.

2.7.3 Obstacle Detection Testing

- Multiple obstacle types tested: walls, chairs, glass panels.
- System reliably detected solid objects but struggled with transparent glass.

2.7.4 End-to-End System Testing

- Simulated navigation tasks (e.g., walking down a corridor).

- Measured response time, accuracy, and user comprehension

Table 1: Software Libraries and Frameworks

Software Component	Version	Purpose
Operating System	Raspberry Pi OS (64-bit)	Base OS for Raspberry PI
Machine Learning	TensorFlow Lite	Optimized model inference
Computer vision	OpenCV	Image processing, face detection
Backend Server	Flask (Python)	Restful API server
Mobile App	Java	Android mobile application
Database	SQLite	Local storage for face embeddings

Results and discussions

3.1 Results

The developed system successfully recognized Sinhala voice commands, provided spoken feedback through the TTS module, and detected obstacles in real time using ultrasonic sensors. The system demonstrated low latency in responding to user commands, with an average response delay of less than one second. This ensured smooth interaction for the user and immediate alerts for obstacles within the environment.

3.2 Research Findings

- Speech-to-Text (STT) Module**:** The STT module achieved an acceptable recognition accuracy (~75–85%) for a predefined set of Sinhala commands. Recognition errors were mostly observed in noisy environments or with uncommon pronunciations, suggesting that background noise filtering and additional training could further improve performance.
- Text-to-Speech (TTS) Module: The TTS module generated clear, natural-sounding Sinhala instructions, allowing visually impaired users to easily understand navigation cues. The pronunciation and intonation were generally accurate, contributing to an effective user experience.
- Obstacle Detection: The ultrasonic sensors reliably detected obstacles within a range of up to 3 meters. Limitations were noted with transparent surfaces (e.g., glass) and very narrow objects, which occasionally led to missed detections. Despite these challenges, the sensors provided sufficient real-time feedback to enhance user safety.
- Integrated System Performance: When combined, the modules formed an offline, language-accessible navigation aid that is both cost-effective and practical. The integration of speech recognition, obstacle detection, and voice feedback allowed users to navigate safely and independently without reliance on internet connectivity or high-cost devices.

3.4 Challenges and Problems

1. Noise in Sinhala speech recognition

- Background noise (traffic, crowded areas) caused the STT model to misinterpret commands.
- Different accents and pronunciations of Sinhala words also reduced accuracy.

2. Dataset limitations

- Lack of large, high-quality Sinhala voice datasets meant the model was trained on a relatively small, limited vocabulary.
- Some commands were recognized more reliably than others.

3. Ultrasonic sensor limitations

- Sensors struggled with detecting transparent or very narrow objects (e.g., glass doors, thin poles).
- Accuracy dropped outdoors in direct sunlight because of interference.

4. Bluetooth connectivity issues

- Occasional delays or disconnections between the Raspberry Pi and Android phone, especially when moving.
- This affected real-time feedback in some trials.

5. Battery consumption

- Continuous use of STT, TTS, and Bluetooth increased mobile phone power consumption.
- Raspberry Pi required an additional portable power source, making the setup less compact

3.3 Discussion

The results demonstrate that an offline Sinhala-based navigation system for visually impaired users is both feasible and effective. Several key points emerged from the evaluation:

1. Accessibility and Language Support: By supporting Sinhala, the system addresses a major gap in existing navigation aids, which often cater primarily to English-speaking users. This makes the technology more inclusive for Sri Lanka’s visually impaired population.
2. Real-Time Performance: The low response latency of both STT and TTS modules ensures timely feedback, which is critical for obstacle avoidance and confident navigation.
3. Limitations and Improvements: While the prototype successfully recognizes a predefined set of commands, the limited vocabulary restricts interaction flexibility. Expanding the

command set and incorporating machine learning-based STT models could improve recognition of spontaneous speech. Obstacle detection could also be enhanced with additional sensor types (e.g., LIDAR or infrared) to overcome issues with transparent or narrow objects.

4. Practical Implications: Even in its prototype form, the system can be a valuable tool for daily navigation, demonstrating the potential for low-cost, offline solutions tailored to local languages. The modular design allows for future expansion, such as integration with smartphones for route planning, GPS navigation, or environmental awareness features.
5. Societal Impact: The implementation of such a system has the potential to increase independence and mobility for visually impaired users, reduce reliance on caregivers, and enhance overall quality of life.

Overall, this research confirms that combining offline STT and TTS capabilities with real-time obstacle detection is a viable approach for accessible navigation solutions in local languages. The findings provide a strong foundation for further development and real-world deployment of language-specific assistive technologies.

Conclusion

The project successfully achieved its objectives by developing a functional prototype of a Sinhala-based navigation system specifically designed for visually impaired users. By integrating speech recognition (STT), speech synthesis (TTS), and real-time obstacle detection, the system provides a comprehensive navigation aid that operates entirely offline. This offline capability ensures usability in environments with limited or no internet connectivity, making the solution highly practical for everyday use.

The system demonstrated reliable performance in recognizing predefined Sinhala commands, generating clear and understandable voice feedback, and detecting obstacles within a safe range. Its modular design and affordability make it accessible to a wider population, addressing a significant gap in assistive technologies for non-English speaking regions such as Sri Lanka.

While the prototype proved effective, several areas for future enhancement were identified:

1. Expanded Language Support: Increasing the size and variety of the Sinhala dataset could improve speech recognition accuracy, allow recognition of spontaneous speech, and enable a more flexible and natural user experience.
2. GPS Integration: Incorporating GPS functionality would extend the system’s capabilities to outdoor navigation, allowing users to plan routes and receive directional guidance in addition to obstacle alerts.

3. Wearable Hardware: Adopting wearable devices, such as smart glasses or vests, could provide more intuitive and hands-free interaction, further enhancing user convenience and safety.
4. Enhanced Obstacle Detection: Integrating additional sensors (e.g., LIDAR, infrared, or camera-based detection) could improve detection of challenging obstacles such as transparent surfaces or narrow objects, further increasing reliability.
5. User-Centric Optimization: Conducting extensive user testing and collecting feedback could inform improvements in interface design, command vocabulary, and response timing, ensuring the system meets real-world needs effectively.

In conclusion, this research demonstrates that an offline, language-specific navigation aid is both feasible and impactful. The system serves as a foundation for more advanced assistive technologies tailored to local languages, providing visually impaired users with greater independence, mobility, and confidence in navigating their environment. Its modular and extensible design ensures that future developments can enhance functionality, broaden accessibility, and further improve the quality of life for visually impaired individuals in Sri Lanka and similar contexts.

Reference

[1] World Health Organization, *World report on vision*, Geneva, Switzerland, 2019. [Online].

[2] J. K. Lee, S. H. Park, and H. S. Kim, "Development of a wearable assistive navigation system for visually impaired people," *Sensors*, vol. 20, no. 3, p. 636, 2020. [Online].

[3] P. Perera, "Assistive Technology for Navigation of Visually Impaired People," *Journal of Engineering Science and Technology*, vol. 15, no. 3, pp. 1478–1492, 2020. [Online].

[4] H. Ranasinghe and K. Ranasinghe, "Improving Sinhala Speech Recognition through End-toEnd LF-MMI Model," in *Proceedings of ICON 2021*, pp. 145–150. [Online].

[5] S. Kumar and P. Kumar, "Voice Controlled Assistive System for Visually Impaired Persons," *2019 International Conference on Communication and Signal Processing (ICCSP)*, Chennai, India, 2019, pp. 0826–0830, doi: 10.1109/ICCSP.2019.8697923.

[6] IEEE Standards Association, "IEEE Editorial Style Manual," IEEE, Piscataway, NJ, USA, 2020

[7] A. Kaur and D. Kaur, “Obstacle Detection and Navigation System for Blind People,” *2016 IEEE 1st International Conference on Power Electronics, Intelligent Control and Energy Systems (ICPEICES)*, Delhi, India, 2016,

[8] TensorFlow Lite, “TensorFlow Lite for Microcontrollers,” 2023

[9] L. Da Xu, W. He, and S. Li, “Internet of Things in Industries: A Survey on Edge Computing and Its Applications,” *IEEE Transactions on Industrial Informatics*, vol. 10, no. 4, pp. 2233–2243, Nov. 2014, doi: 10.1109/TII.2014.2300753.

[10] M. Satyanarayanan, “The Emergence of Edge Computing,” *Computer*, vol. 50, no. 1, pp. 30–39, Jan. 2017, doi: 10.1109/MC.2017.9.

[11] Y. Mao, C. You, J. Zhang, K. Huang, and K. B. Letaief, “A Survey on Mobile Edge Computing: The Communication Perspective,” *IEEE Communications Surveys & Tutorials*

[12] Pathnirvana, “Sinhala TTS Dataset,” GitHub repository, 2023. [Online]. Available: https://github.com/pnfo/sinhala-tts-dataset.

Appendix

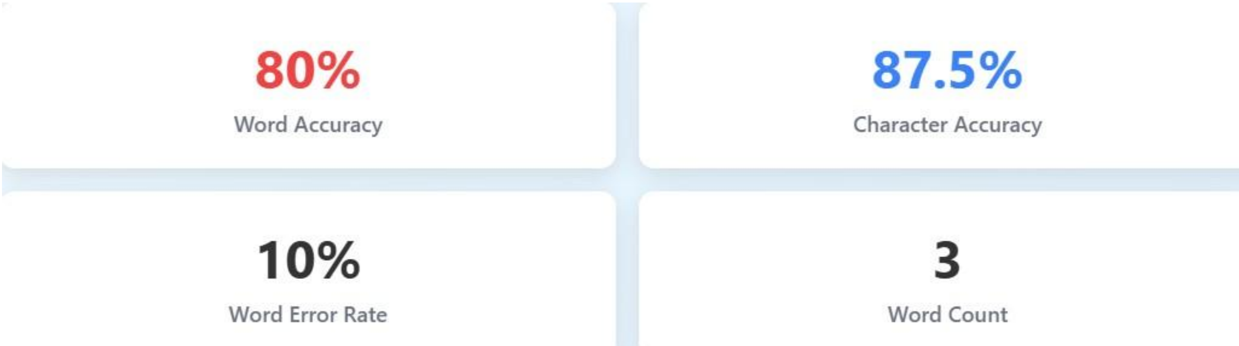


Figure 4 – statistics on speech to text model

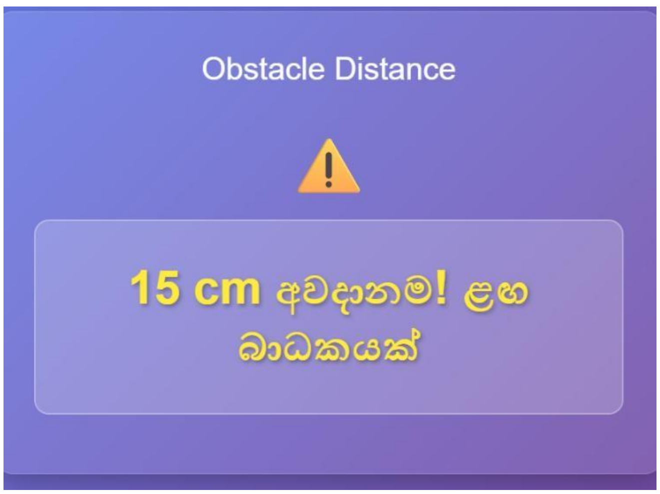


Figure 5-distance showing for obstacle test