Sinhala Text Recognition and Voice Feedback for Visually Impaired - Optical Character Recognition (OCR)

Project ID: R25-012

Project Proposal Report

Dilini K.D. | IT21826740

Supervisor: Dr. Dharshana Kasthurirathna

Co-Supervisor: Ms. Hansi De Silva

B.Sc. (Hons) Degree in Information Technology Specialization in

Information Technology

Department of Information Technology

Sri Lanka Institute of Information Technology

Sri Lanka

January 2025

DECLARATION

We declare that this is our own work, and this proposal does not incorporate without acknowledgement any material previously submitted for a degree or diploma in any other university or Institute of higher learning and to the best of our knowledge and belief it does not contain any material previously published or written by another person except where the acknowledgement is made in the text.

Name	Student Id	Signature
Dilini K.D.	IT21826740	

	2025/02/03
Signature of the Supervisor	Date

	2025/02/03
Signature of the Co-Supervisor	Date

ABSTRACT

This research aims to develop an affordable and effective assistive platform for visually impaired individuals in Sri Lanka, enabling them to capture and read Sinhala text using Tiny ML and Edge AI technologies. The primary objective is to address the lack of accessible, real-time text recognition systems tailored for Sinhala, as existing solutions fail to meet the specific needs of the visually impaired community, particularly in resource-constrained environments. This study proposes a system involving smart glasses and a mobile application that integrates Optical Character Recognition (OCR) and text-to-speech (TTS) technologies to provide real-time, hands-free reading.

The smart glasses will be equipped with a Raspberry Pi camera, which will solely be used for image capture. The mobile application will process the captured text using AI-powered OCR and convert it into synthesized speech, providing immediate auditory feedback to the user. A key feature of this solution is real-time voice guidance, which assists users in positioning the text correctly within the camera's frame, ensuring effective text capture while minimizing errors in recognition. Interactive audio cues will further enhance accessibility by guiding users through the application interface. By leveraging Tiny ML and Edge AI, the system will function without reliance on cloud computing, ensuring privacy, low latency, and cost efficiency.

The proposed solution will utilize low-cost hardware components and lightweight AI models optimized for mobile processing, making it both affordable and accessible. Unlike traditional assistive technologies that rely on internet connectivity or expensive proprietary systems, this approach focuses on real-time, offline processing, allowing users to read printed text without external dependencies. The mobile application will be developed using frameworks like React Native, integrating Google ML Kit or Tesseract OCR for text recognition, and employing text-to-speech synthesis for clear and natural voice output.

Through a user-centered design approach, usability testing will be conducted with visually impaired individuals to evaluate the system's effectiveness, accessibility, and ease of use. Key performance indicators will include text recognition accuracy, response time, user satisfaction, and overall impact on daily reading experiences. The expected outcome is a cost-effective, user-friendly assistive device that significantly enhances the independence and quality of life for visually impaired individuals in Sri Lanka by making printed materials more accessible and improving access to information.

Ultimately, this research aims to contribute to the development of localized assistive technologies in Sri Lanka and provide a scalable model for similar environments worldwide. By addressing the unique challenges faced by Sinhala-speaking visually impaired users, this system has the potential to bridge the accessibility gap and empower individuals with greater autonomy in education, work, and daily life.

Table of Contents

1. INTRODUCTION

1.1. Background & Literature Survey

Assistive technology has played a significant role in improving accessibility for visually impaired and blind individuals worldwide. Optical Character Recognition (OCR) is one of the most important technologies in this domain, enabling the conversion of printed and handwritten text into machine-readable formats [1]. This has led to the development of various reading aids, such as text-to-speech systems and Braille conversion tools, which help visually impaired users access printed content with greater independence.

Early OCR systems focused on converting scanned documents into editable text for industrial and academic use. The Kurzweil Reading Machine (1976) was a pioneering assistive device that introduced text-to-speech technology, allowing visually impaired individuals to listen to printed text [2]. In recent years, AI-driven solutions such as Microsoft's Seeing AI and OrCam MyEye have provided real-time text recognition, allowing users to capture and interpret printed text instantly through wearable and smartphone-based systems [3].

In Sri Lanka, research into OCR for the Sinhala language has gained traction due to the unique challenges posed by its script. The University of Colombo School of Computing developed an OCR system for Sinhala in 2008 to enhance digital text accessibility [4]. Similarly, in 2019, Sabaragamuwa University introduced the Optical Braille Recognition Software Prototype for Sinhala, aimed at recognizing and converting Sinhala Braille text into readable formats [5]. Despite these advancements, there is still a lack of affordable, real-time, wearable assistive devices tailored to the needs of Sinhala-speaking visually impaired individuals.

OCR systems designed for Western languages often struggle with the complexities of non-Latin scripts like Sinhala. The Sinhala script consists of a large number of character ligatures, making accurate text recognition a challenging task [6]. Furthermore, most existing Sinhala OCR solutions are designed for desktop or cloud-based applications, which are not suitable for real-time use in mobile or wearable devices.

Recent advancements in Tiny ML and Edge AI have revolutionized real-time AI applications by enabling on-device processing without reliance on cloud computing [7]. These technologies have made it possible to develop lightweight, efficient, and low-cost AI models that can run on embedded devices like Raspberry Pi and mobile processors, making assistive technology more accessible [8].

However, existing solutions still present limitations, including high costs, dependence on cloud services, and limited customization for regional languages. This research aims to bridge these gaps by developing an affordable, portable smart glass integrated with a mobile-based Sinhala book reader. The proposed system will incorporate real-time OCR, Edge AI, and text-to-speech conversion, offering a localized solution that empowers visually impaired individuals in Sri Lanka to access printed content independently.

By leveraging machine learning algorithms and advanced image processing techniques, the OCR model in this study will be trained to recognize Sinhala characters with high accuracy. The expected outcome is a low-cost, real-time assistive platform that enhances literacy, accessibility, and independence for visually impaired individuals in Sri Lanka.

1.2. Relevant Studies or Techniques

To successfully develop this solution, mastery in the following areas is essential:

1. **Tiny ML and Edge AI:** These enable running AI models on resource-constrained devices. Tools like TensorFlow Lite and Edge Impulse are critical for model optimization.

2. **OCR Algorithms:** Knowledge of techniques such as convolutional neural networks (CNNs) for text recognition and pre-processing methods for handling complex scripts like Sinhala is required (Smith, 2007).

3. **Speech Synthesis and Natural Language Processing (NLP):** Using text-to-speech libraries like Google TTS or open-source alternatives is key to generating real-time voice feedback.

4. **Human-Computer Interaction (HCI):** Designing intuitive interfaces and seamless voice command systems tailored to visually impaired users is essential.

5. **Embedded Systems and Wearables:** Familiarity with microcontrollers (e.g., Raspberry Pi, Arduino) and integration of camera modules for real-time image capture.

By addressing the limitations of existing solutions, this project aims to provide an inclusive and scalable technology that significantly enhances the quality of life for visually impaired individuals in Sri Lanka.

2.REREARCH GAP

Assistive technology has played a transformative role in improving accessibility for visually impaired individuals, particularly through Optical Character Recognition (OCR) systems. However, despite advancements in AI-powered OCR applications, a critical research gap exists in the development of an affordable, real-time, and localized wearable OCR solution optimized for the Sinhala language. Most existing solutions either lack support for non-Latin scripts like Sinhala, do not function efficiently in real-time, or are prohibitively expensive, making them inaccessible to many users in Sri Lanka.

Early research efforts, such as the Sinhala OCR System (UCSC, 2008), focused on digitizing Sinhala text using desktop-based software. While this was a significant step in Sinhala text recognition, it was not designed for real-time use or visually impaired users. Similarly, the Mobile-Based Sinhala Book Reader (SLIIT, 2023) introduced a Sinhala text-to-speech conversion system but relied on a smartphone-based interface, requiring users to manually capture images for text recognition, which is not ideal for seamless accessibility. These projects highlight the lack of real-time, hands-free solutions for Sinhala-speaking visually impaired individuals.

In contrast, global assistive technologies such as Microsoft's Seeing AI and OrCam MyEye (2015) offer real-time OCR capabilities. Seeing AI is a mobile-based application that provides text-to-speech feedback but does not support Sinhala text recognition, making it unsuitable for Sri Lankan users. Meanwhile, OrCam MyEye is a wearable AI-powered text recognition device, offering hands-free operation. However, the high cost (around $3,500) makes it inaccessible to most visually impaired individuals in developing countries. The absence of a low-cost, Sinhala-specific, real-time OCR solution highlights a major gap in assistive technology for Sri Lanka.

Another key limitation of existing assistive devices is their dependence on cloud-based processing. Advanced OCR applications often require internet connectivity to perform complex text recognition tasks, which restricts accessibility in rural areas with limited connectivity. This issue could be addressed using Edge AI and Tiny ML, enabling real-time text recognition directly on embedded devices, reducing latency and eliminating the need for constant internet access.

To bridge these gaps, this research proposes a low-cost, real-time wearable smart glass integrated with a mobile-based Sinhala book reader. Unlike previous Sinhala OCR research, this solution will feature a custom-trained AI model for Sinhala text recognition, optimized for real-time use. By incorporating Edge AI, the system will process text locally, eliminating cloud dependencies and ensuring smooth functionality even in offline environments. Furthermore, its wearable form factor will allow hands-free operation, a significant advantage over smartphone-dependent solutions like Seeing AI and the Sinhala book reader.

Affordability is another major focus of this research. Unlike high-end wearable devices such as OrCam MyEye, which remain out of reach for many users due to their prohibitive cost, this project aims to develop a cost-effective, locally manufactured solution. By leveraging low-power hardware and optimized AI models, the system will offer an affordable alternative without compromising performance.

Additionally, the proposed research will introduce smart reading modes and voice navigation, enabling context-aware reading assistance, topic-based text extraction, and voice-guided navigation. These features will enhance the usability of the device for visually impaired individuals, providing a more interactive and user-friendly reading experience.

By addressing the technical, linguistic, and financial limitations of existing assistive solutions, this research aims to empower visually impaired individuals in Sri Lanka with a practical and sustainable smart glass system. The project will not only improve literacy and accessibility but also enhance independence by enabling visually impaired users to access printed text seamlessly in their daily lives.

Research Reference	Target Audience	Technology Used	Real-Time Processing	Supports Sinhala Language	Wearable/Portable	Smart Reading Modes	Affordability
Research A: Sinhala OCR System (UCSC, 2008)	General users, researchers	OCR, Image Processing	✗	✓	✗	✗	✗
Research B: Seeing AI (Microsoft)	Visually impaired users	AI, OCR, NLP, Cloud Computing	✓	✗	✗	✗	✓ (Free but requires internet)
Research C: OrCam MyEye (2015)	Visually impaired users	AI, Wearable OCR, Computer Vision	✓	✗	✓	✗	✗ (Very Expensive)
Research D: Mobile-Based Sinhala Book Reader (SLIIT, 2023)	Sinhala-speaking visually impaired users	Mobile OCR, TTS	✗	✓	✗	✗	✓ (Moderate)
Proposed system	Visually impaired Sinhala users	Edge AI, Tiny ML, OCR, NLP, Image Processing, Wearable Tech	✓	✓	✓	✓	✓ (Low-Cost)

3. RESEARCH PROBLEM

The main research problem is that visually impaired individuals in Sri Lanka lack access to an affordable, real-time, and wearable solution that enables them to independently read printed Sinhala text. Existing solutions, such as mobile-based OCR applications, are limited in their ability to provide seamless interaction due to the need for manual handling, dependency on cloud services, and lack of real-time feedback. High-end assistive technologies, such as OrCam MyEye, remain financially inaccessible for the majority of users in developing regions.

One of the primary challenges is the complexity of the Sinhala script. Unlike Latin-based languages, Sinhala consists of a large number of character ligatures and contextual variations, making accurate OCR recognition difficult. Many existing OCR systems struggle with multi-font and multi-size text variations, leading to errors in text conversion. Furthermore, text positioning and alignment issues often hinder the effectiveness of current OCR models, as visually impaired users may struggle to correctly position the text within the camera's capture frame.

Another significant limitation is the real-time processing capability of current assistive technologies. Most Sinhala OCR applications require internet connectivity for cloud-based processing, causing delays and making them unusable in offline environments. Visually impaired individuals need an instant audio response when reading text, but cloud-dependent solutions introduce latency and limit accessibility in low-connectivity areas.

Additionally, portability and usability pose major challenges. Mobile-based OCR apps require users to hold and adjust the smartphone camera manually, which can be difficult for visually impaired individuals. In contrast, wearable assistive devices like OrCam MyEye provide hands-free functionality but are prohibitively expensive. There is a clear need for a low-cost, lightweight, and user-friendly wearable device that enables effortless real-time text recognition and speech conversion without requiring complex interactions.

By addressing these limitations, this research aims to develop a smart glass-based Sinhala text reader, integrated with Edge AI and Tiny ML, to provide real-time text recognition and voice feedback. The proposed system will optimize OCR for the Sinhala script, implement smart reading modes, and ensure affordability through lightweight on-device processing. This solution has the potential to enhance the independence and literacy of visually impaired individuals in Sri Lanka by providing them with an accessible, wearable, and real-time assistive technology.

Research Questions:

1. What are the most effective OCR techniques for accurately recognizing Sinhala characters in real-time?
2. How can the OCR system be optimized for multi-font and multi-size Sinhala text using Edge AI?
3. What techniques can be used to provide real-time audio feedback to ensure seamless interaction for visually impaired users?
4. How can the device assist users in correctly positioning the camera for better text capture without requiring manual intervention?
5. What are the best approaches to developing a low-cost, offline-capable wearable assistive device for Sinhala text recognition?

To address these research questions, the study will utilize a combination of computer vision, deep learning-based OCR, natural language processing (NLP), and Tiny ML for real-time, on-device processing. The development process will involve training AI models with large datasets of Sinhala text, optimizing for low-power consumption and affordability, and ensuring a user-friendly wearable design. The final outcome will be a cost-effective, real-time assistive smart glass solution that significantly improves the reading accessibility of visually impaired individuals in Sri Lanka.

4. OBJECTIVES

4.1. Main Objective

The primary objective of this research is to develop a **smart glass-based Sinhala book reader** for visually impaired individuals in Sri Lanka, integrating **Optical Character Recognition (OCR), Tiny ML, and real-time Text-to-Speech (TTS) conversion**. The device will use a **Raspberry Pi camera** to capture printed Sinhala text and provide **instant audio feedback** via a mobile application, ensuring a **hands-free and real-time reading experience**. This research aims to **enhance accessibility, literacy, and independence** for visually impaired users by offering an **affordable, wearable, and AI-powered solution** that does not require internet connectivity for basic operations.

4.2. Sub Objectives

Sinhala OCR System Development

• Create localized Sinhala OCR system for accurate recognition of Sinhala characters and words.

• Enhance system for real-time operation on mobile devices.

Integration of Tiny ML and Edge AI for Real-Time Processing

• Implement Tiny ML and Edge AI models for real-time OCR and text-to-speech conversion on wearable devices.

Design of Smart Glasses with Text Capture Functionality

• Develop prototype integrating camera, microphone, and speaker for text capture and real**-time voice feedback.**

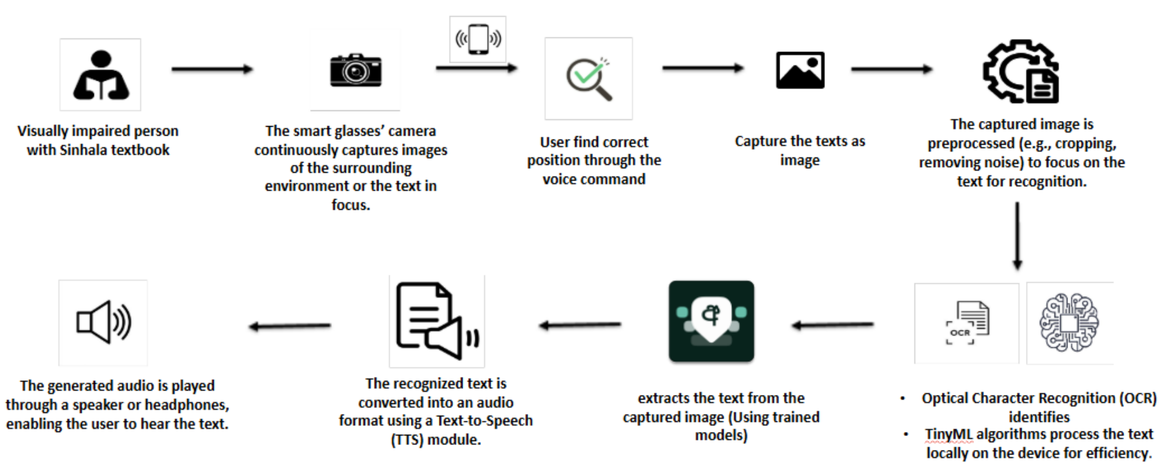
5. METHODOLOGY

5.1 System Design

This research aims to develop a smart glasses-based assistive device integrated with a mobile application to enable visually impaired individuals to read Sinhala text through real-time text recognition and voice feedback. The methodology consists of several phases, including system design, hardware and software development, data collection, and evaluation.

The proposed system consists of **smart glasses** equipped with a **Raspberry Pi camera**, a **mobile application**, and **AI-based Optical Character Recognition (OCR) and Text-to-Speech (TTS) processing**. The **system diagram** (Figure 2) illustrates the interaction between these components:

1. **Text Capture**: The Raspberry Pi camera embedded in the smart glasses captures an image of printed Sinhala text.
2. **Preprocessing**: The captured image is transmitted to the mobile application for enhancement (noise removal, contrast adjustment, and segmentation).
3. **Text Recognition (OCR Processing)**: AI-based OCR extracts Sinhala characters from the image.
4. **Text-to-Speech Conversion**: Recognized text is converted into natural-sounding speech using Sinhala TTS models.
5. **Voice Output & User Interaction**: The synthesized speech is played through the mobile device, with **real-time voice guidance** assisting users in positioning the text correctly.



5.2 System Overview

The proposed system is a mobile-based Sinhala book reader designed for visually impaired individuals, utilizing smart glasses equipped with a Raspberry Pi camera and a mobile application powered by Tiny ML and Edge AI. The system aims to provide real-time text recognition and audio feedback, enabling hands-free reading of printed Sinhala text. The smart glasses capture images of printed material, which are processed by the mobile application to extract text using Optical Character Recognition (OCR) and convert it into synthesized speech for the user. The integration of real-time voice guidance ensures correct text positioning, improving recognition accuracy and ease of use.

The system consists of multiple components working together to deliver an efficient and accessible reading experience. The Preprocessing Component enhances captured images by applying techniques such as de-skewing, binarization, noise reduction, and contrast adjustment to improve OCR accuracy. The Segmentation Component isolates individual characters or words, addressing the challenge of connected Sinhala script using line and word segmentation techniques. The Feature Extraction Component identifies unique character attributes such as stroke width, height, and orientation, enabling the OCR engine to recognize Sinhala text accurately. The Recognition Component processes segmented characters using a trained AI-based OCR model to convert images into digital text.

Once the text is recognized, the Postprocessing Component refines the output by correcting errors through language modeling and spell-checking techniques. The Text-to-Speech (TTS) Conversion Component then transforms the recognized text into natural-sounding Sinhala speech, which is played back to the user through the mobile application. Additionally, the system includes an Interactive Voice Guidance Component, which provides real-time auditory feedback to help users align the text correctly within the camera's frame, ensuring accurate text capture.

The system is designed for real-world usability, prioritizing offline functionality to ensure accessibility without internet dependence. By integrating lightweight AI models optimized for edge computing, the system provides low-latency performance while maintaining affordability. The mobile application, developed using React Native, ensures cross-platform compatibility, allowing users to interact seamlessly with the system. Through user-centered design and iterative testing, the system aims to improve the independence and quality of life of visually impaired individuals by making Sinhala printed materials more accessible.

6. PROJECT REQUIRESMENT

6.1. Functional Requirements

• Text Capture: The system must allow users to capture images of printed Sinhala text using the Raspberry Pi camera integrated into smart glasses.

• Optical Character Recognition (OCR): The system must extract Sinhala text from captured images and convert it into digital text using AI-based OCR models.

• Text-to-Speech (TTS) Conversion: The system must generate clear and natural Sinhala speech from the recognized text.

• Offline Processing: The system must perform OCR and TTS tasks without requiring an internet connection to ensure accessibility.

• Mobile Application Interface: The system must include a mobile application for user interaction, providing options to adjust volume, playback speed, and other settings.

6.2. Non-Functional Requirements

• Performance: The system should process text and generate speech output in real-time with minimal delay.

• Accuracy: The OCR model should achieve accuracy in Sinhala text recognition.

• Usability: The system should have a user-friendly interface with simple navigation and voice-based guidance.

• Reliability: The system must function consistently without frequent crashes or failures.

• Security: User data, including saved text and settings, must be securely stored and protected.

• Portability: The system should be lightweight and should run on mobile devices with minimal processing power.

• Scalability: The system should allow future enhancements, such as adding additional assistive features.

6.3. Software Requirement

• Mobile Operating System: The application should be compatible with Android operating systems.

• OCR Engine: The application should use a reliable and high-quality OCR engine like Tesseract OCR

6.4. User Requirements

• Sinhala Language Support: The application should support the Sinhala language and OCR should identify the Sinhala characters.

• Simple and intuitive voice-guided interaction.

• Offline functionality without requiring an internet connection.

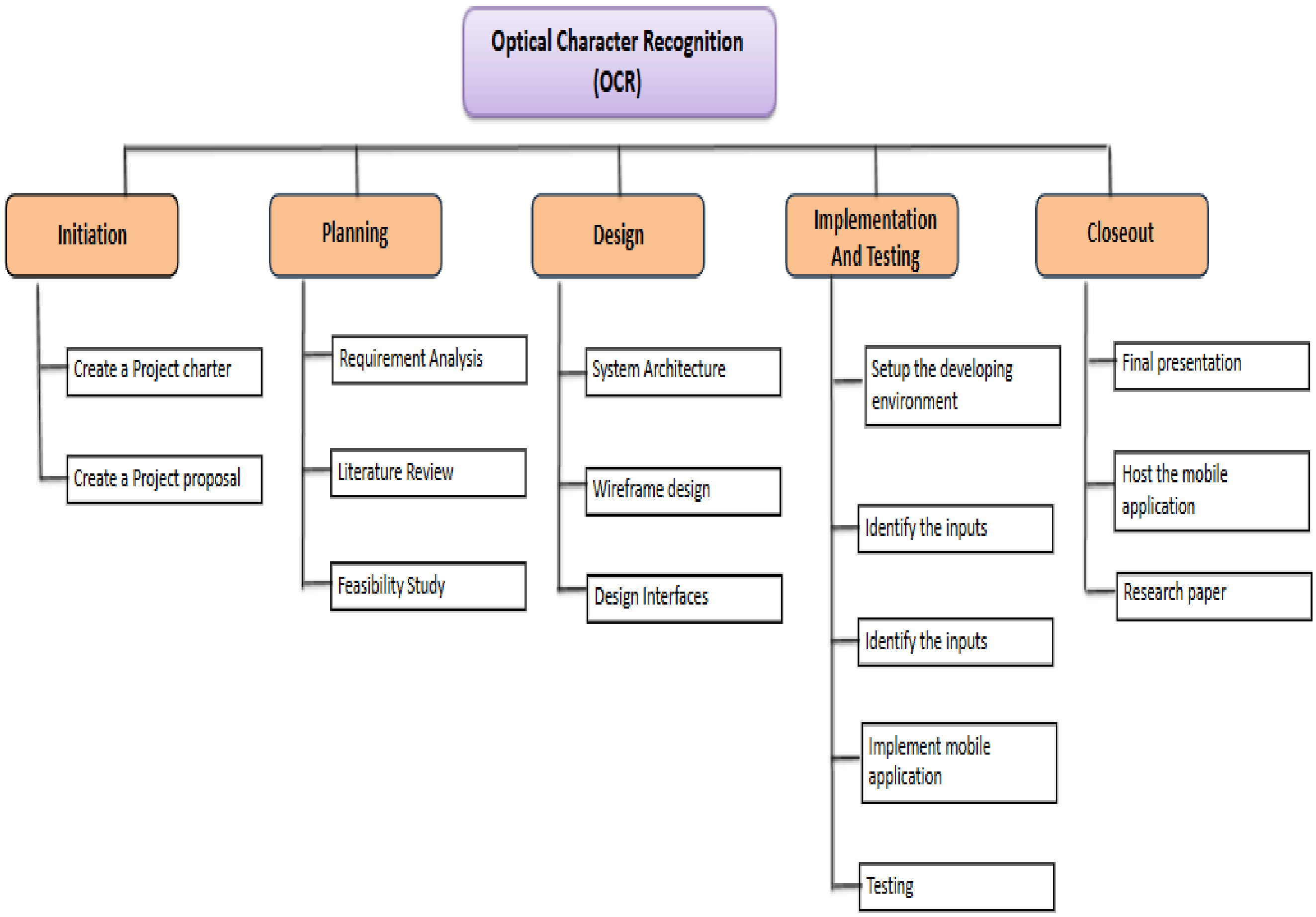
Developers & Researchers:

• Ability to update and retrain the OCR model to improve accuracy.

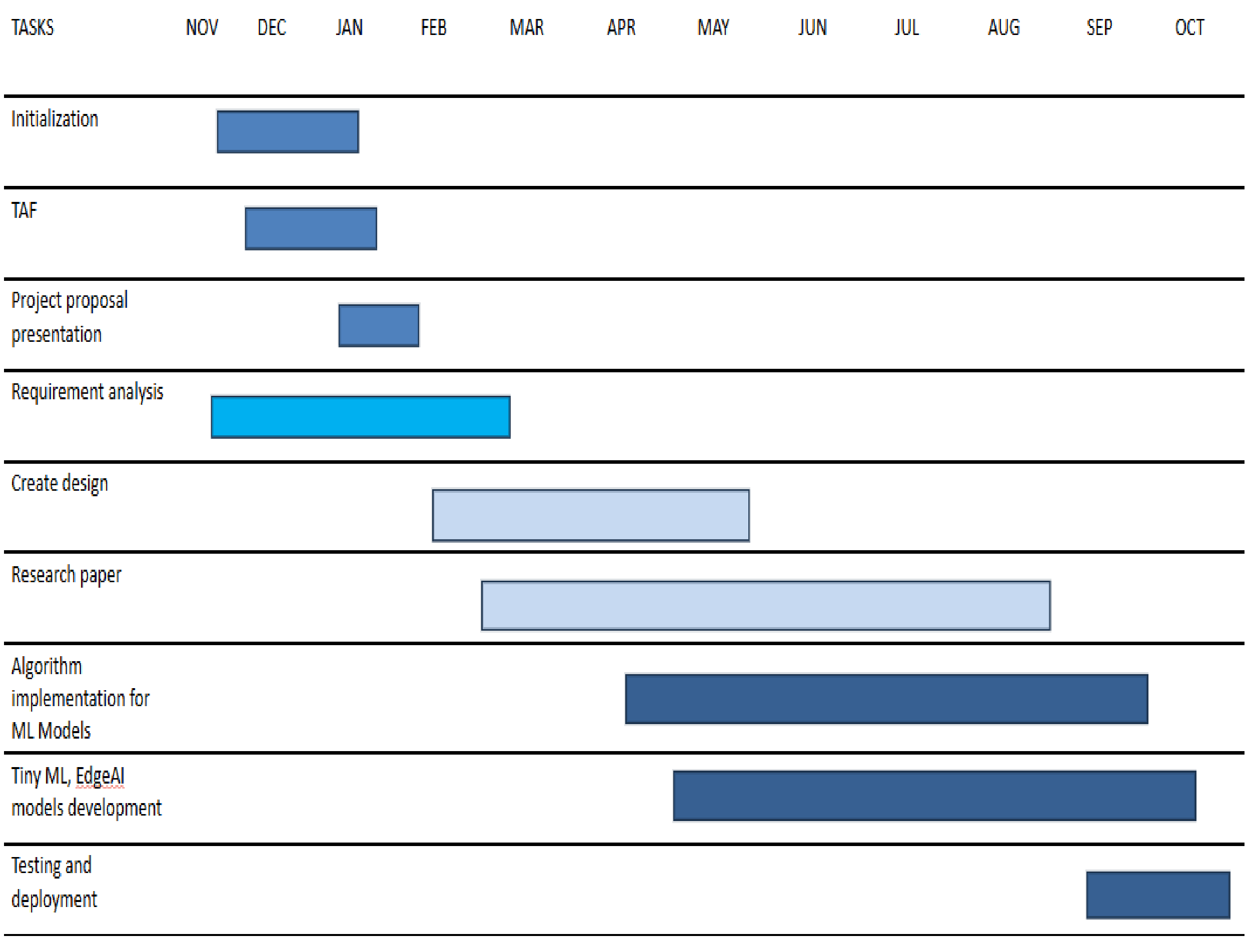
• Compatibility with additional assistive technologies.

• Secure access to user feedback for system improvement.

7. WORK BREAKDOWN CHART



9. GANTT CHART



8. REFERENCES

[1] T. Madhusha, H. D. M. Bhagya, S. Godakanda, J. P. D. Semini, K. Pulasinghe, and P. K. P. G. Panduwawala, "Mobile Base Sinhala Book Reader for Visually Impaired Students," Faculty of Computing, Sri Lanka Institute of Information Technology, Malabe, Sri Lanka. Available: https://irjiet.com/common_src/article_file/1699450424_6d2ebfa895_7_irjiet.pdf.

[2] S. Vasanthapriyan and M. De Silva, "Optical Braille Recognition Software Prototype for the Sinhala Language," Department of Computing & Information Systems, Sabaragamuwa University of Sri Lanka, Belihuloya, Sri Lanka, 2019. Available: https://www.astesj.com/publications/ASTESJ_040427.pdf.

[3] University of Colombo School of Computing, "Development of Sinhala OCR Systems," 2008.

[4] University of Moratuwa, "Development of Sinhala Text-to-Speech (TTS) Systems."

[5] OrCam Technologies, "OrCam MyEye: Portable AI-based Text Recognition Device," 2015. Available: https://www.orcam.com.

[6] N. de Silva, "Survey on Publicly Available Sinhala Natural Language Processing Tools and Research," University of Moratuwa, Revised January 4, 2024.

[7] K. Barathkumar, S. Balaji, K. N. Desikan, and S. P. Iyappan, "Raspberry Pi-based Smart Reader for Visually Impaired People," Valliammai Engineering College, Chennai, India.

[8] R. Smith, "An Overview of the Tesseract OCR Engine," pp. 629-633, Sep 2007.

[9] R. N. and N. Fonseca, "Camera Reading for Blind People," Polytechnic Institute of Leiria, Portugal.

[10] Ministry of Health, "National Survey of Blindness, Visual Impairment, Ocular Morbidity and Disability in Sri Lanka," 2015. Available: https://www.iapb.org/wp-content/uploads/National-Survey-of-Blindness-A-Report-2014-2015.pdf.

[11] A. S. and B. Kubendran, "Optical Character Recognition of Printed Tamil Characters," 2000.

[12] R. J. Perera et al., "Edge AI for Resource-Constrained Devices: A Case Study on Sinhala OCR," IEEE Access, 2020.

[13] A. Hassan et al., "Enhancing OCR Using TinyML for Low-Power Devices," IEEE Transactions on Neural Networks and Learning Systems, 2021.

[14] N. Fernando et al., "Affordable Assistive Technologies for Visually Impaired in Sri Lanka," Proceedings of the 2022 International Conference on Assistive Technologies, 2022.

[15] J. Dissanayake, "Improving Accessibility for Blind Communities Through IoT-Based Wearable Devices," Proceedings of the International Symposium on Assistive Technologies, 2019.

[16] A. Gunasekara et al., "Text-to-Audio Conversion for Sinhala Language Using AI," National ICT Conference, 2021.

[17] Faculty of Computing, "A Study on Accessibility Solutions for Visually Impaired in Sri Lanka," University of Peradeniya, 2018.

[18] R. Silva et al., "Innovations in Assistive Technology: Smart Glasses for Sinhala Text Recognition," Proceedings of the 2024 International Conference on Emerging Technologies, 2024.