

**Mobile Base Sinhala Book Reader for
Visually Impaired Individuals**

TMP-23-198

Status Document 2

Semini J.P.D.L. – IT20241346

Bachelor of Science (Hons) Degree in Information Technology
Specializing in Information Technology

Department of Information Technology
Faculty of Computing

Sri Lanka Institute of Information Technology
Sri Lanka

September 2023

Mobile Base Sinhala Book Reader forVisually Impaired Individuals.....	1
1. Microsoft Teams Details.....	4
1.1 Teams group channels	4
1.2 Team Chat Meetings.....	6
1.3 Teams chat group with the supervisor.....	7
1.4 Teams calls with research team.....	8
1.5 Teams tasks planner	11
2. Whatsapp group and calls	13
3. Gantt Chart.....	16
4. Work Breakdown Chart	16
5. Gitlab.....	17
6. System methodology diagram	18
7. Current Progress up to now	19
8. Our Research Paper	23
9. Extra helping resources.....	24

Figure 1-Team members	4
Figure 2-Team documents	5
Figure 3-Team chat meetings	6
Figure 4-Team group with the supervisor	7
Figure 5-Team Calls in teams.....	10
Figure 6-Teams tasks planner.....	12
Figure 7-Teams whatsapp calls	14
Figure 8-Teams whatsapp group	15
Figure 9-Grant chart	16
Figure 10-Work Breakdown Chart	16
Figure 11-Gitlab	17
Figure 12-System diagram	18
Figure 13-dataset	19
Figure 14-dataset training.....	20
Figure 15-backend postman API	21
Figure 16-frontend.....	22
Figure 17-pp2 presentation.....	22
Figure 18-research paper	23
Figure 19-you tube tutorials	24
Figure 20-chat gpt	24
Figure 21-research papers	25
Figure 22-W3 School	25

1. Microsoft Teams Details

1.1 Teams group channels

The screenshot shows the Microsoft Teams interface for a team named 'Blind App' (Research Project Group 2023). The left sidebar contains navigation icons for Activity, Chat, Teams, Assignments, Calendar, Calls, Files, and Apps. The main pane displays the 'Members' tab with a search bar and an 'Add member' button. Below the search bar, there are two sections: 'Owners (1)' and 'Members and guests (3)'. The 'Owners' section lists one member: Jayathunga T. M. (it20146238) with the role of Owner. The 'Members and guests' section lists three members: Semini J.P.D.L. (it20241346), Bhagya H.D.M. (it20254520), and Godakanda P. G. S. (it20129712), all with the role of Member.

Name	Title	Location	Tags	Role
JL Jayathunga T. M. it20146238				Owner

Name	Title	Location	Tags	Role
SL Semini J.P.D.L. it20241346				Member
BL Bhagya H.D.M. it20254520				Member
GL Godakanda P. G. S. it20129712				Member

The screenshot shows the Microsoft Teams interface for a team named 'Mobile Base Sinhala Book Reader for Visually Impaired Individuals' (2023 Research Project Group). The left sidebar contains navigation icons for Activity, Chat, Teams, Assignments, Calendar, Calls, Files, and Apps. The main pane displays the 'Members' tab with a search bar and an 'Add member' button. Below the search bar, there are two sections: 'Owners (1)' and 'Members and guests (4)'. The 'Owners' section lists one member: Jayathunga T. M. (it20146238) with the role of Owner. The 'Members and guests' section lists four members: Semini J.P.D.L. (it20241346), Bhagya H.D.M. (it20254520), Godakanda P. G. S. (it20129712), and Poorna Panduwawala (Assistant Lecturer, Malabe), all with the role of Member.

Name	Title	Location	Tags	Role
JL Jayathunga T. M. it20146238				Owner

Name	Title	Location	Tags	Role
SL Semini J.P.D.L. it20241346				Member
BL Bhagya H.D.M. it20254520				Member
GL Godakanda P. G. S. it20129712				Member
PL Poorna Panduwawala	Assistant Lecturer	Malabe		Member

Figure 1-Team members

The screenshot displays the Microsoft Teams interface. On the left, a sidebar shows navigation options: Activity, Chat, Teams, Assignments, Calendar, Calls, Files, Apps, and Help. The main area is divided into two panes. The left pane shows the 'General' channel of a team named 'Mobile Base Sinhala Book Reading...'. The right pane displays a list of documents in the 'General' channel. The documents are organized in a table with columns for Name, Modified, and Modified By.

Name	Modified	Modified By
Project Charter	March 22	Jayathunga T. M. it...
Proposal Presentation	April 2	Jayathunga T. M. it...
Proposal Presentation 1	May 26	Jayathunga T. M. it...
Proposal Presentation(Draft)	March 22	Jayathunga T. M. it...
Recordings	6 days ago	Semini J.P.D.L. it20...
Topic Assignment Form	February 15	Jayathunga T. M. it...
Madushi.txt	Monday at 10:39 PM	Jayathunga T. M. it...
research papers.zip	August 30	Semini J.P.D.L. it20...
Screenshot_25.png	August 14	Jayathunga T. M. it...
Screenshot_277.png	August 30	Jayathunga T. M. it...
Screenshot_277.png	Sunday at 9:56 PM	Jayathunga T. M. it...

Figure 2-Team documents

1.2 Team Chat Meetings

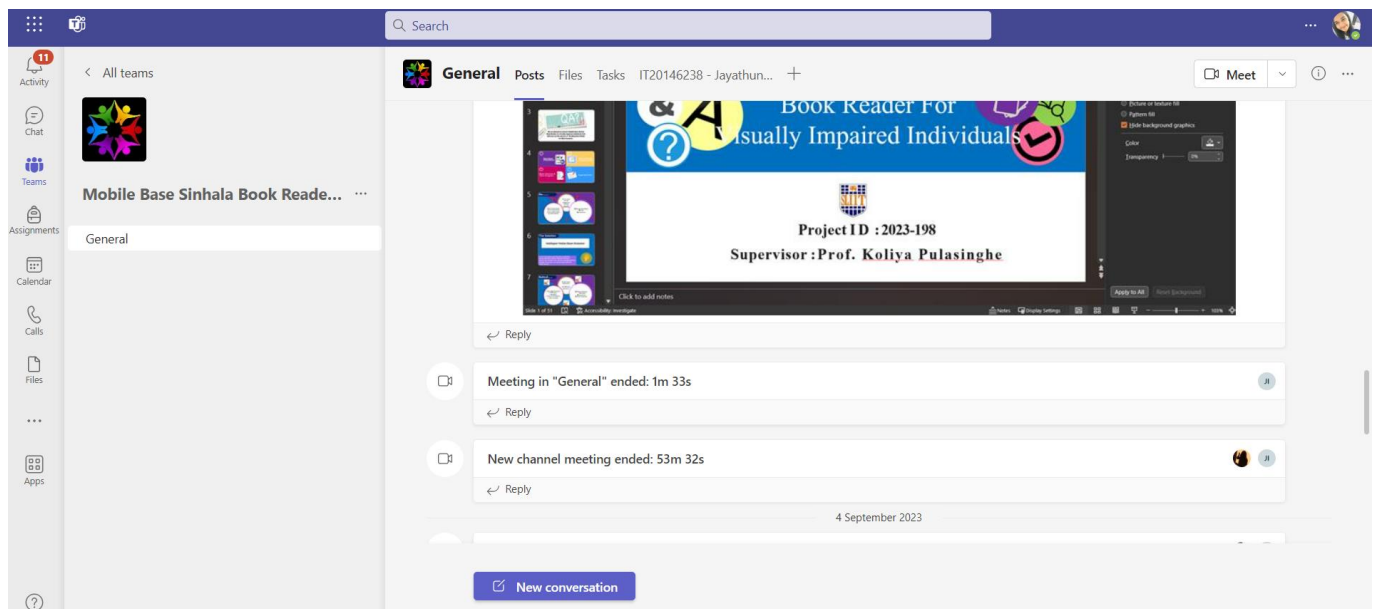
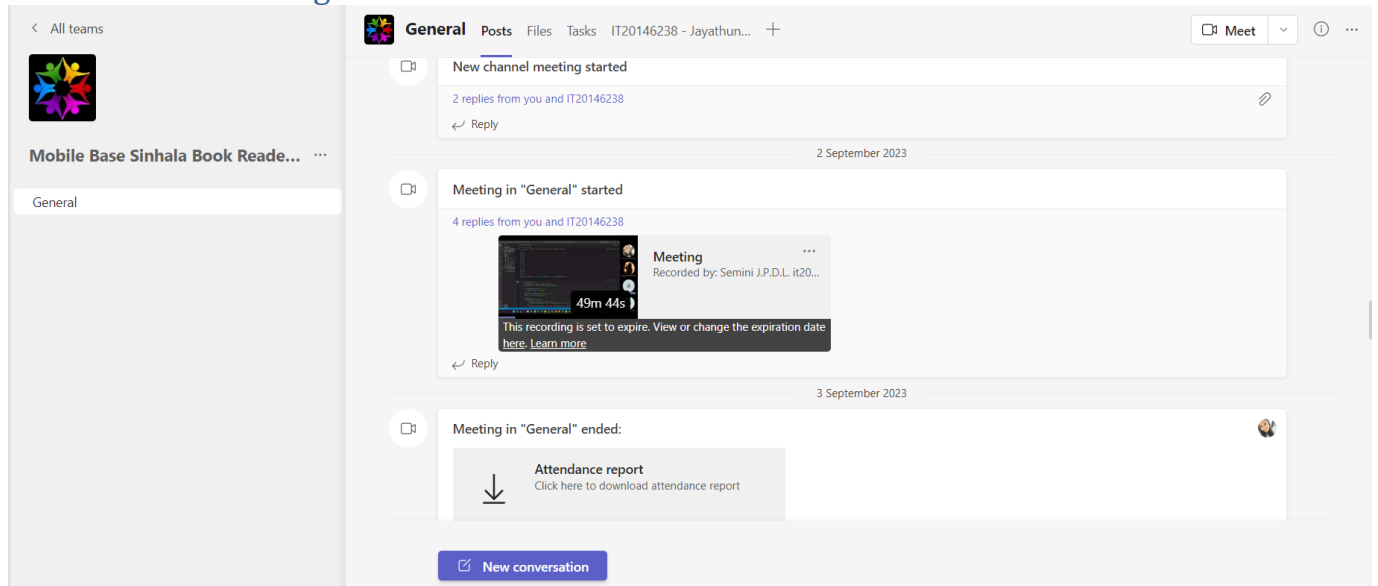


Figure 3-Team chat meetings

1.3 Teams chat group with the supervisor

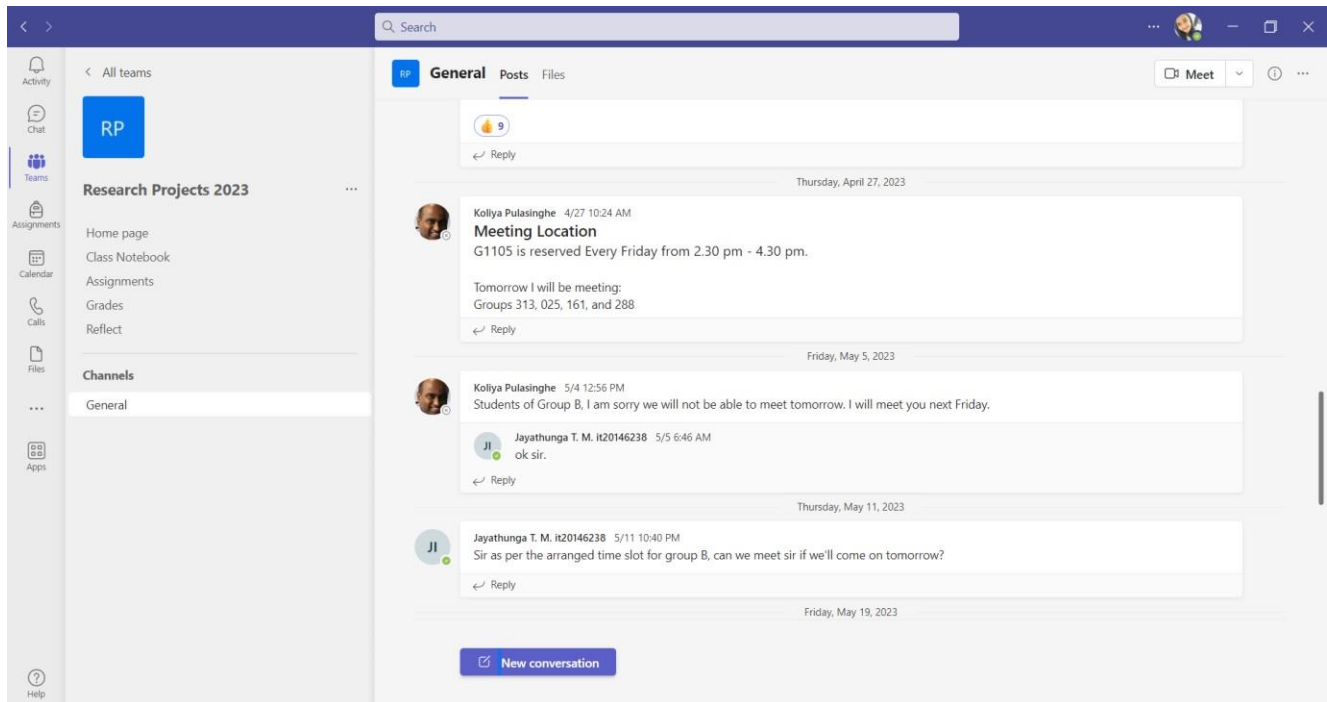
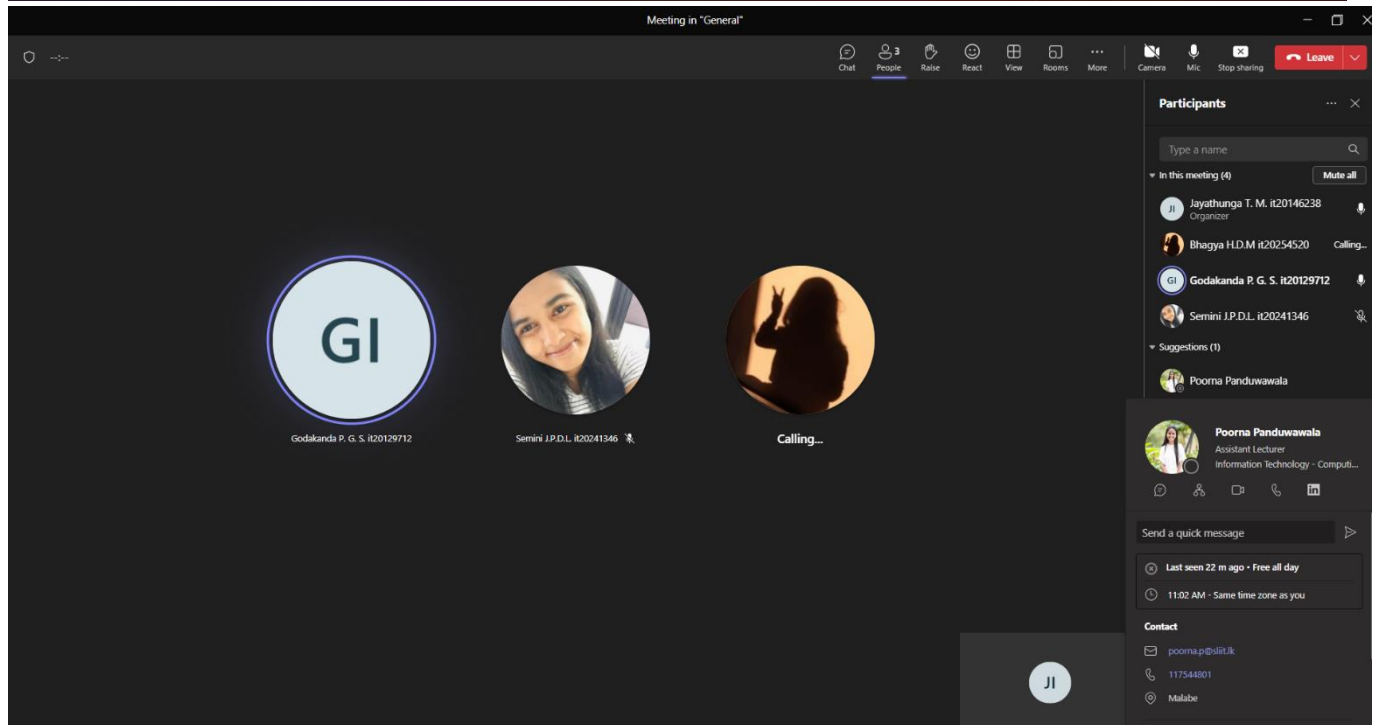
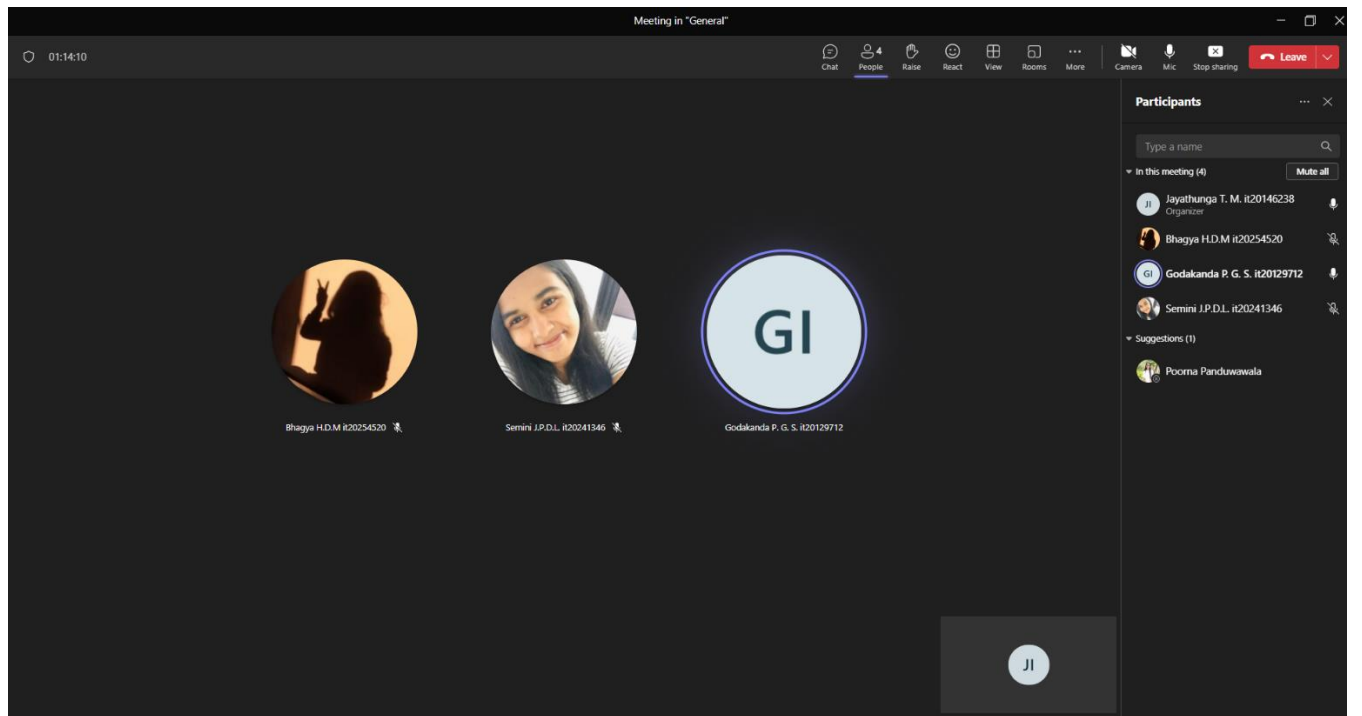
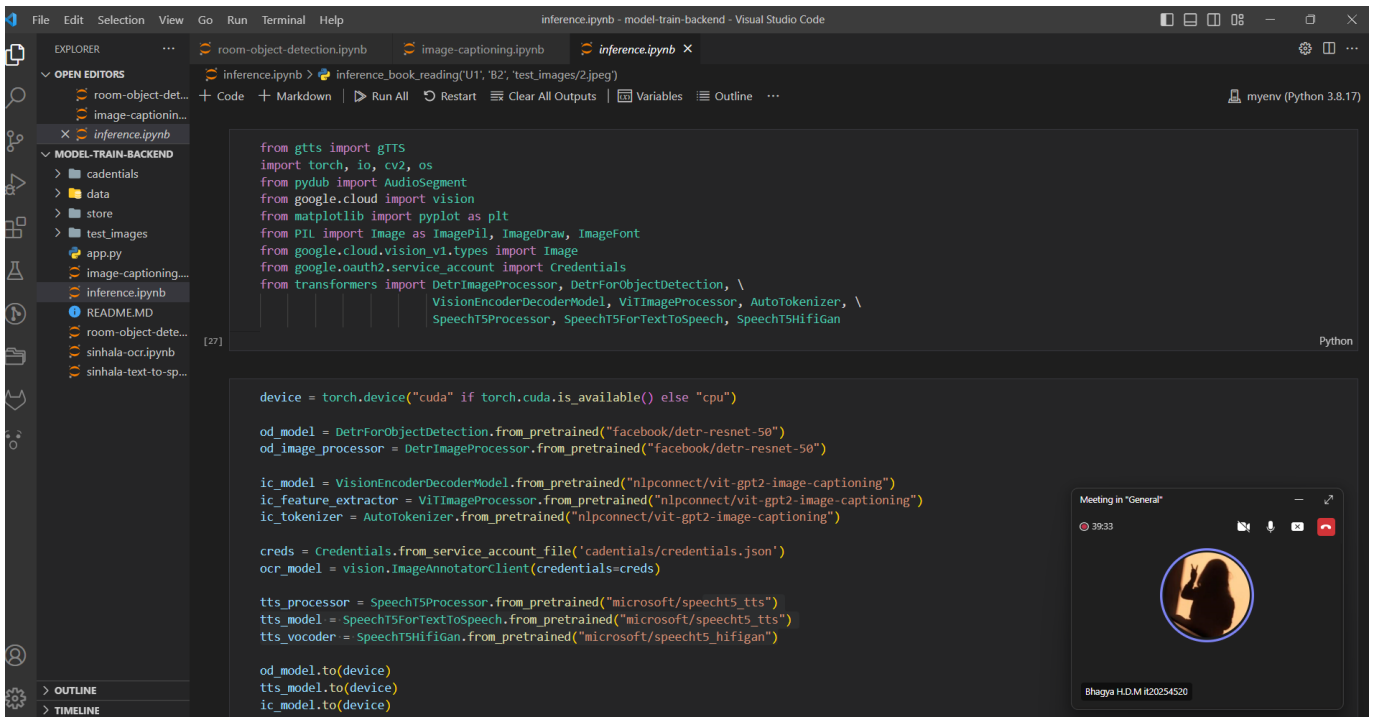
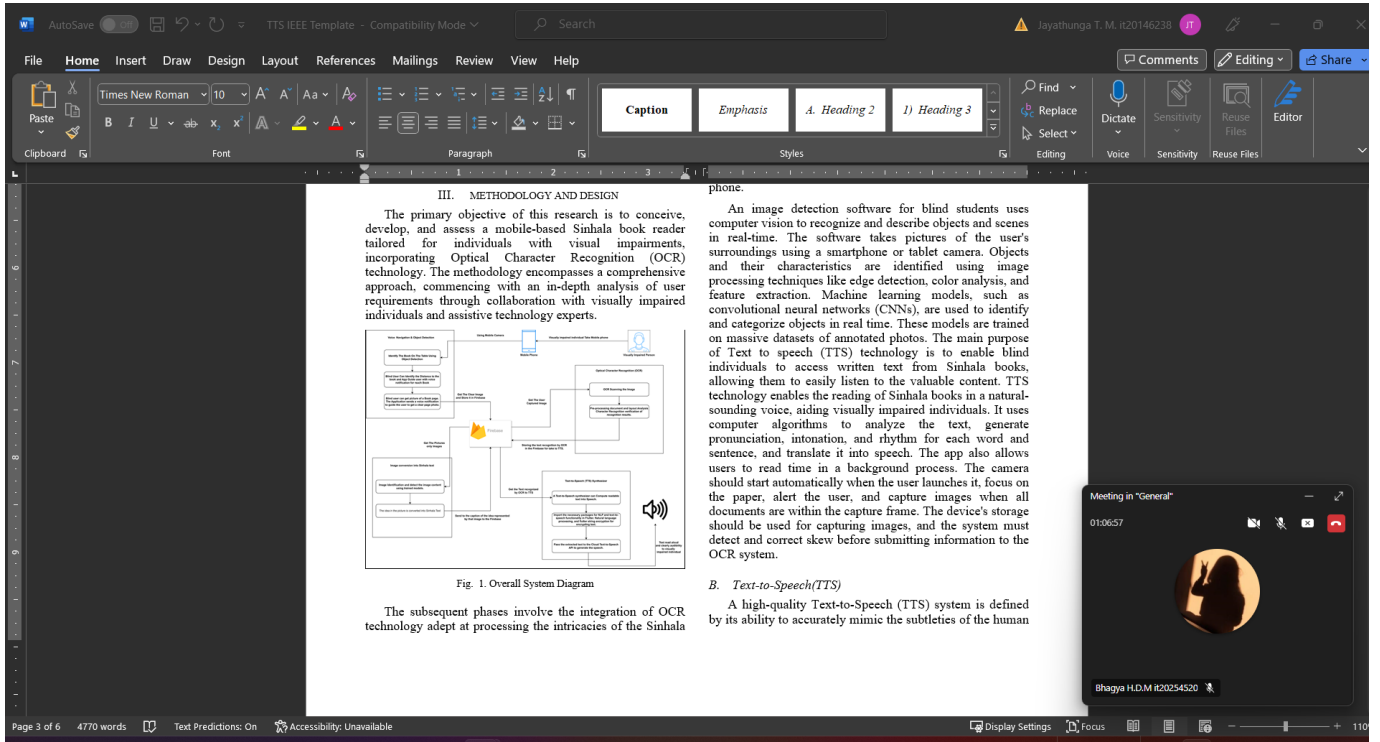


Figure 4-Team group with the supervisor

1.4 Teams calls with research team





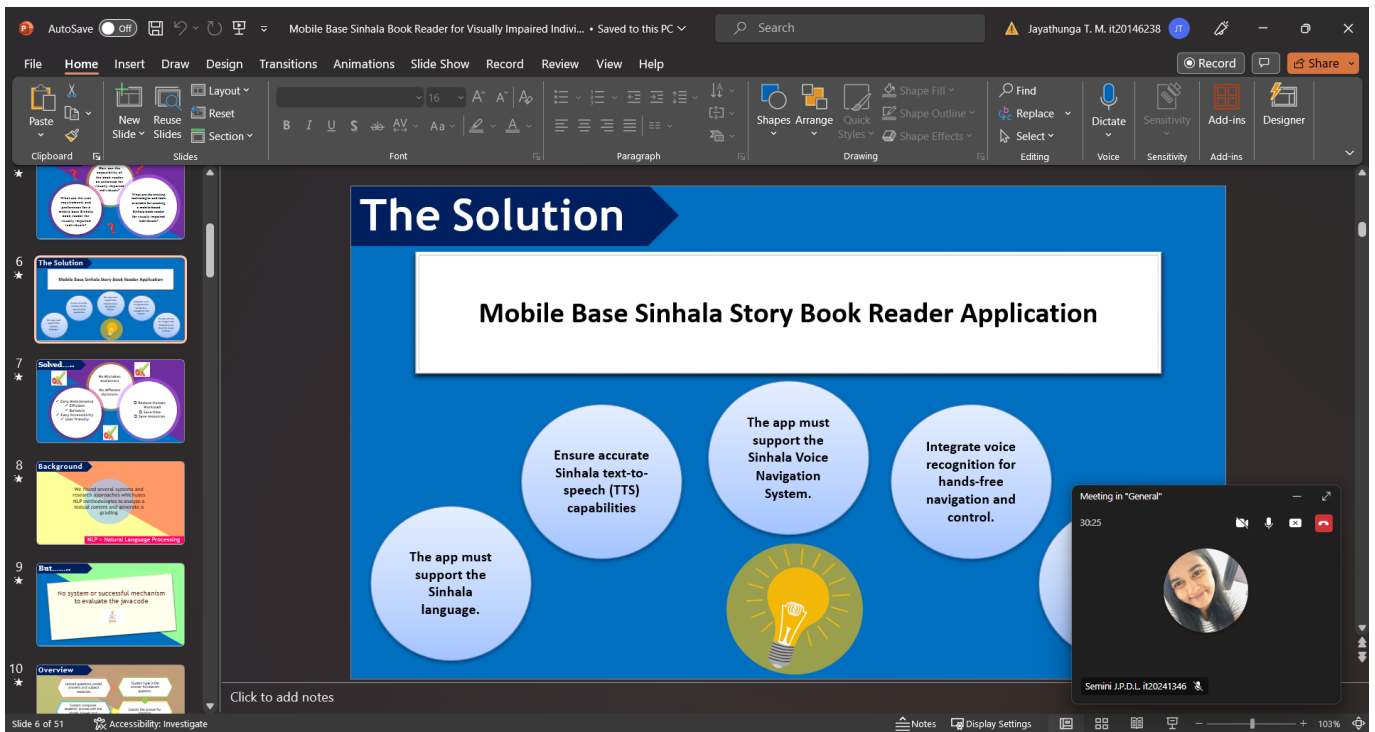
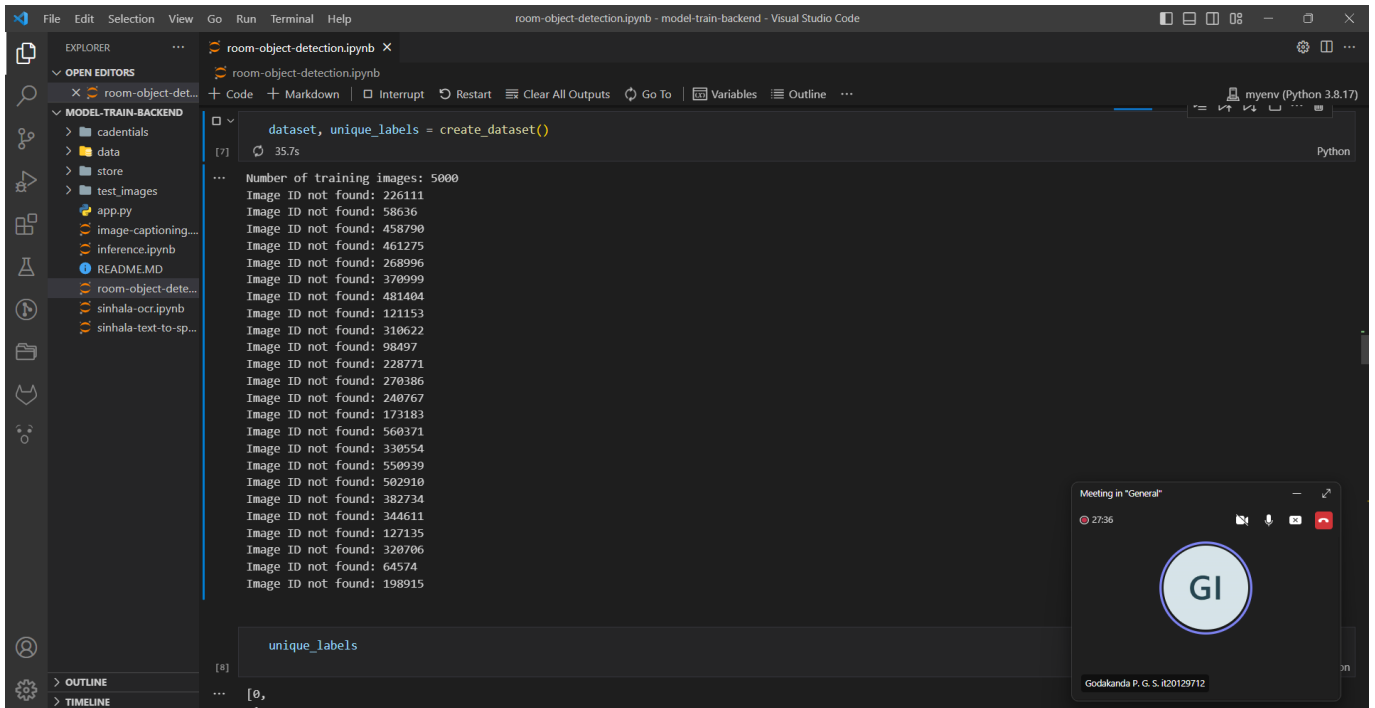
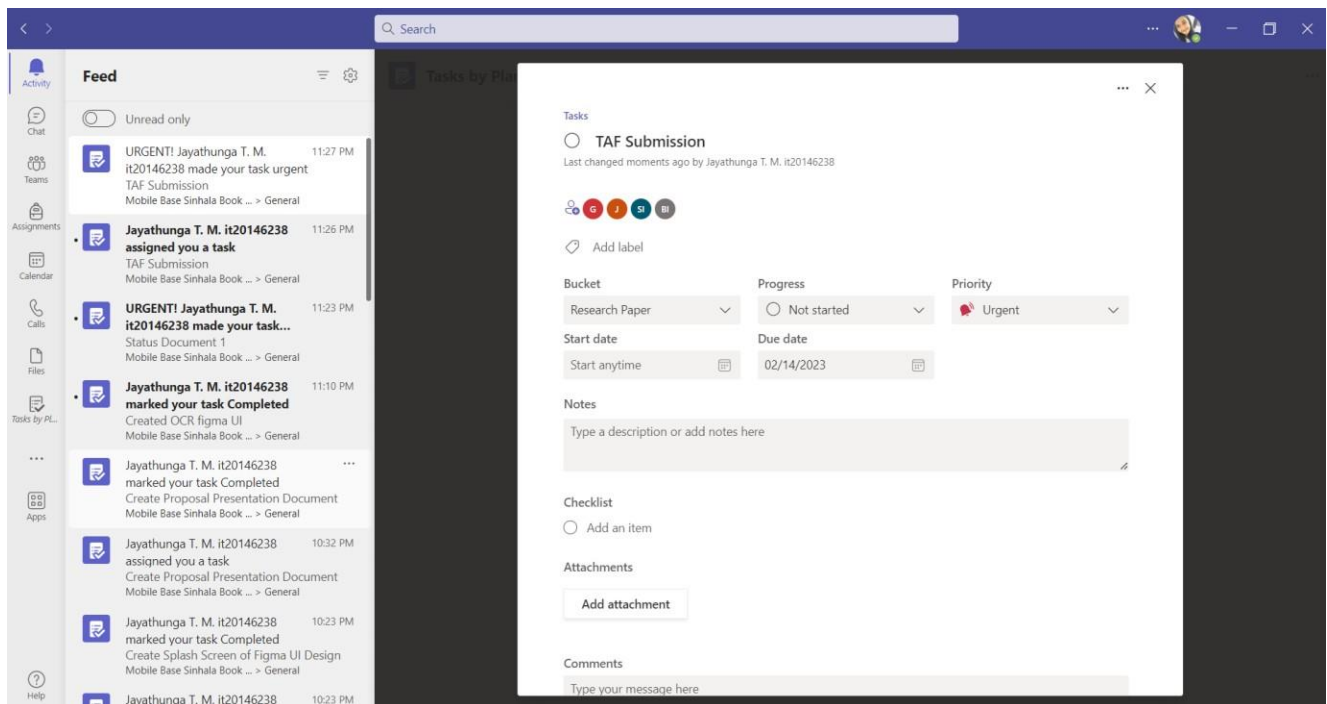
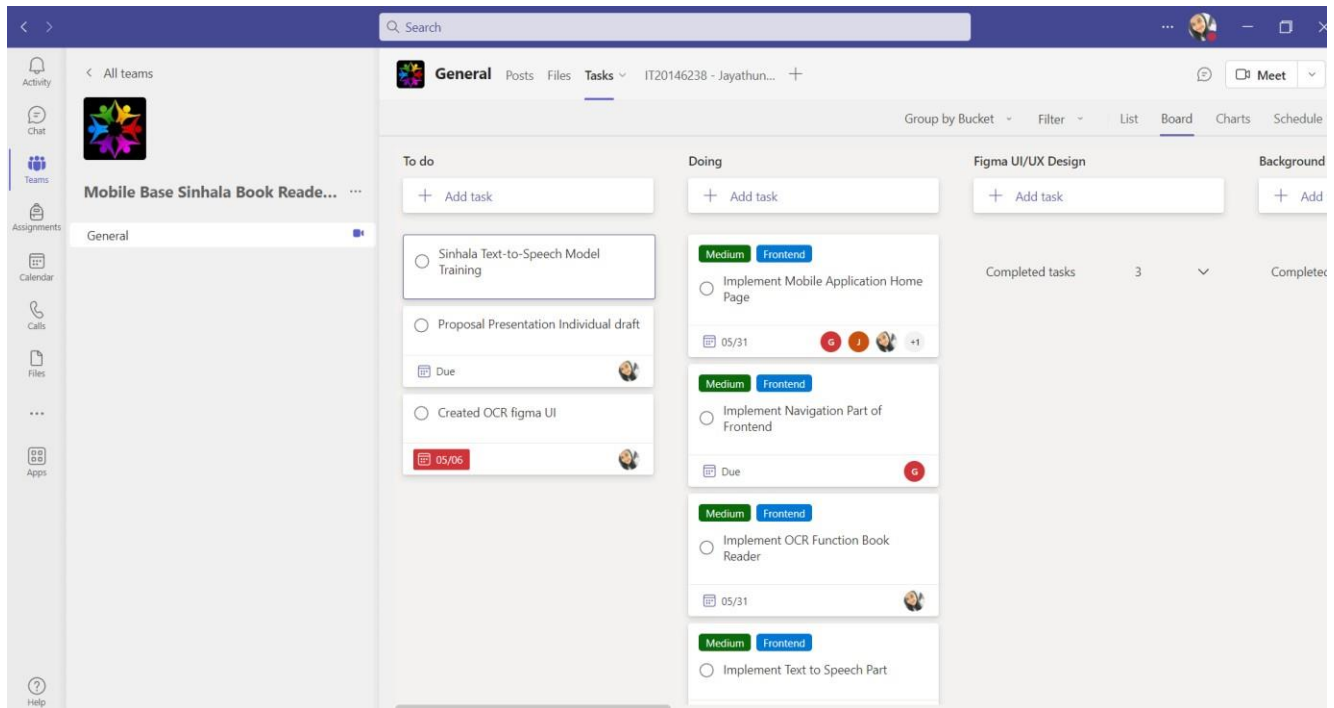


Figure 5-Team Calls in teams

1.5 Teams tasks planner



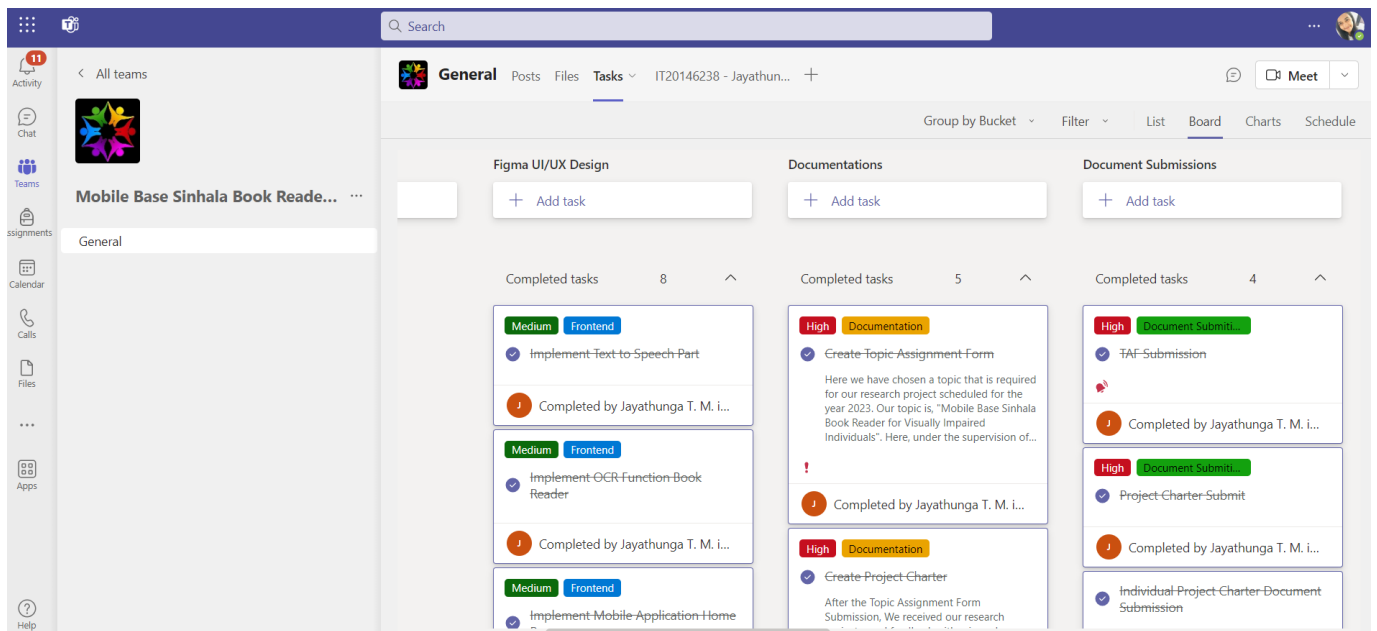
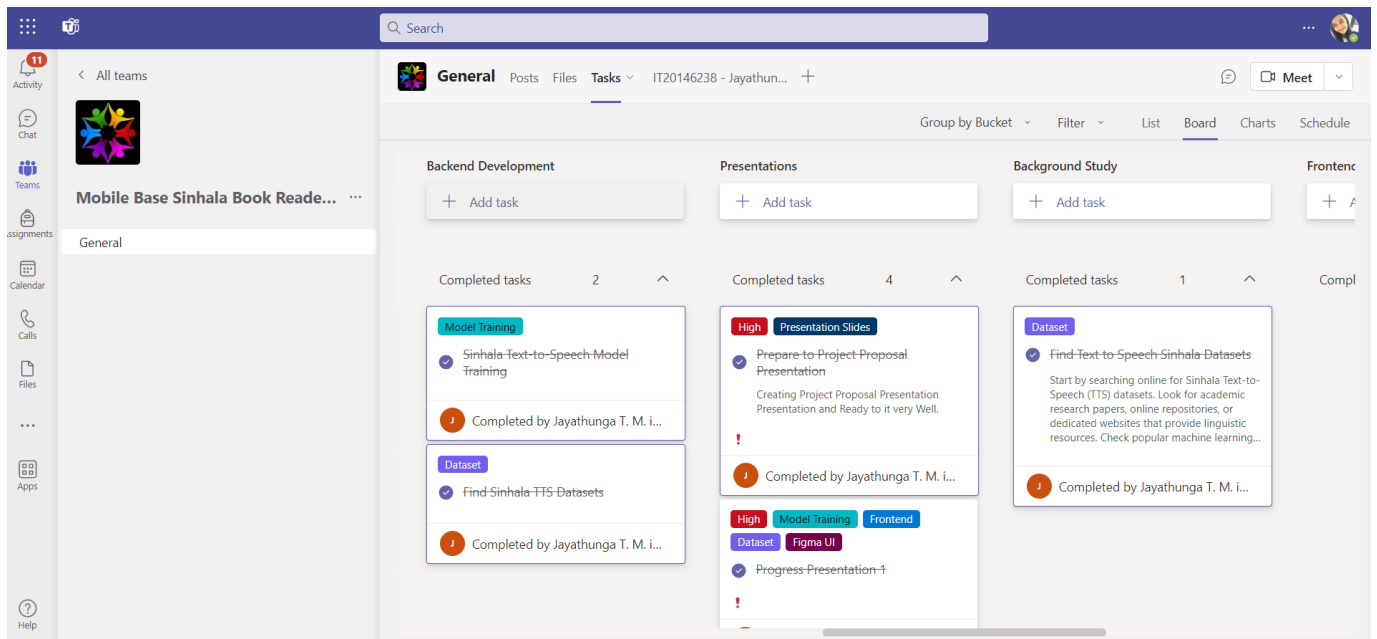
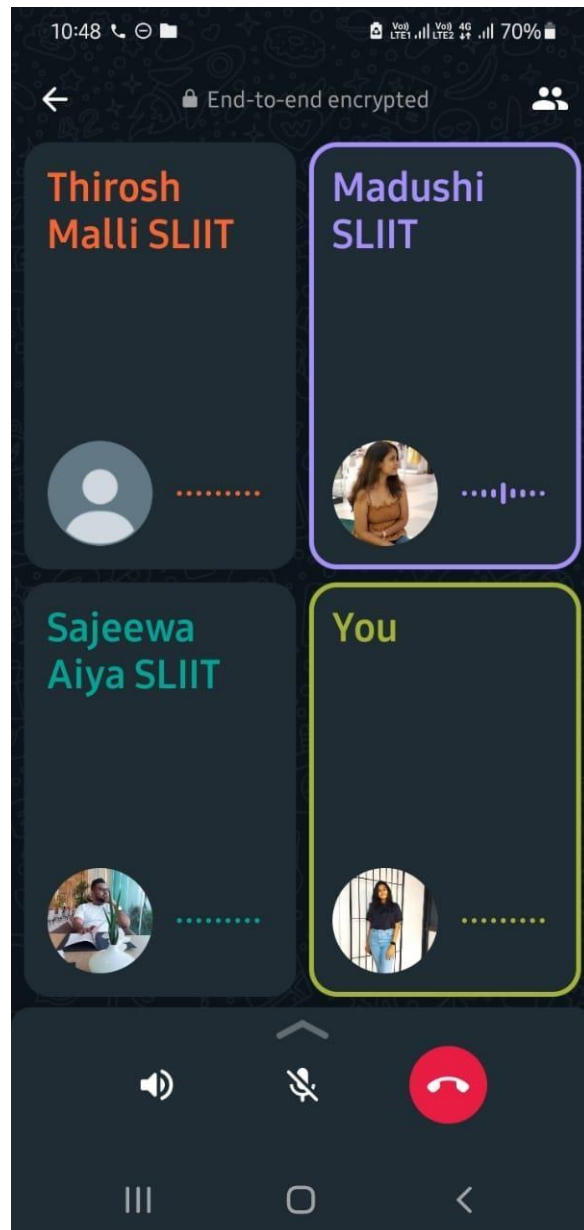


Figure 6-Teams tasks planner

2. Whatsapp group and calls



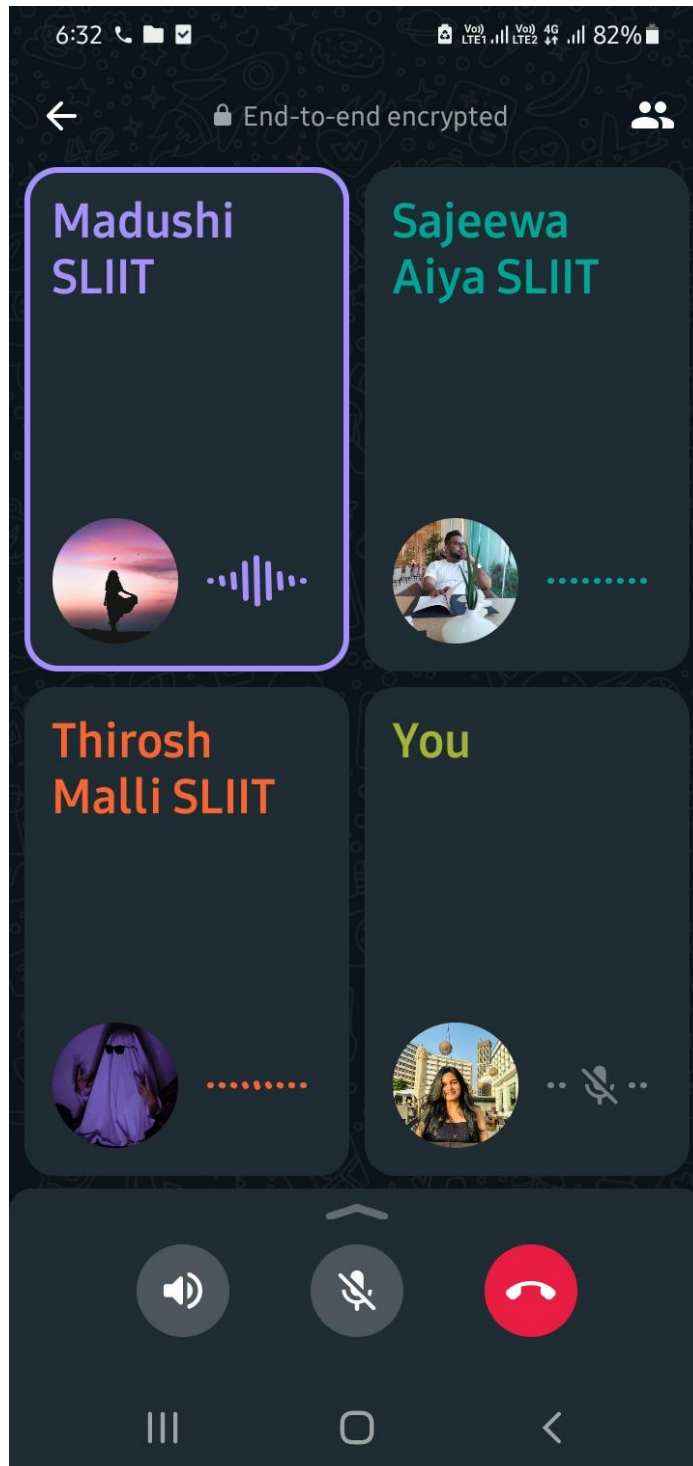


Figure 7-Teams whatsapp calls

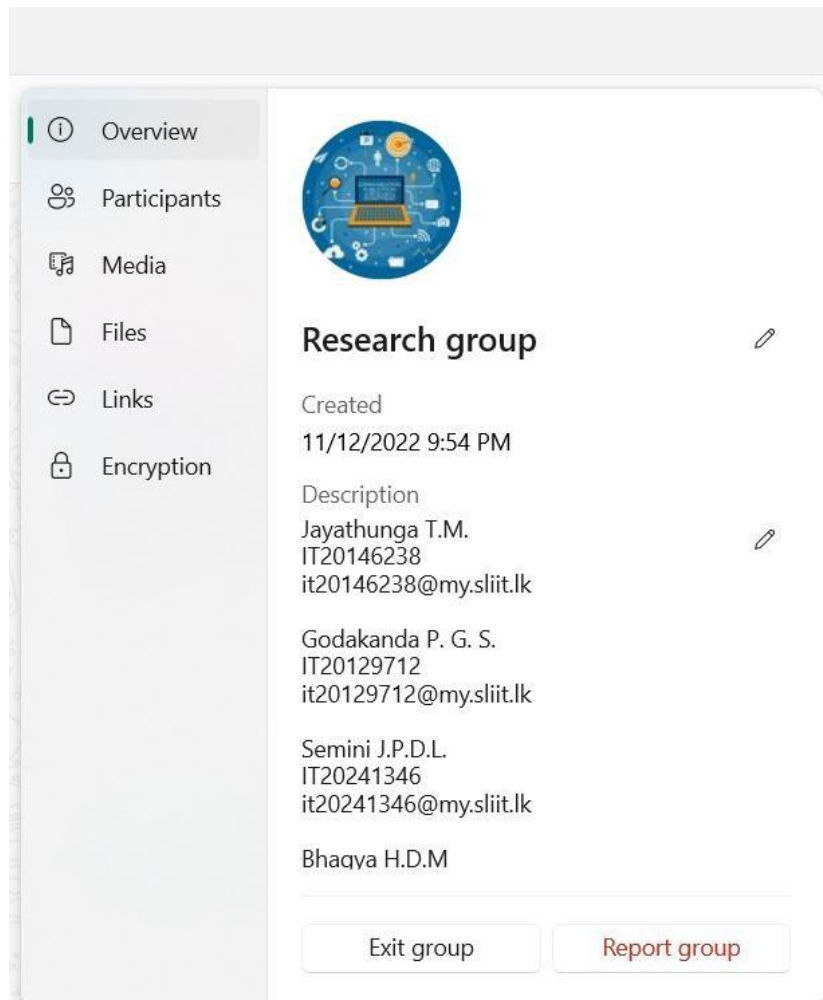


Figure 8-Teams whatsapp group

3. Gantt Chart

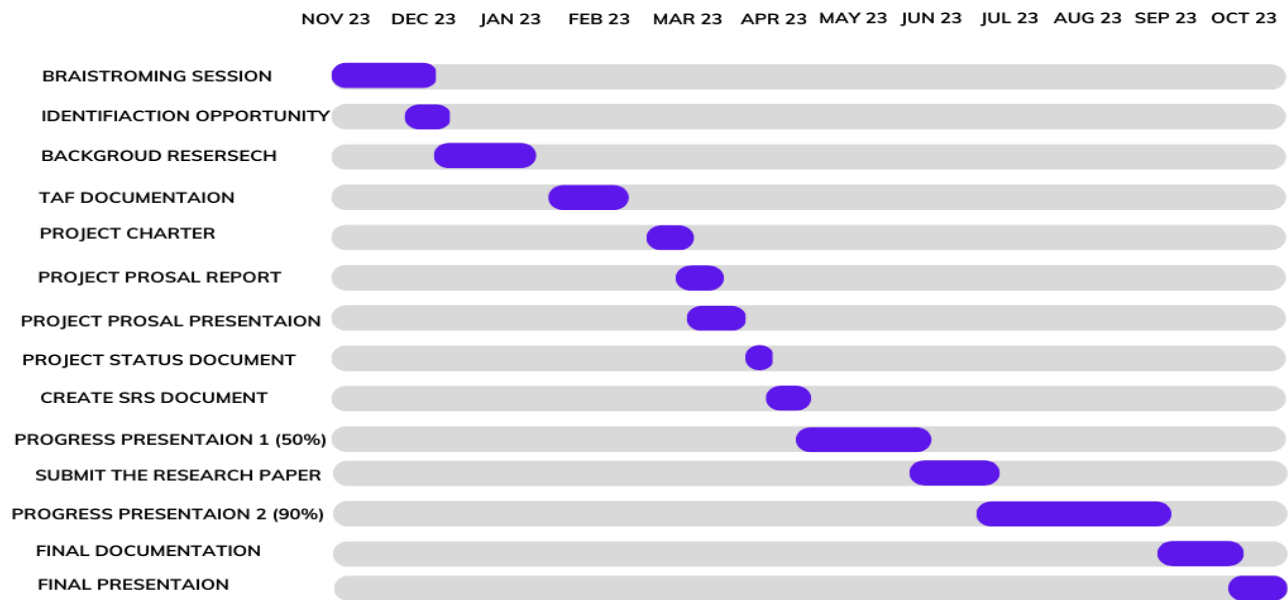


Figure 9-Grant chart

4. Work Breakdown Chart

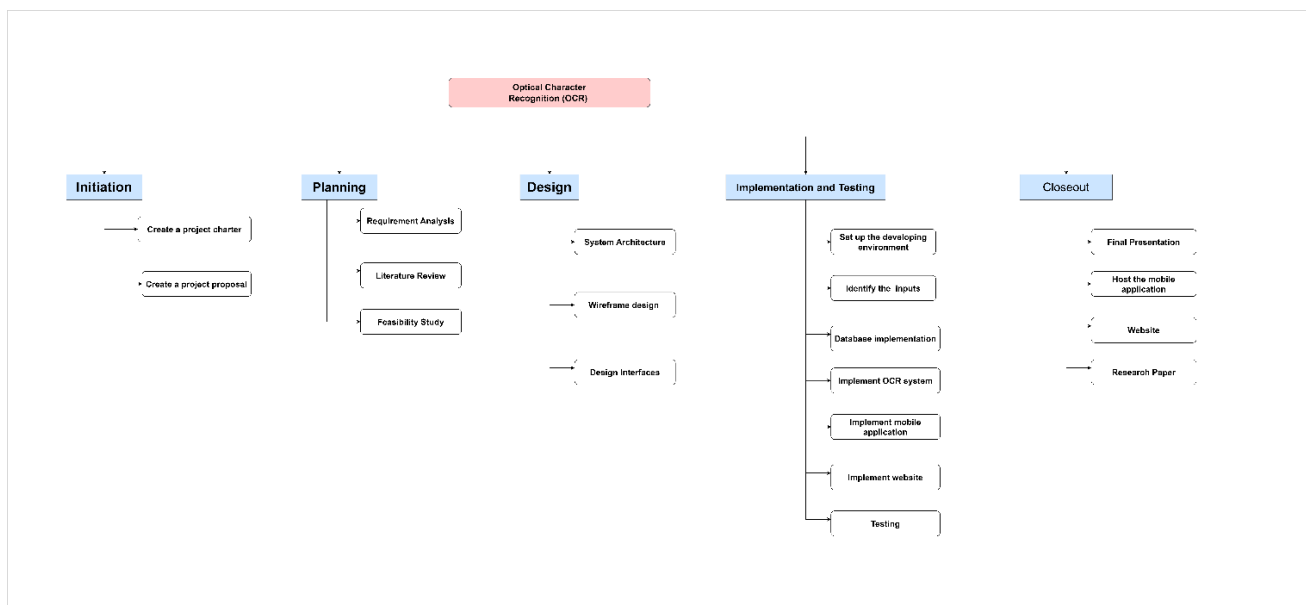
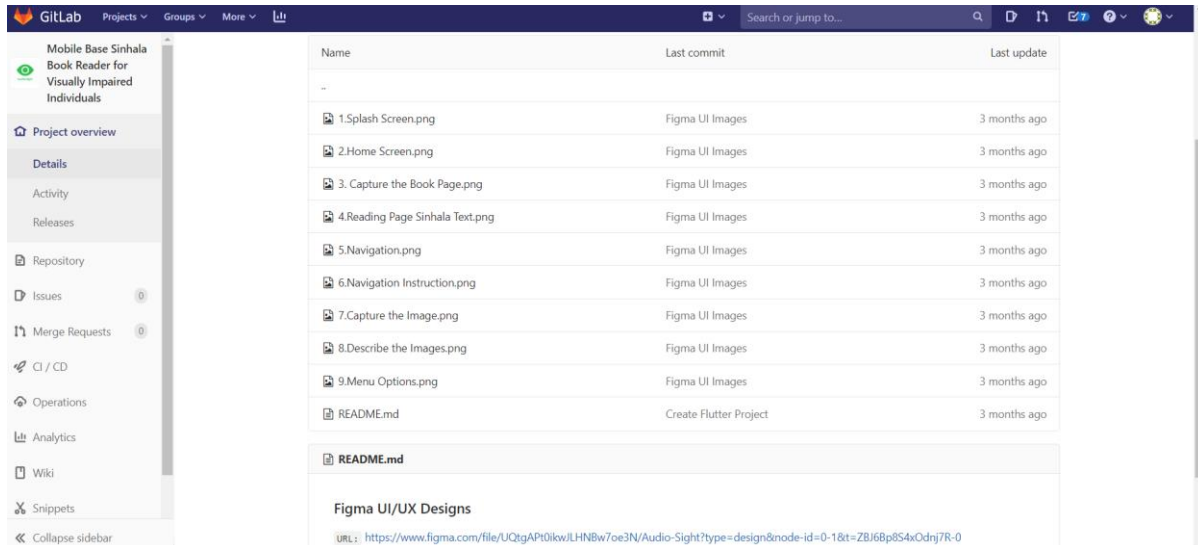


Figure 10-Work Breakdown Chart

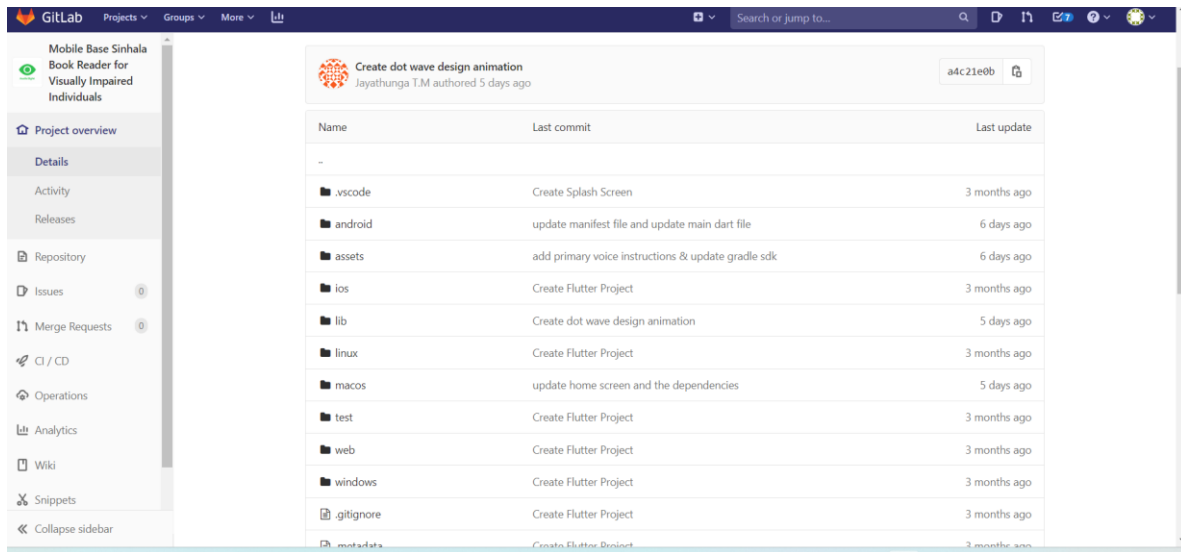
5. Gitlab



The screenshot shows the GitLab interface for a repository named "Mobile Base Sinhala Book Reader for Visually Impaired Individuals". The left sidebar contains navigation links: Project overview, Details, Activity, Releases, Repository, Issues (0), Merge Requests (0), CI / CD, Operations, Analytics, Wiki, and Snippets. The main content area displays a table of files with columns for Name, Last commit, and Last update.

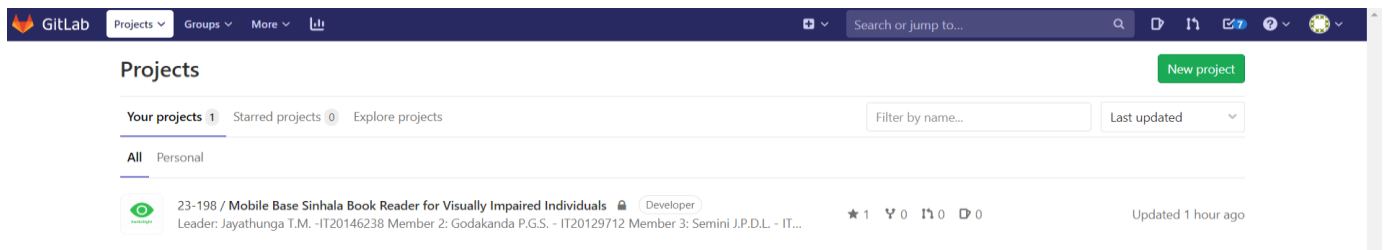
Name	Last commit	Last update
1.Splash Screen.png	Figma UI Images	3 months ago
2.Home Screen.png	Figma UI Images	3 months ago
3. Capture the Book Page.png	Figma UI Images	3 months ago
4.Reading Page Sinhala Text.png	Figma UI Images	3 months ago
5.Navigation.png	Figma UI Images	3 months ago
6.Navigation Instruction.png	Figma UI Images	3 months ago
7.Capture the Image.png	Figma UI Images	3 months ago
8.Describe the Images.png	Figma UI Images	3 months ago
9.Menu Options.png	Figma UI Images	3 months ago
README.md	Create Flutter Project	3 months ago

Below the table, the README.md file content is displayed, titled "Figma UI/UX Designs". It includes a URL: <https://www.figma.com/file/UQtgAPt0ikwJLHN8w7oe3N/Audio-Sight?type=design&node-id=0-1&t=ZBj68p8S4xOdj7R-0>.



This screenshot shows the commit history of the same repository. The commit titled "Create dot wave design animation" by Jayathunga T.M. is highlighted. Below it, a table lists the files and their corresponding commit messages and dates.

Name	Last commit	Last update
..		
.vscode	Create Splash Screen	3 months ago
android	update manifest file and update main dart file	6 days ago
assets	add primary voice instructions & update gradle sdk	6 days ago
ios	Create Flutter Project	3 months ago
lib	Create dot wave design animation	5 days ago
linux	Create Flutter Project	3 months ago
macos	update home screen and the dependencies	5 days ago
test	Create Flutter Project	3 months ago
web	Create Flutter Project	3 months ago
windows	Create Flutter Project	3 months ago
.gitignore	Create Flutter Project	3 months ago
metadata	Create Flutter Project	3 months ago



The screenshot shows the GitLab Projects page. At the top, there's a "New project" button. Below it, the "Your projects" section shows a list of projects. The project "23-198 / Mobile Base Sinhala Book Reader for Visually Impaired Individuals" is selected, showing its details. The project is owned by Jayathunga T.M. and has a leader, Jayathunga T.M. The project is updated 1 hour ago.

Projects

Your projects 1 Starred projects 0 Explore projects

Filter by name... Last updated

All Personal

23-198 / Mobile Base Sinhala Book Reader for Visually Impaired Individuals Developer

Leader: Jayathunga T.M. - IT20146238 Member 2: Godakanda P.G.S. - IT20129712 Member 3: Semini J.P.D.L. - IT...

Updated 1 hour ago

Figure 11-Gitlab

6. System methodology diagram

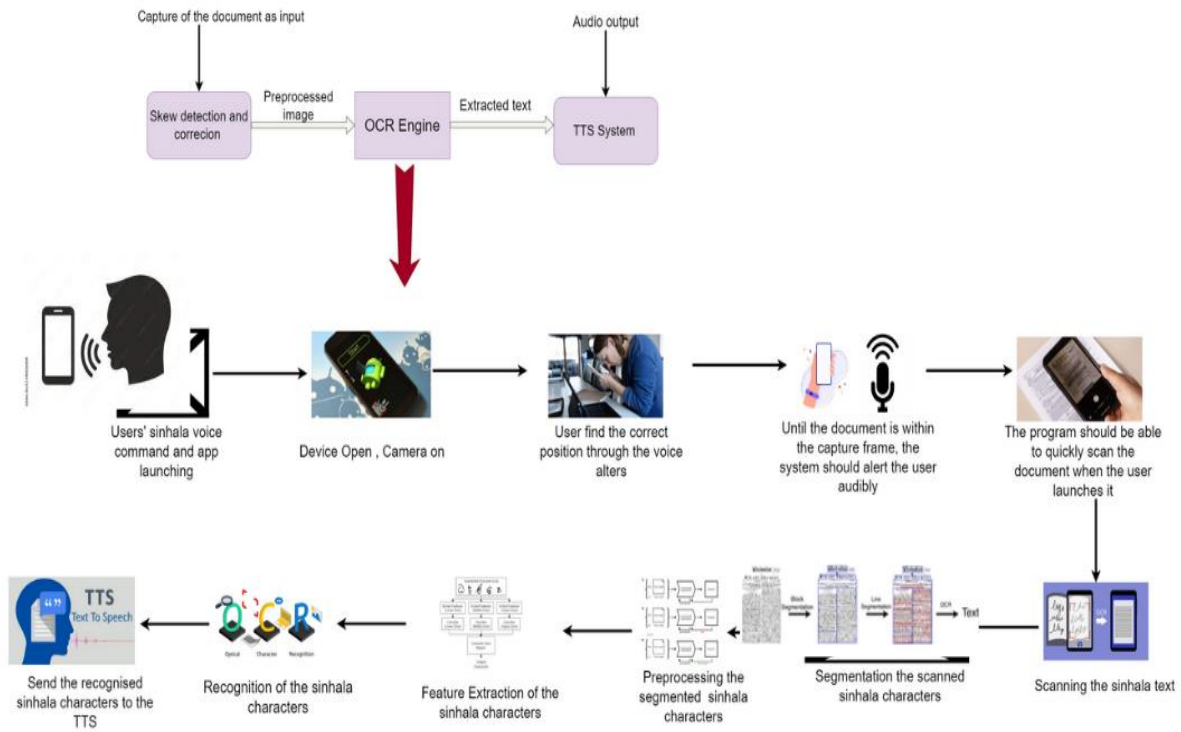


Figure 12-System diagram

7. Current Progress up to now

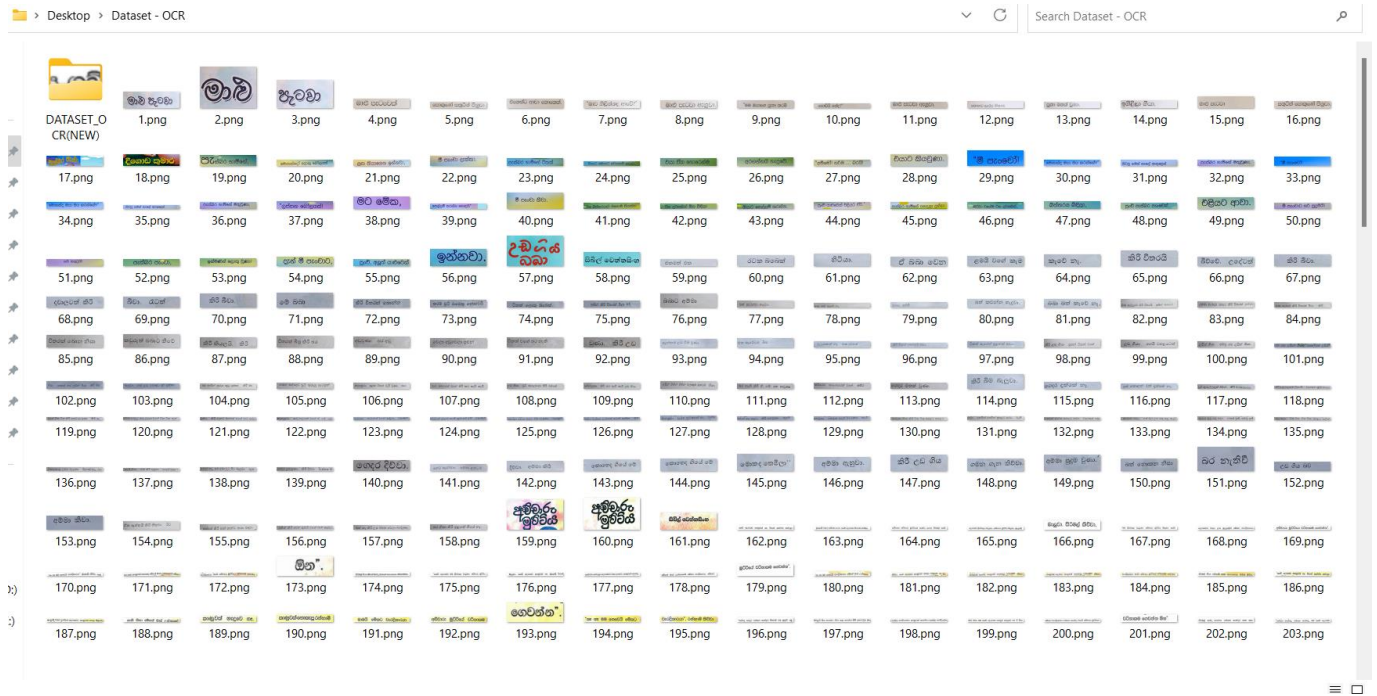
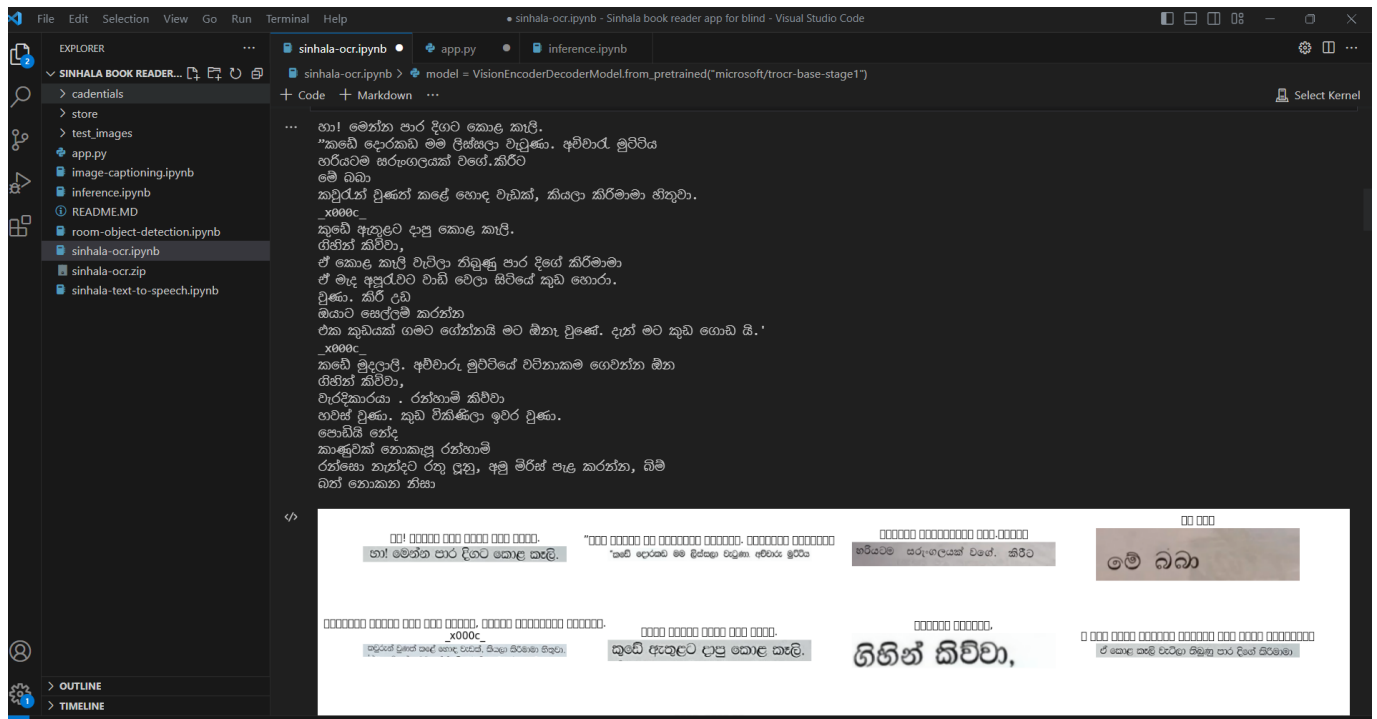


Figure 13-dataset



```

# pip install jiwer

import cv2, re
import warnings
import torch, io
import cv2 as cv
import numpy as np
import pandas as pd
from PIL import Image as ImagePIL
from google.cloud import vision
from datasets import load_metric
from torch.utils.data import Dataset
from matplotlib import pyplot as plt
from google.cloud.vision.v1.types import Image
from sklearn.model_selection import train_test_split
from google.oauth2.service_account import Credentials
from transformers import Seq2SeqTrainer, Seq2SeqTrainingArguments, \
    TrOCRProcessor, VisionEncoderDecoderModel, default_data_collator, AutoTokenizer

warnings.filterwarnings('ignore')

df = pd.read_excel('data/ocr/labels.xlsx')

df['ImageID'] = df.index + 1
df['ImageName'] = df['ImageID'].apply(lambda x: str(x) + '.png')
del df['ImageID']

# drop rows with empty text

```

```

trainer.train()

```

67% 1654/2476 [4:54:22<2:34:56, 11.31s/it]

```

{'loss': 4.7256, 'learning_rate': 9.192245557350567e-06, 'epoch': 0.08}
{'loss': 4.871, 'learning_rate': 8.990306946688207e-06, 'epoch': 0.1}
{'loss': 4.0503, 'learning_rate': 8.788368336025849e-06, 'epoch': 0.12}
{'loss': 4.1091, 'learning_rate': 8.58642972536349e-06, 'epoch': 0.14}
{'loss': 3.6878, 'learning_rate': 8.384491114701132e-06, 'epoch': 0.16}
{'loss': 4.3689, 'learning_rate': 8.182552504038774e-06, 'epoch': 0.18}
{'loss': 3.9379, 'learning_rate': 7.980613893376414e-06, 'epoch': 0.2}
{'loss': 3.5528, 'learning_rate': 7.778675282714055e-06, 'epoch': 0.22}
{'loss': 3.6716, 'learning_rate': 7.576736672051697e-06, 'epoch': 0.24}
{'loss': 3.6768, 'learning_rate': 7.374790061389338e-06, 'epoch': 0.26}
{'loss': 3.5433, 'learning_rate': 7.17285945072698e-06, 'epoch': 0.28}
{'loss': 3.5512, 'learning_rate': 6.970920840064621e-06, 'epoch': 0.3}
{'loss': 3.3651, 'learning_rate': 6.768982229402262e-06, 'epoch': 0.32}
{'loss': 3.609, 'learning_rate': 6.567043618739903e-06, 'epoch': 0.34}
{'loss': 3.0503, 'learning_rate': 6.365105000077545e-06, 'epoch': 0.36}
{'loss': 3.0532, 'learning_rate': 6.163166397415186e-06, 'epoch': 0.38}
{'loss': 3.0921, 'learning_rate': 5.961227786752828e-06, 'epoch': 0.4}
{'loss': 3.0626, 'learning_rate': 5.759289176090469e-06, 'epoch': 0.42}
{'loss': 2.8606, 'learning_rate': 5.5573505654281104e-06, 'epoch': 0.44}
{'loss': 3.0606, 'learning_rate': 5.355411954765751e-06, 'epoch': 0.46}
{'loss': 2.9124, 'learning_rate': 5.153473344103394e-06, 'epoch': 0.48}
{'loss': 2.6757, 'learning_rate': 4.951534733441034e-06, 'epoch': 0.5}
{'loss': 2.9099, 'learning_rate': 4.749596122778676e-06, 'epoch': 0.53}
{'loss': 2.9063, 'learning_rate': 4.547657512116317e-06, 'epoch': 0.55}
{'loss': 2.6489, 'learning_rate': 4.3457189014539584e-06, 'epoch': 0.57}
{'loss': 2.9069, 'learning_rate': 4.1437802897916e-06, 'epoch': 0.59}
{'loss': 2.9022, 'learning_rate': 3.941841680129241e-06, 'epoch': 0.61}
{'loss': 2.7671, 'learning_rate': 3.7399030694668824e-06, 'epoch': 0.63}
{'loss': 2.6325, 'learning_rate': 3.5379644588045236e-06, 'epoch': 0.65}
{'loss': 2.7164, 'learning_rate': 3.336025848142165e-06, 'epoch': 0.67}

```

Figure 14-dataset training

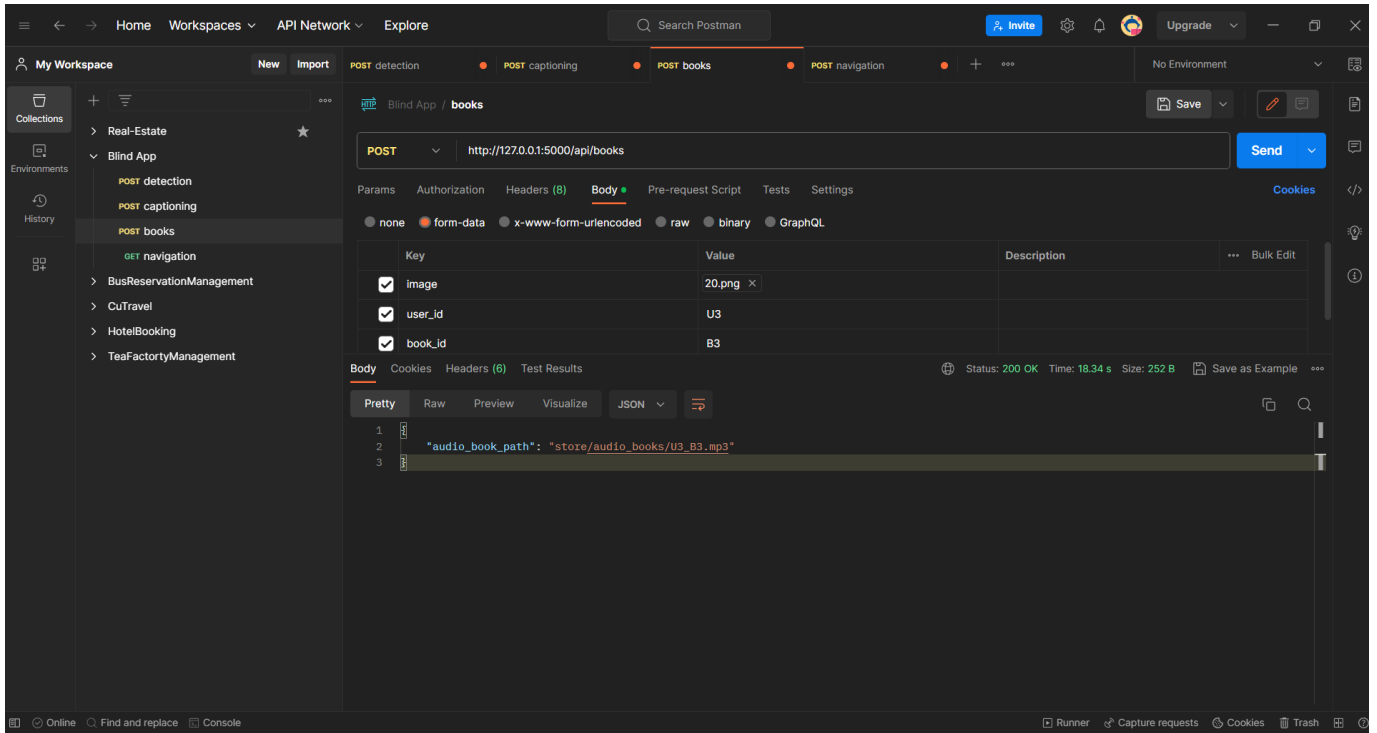
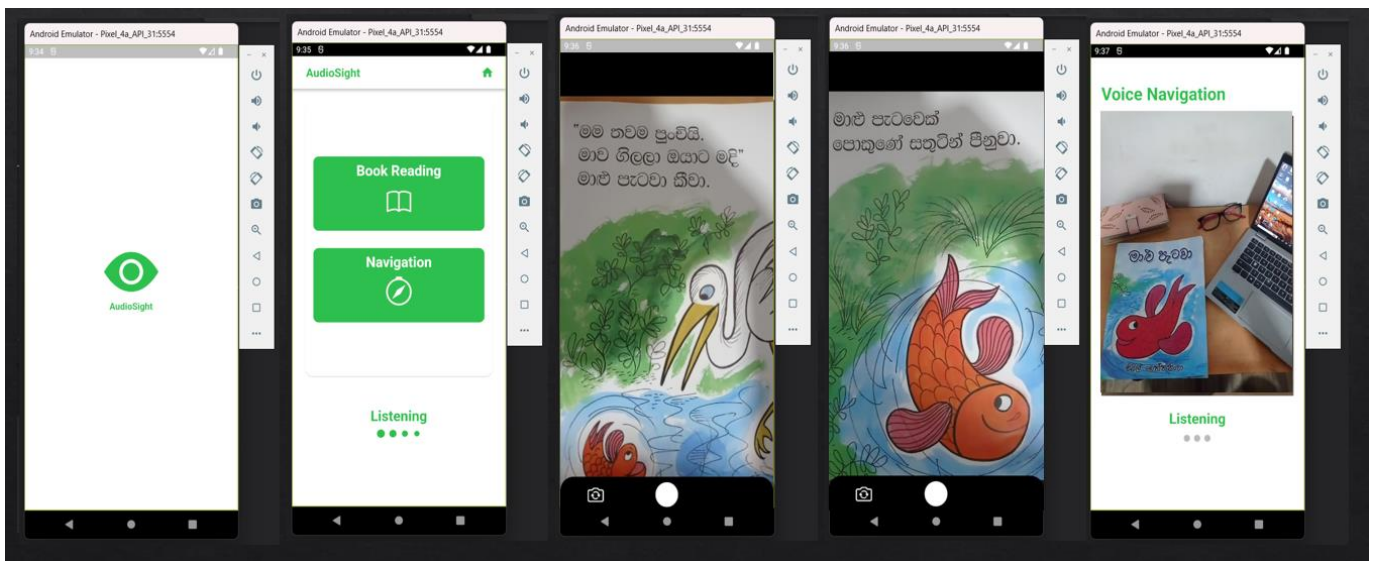


Figure 15-backend postman API



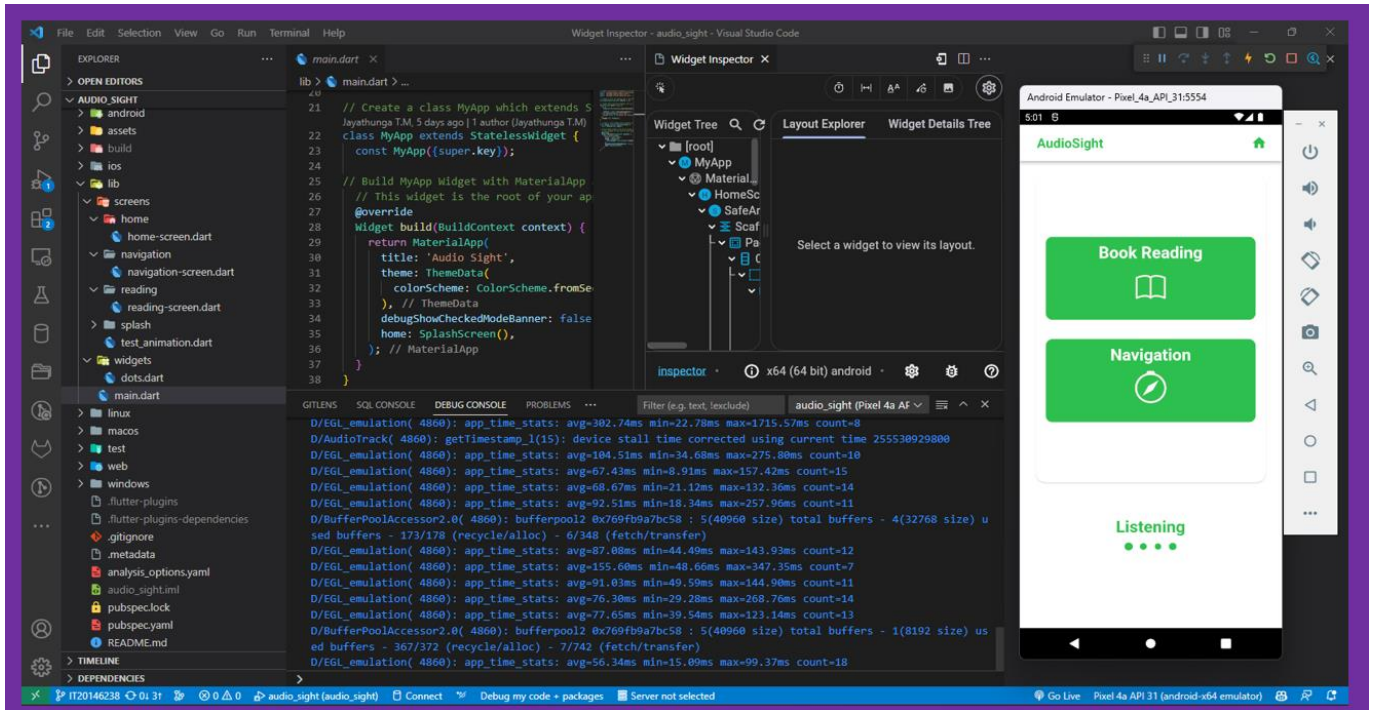


Figure 16-frontend

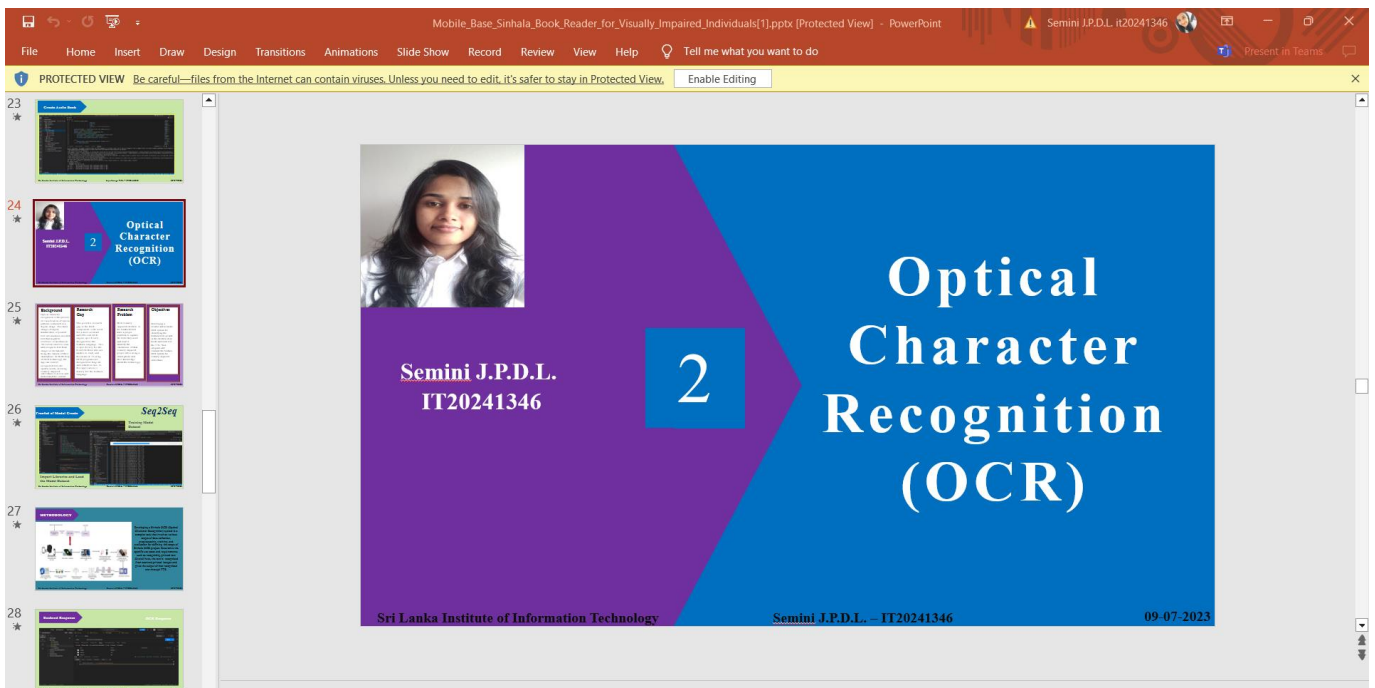


Figure 17-pp2 presentation

8. Our Research Paper

Mobile Base Sinhala Book Reader For Visually Impaired Students

J.Thirosh Madhusha
Faculty of Computing
Sri Lanka Institute of Information
Technology
Malabe, Sri Lanka
thiroshmadhusha0520@gmail.com

J.P.Dulmi Semini
Faculty of Computing
Sri Lanka Institute of Information
Technology
Malabe, Sri Lanka
dulmilaknasewmz9811@gmail.com

H.D.Madushi Bhagya
Faculty of Computing
Sri Lanka Institute of Information
Technology
Malabe, Sri Lanka
madushihd@gmail.com

Prof. Koliya Pulasinghe
Faculty of Computing
Sri Lanka Institute of Information
Technology
Malabe, Sri Lanka
koliya.p@slit.lk

Sajeewa Godakanda
Faculty of Computing
Sri Lanka Institute of Information
Technology
Malabe, Sri Lanka
sajeewagodakanda@gmail.com

P.K.P.G. Panduwawala
Faculty of Computing
Sri Lanka Institute of Information
Technology
Malabe, Sri Lanka
poorna.p@slit.lk

Abstract—The project aims to improve the reading experience and skills of visually impaired students in Sri Lanka by creating a mobile application that allows them to easily read printed books and stationery in Sinhala. The mobile application uses optical character recognition (OCR) technology and voice navigation, incorporating text-to-speech features of the event synthesis framework. The application accurately captures characters on a page of a Sinhala book and distinguishes them using OCR technology, enabling visually impaired people to convert text into accessible digital formats. The extracted text is then made audible via text-to-speech. Sinhala Voice Navigation support is provided for users to navigate the app, get feedback from the user, and identify objects in the surrounding room. The application uses image recognition and description algorithms to describe pictures in Sinhala, helping visually impaired children understand the visual content and improve their reading skills. The platform also offers features to adjust reading speed and choose between male or female voices.

Keywords—Visually Impaired Individuals, Sinhala Text-to-Speech(TTS), Sinhala Optical Character Recognize(OCR), Sinhala Voice Navigation, Image Recognition, Sinhala Object Detection

I. INTRODUCTION

Knowledge is the most important factor for surviving in this century. One way to gain knowledge is through reading, even for those who are visually impaired, who can use the braille system [1]. However, traditional braille systems are becoming outdated as computer-assisted braille systems and text-to-speech systems are becoming more common. Unfortunately, these technologies are not widely available in Sinhala, and these devices are too expensive for the average Sri Lankan. An Android-based solution using OCR, TTS, image recognition, and voice navigation was considered for this study to improve the reading experience and accessibility for the visually impaired "Sinhala Book Reader Mobile Application".

The quality of a Text-to-Speech (TTS) system depends on its ability to imitate human speech and ensure clear understanding. The absence of natural expressions in TTS

output has a substantial influence on application usability. This emphasizes a key issue in TTS development for creating a synthesized speech [2] that closely matches the human voice from the text. TTS technology's major goal is to recreate the complete range of human speech, including different speech patterns, subtleties, and intonations, while reducing the mechanical or robotic quality of the output voice.

The Sinhala language, the mother tongue of most Sri Lankans, is a crucial area for TTS development due to its complexities and nuances. Despite the large number of Sinhala speakers in Sri Lanka, there is a need for research on Sinhala voice recognition. The complexities of the Sinhala language make it difficult for computers to understand and reproduce it. Currently, there is little progress in developing TTS systems for the Sinhala language. However, this is a key research frontier that must be explored. An efficient TTS system for Sinhala would bridge the gap between human language skills and machine-generated speech [1], improving user experiences and bridging the gap between human language skills and machine-generated speech. There have been only a few attempts made to develop a Sinhala language TTS. This is still a major research area that requires investigation, which is one of the key motivations for this research.

In an increasingly digitized world, accessibility to information and literature remains a challenge for visually impaired individuals. Mobile technology and Optical Character Recognition (OCR) can solve this issue. This introduction elucidates the significance of mobile-based Sinhala book readers employing OCR technology as a transformative solution for individuals with visual impairments. By harnessing the power of mobile devices and OCR, these readers offer the potential to convert printed Sinhala text into accessible digital formats, thereby facilitating independent and inclusive access to literature for visually impaired individuals. This section introduces the key components of this paper, including the integration of OCR technology, the unique context of the Sinhala language, and the overarching goal of enhancing accessibility and enriching the reading experiences of visually impaired individuals [3].

Our Sinhala book reader app for blind users delivers a ground-breaking feature: picture detection within Sinhala children's books, at the nexus of accessibility and education. By allowing those who are blind to enjoy the rich world of

Figure 18-research paper

9. Extra helping resources

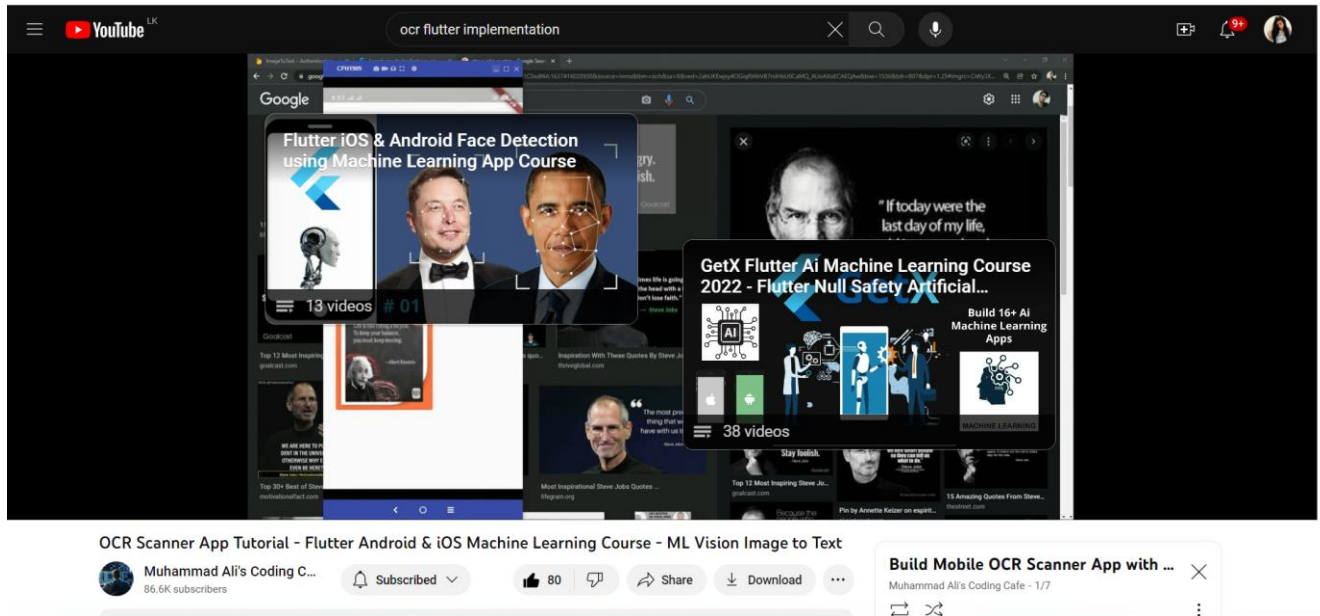


Figure 19-you tube tutorials

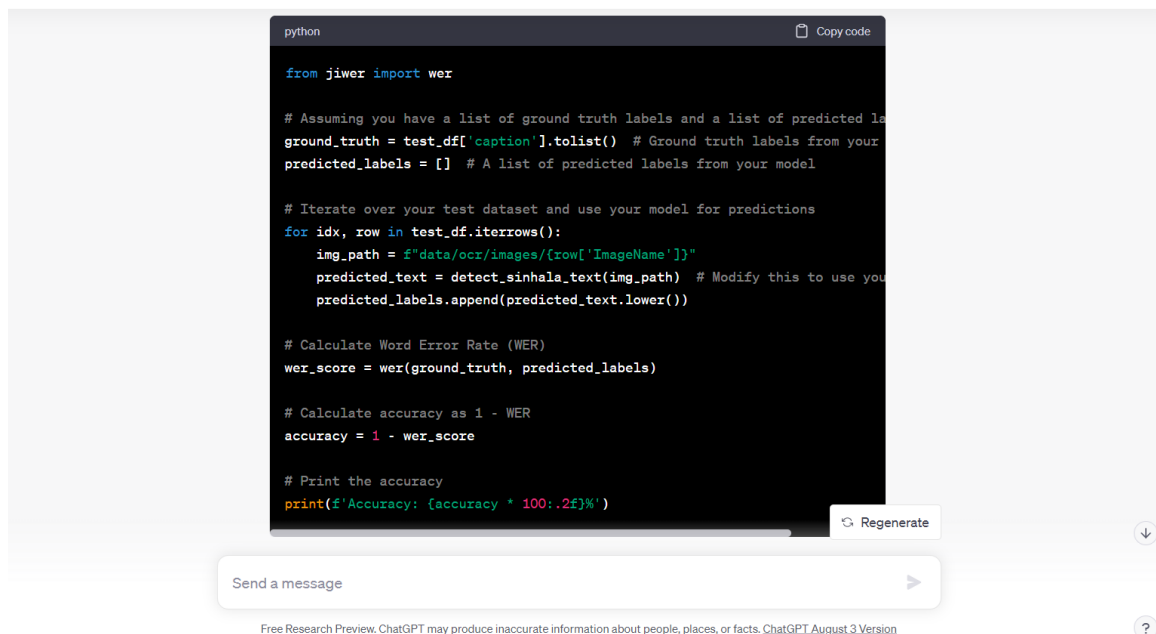


Figure 20-chat gpt

NLP Applications of Sinhala: TTS & OCR

Ruvan Weerasinghe, Asanka Wasala, Dulip Herath and Viraj Welgama

Language Technology Research Laboratory,
University of Colombo School of Computing,
35, Reid Avenue, Colombo 00700, Sri Lanka
{arw, raw, dlh, wvw}@ucsc.cmb.ac.lk

Abstract

This paper brings together the practical applications and the evaluation of the first Text-to-Speech (TTS) system for Sinhala using the Festival framework and an Optical Character Recognition system for Sinhala.

1 Introduction

Language Technology Research Laboratory[†] (LTRL) of the University of Colombo School of Computing (UCSC), was established in 2004 evolving from work engaged in by academics of the university since the early 1990's in local language computing in Sri Lanka.

Under the scope of the laboratory, numerous Natural Language Processing projects are being carried out with the relevant national bodies, international technology partners, local industry and the wider regional collaboration particularly within the PAN Localization Initiative^{*}. The Sri Lankan component of the PAN Localization Project concentrated on developing some of the fundamental resources needed for language processing and some

This paper focuses primarily on the end-user applications developed under the above project; Sinhala TTS system and OCR system. The paper describes the practical applications of these tools and evaluates it in the light of experience gained so far.

The rest of this paper is organized as follows: Section 2 gives an overview of the Sinhala TTS system; Section 3 describes the Sinhala OCR system. A summary along with future research directions and improvements are discussed in the last section.

2 Sinhala Text-to-Speech System

Sighted computer users spend a lot of time reading items on-screen to do their regular tasks such as checking email, fill out spreadsheets, gather information from internet, prepare and edit documents, and much more. However visually impaired people cannot perform these tasks without an assistance from other, or without using assistive technologies.

A TTS (text-to-speech) system takes computer text and converts the words into audible speech (Dutoit, 1997). With a TTS engine, application, and basic computer hardware, one can listen to computer text instead of reading it. A Screen Reader (2007) is a piece of software that attempts

Figure 21-research papers

The screenshot shows a web page with a navigation bar at the top containing links to various technologies: HTML, CSS, JAVASCRIPT, SQL, PYTHON (highlighted), JAVA, PHP, HOW TO, W3.CSS, C, C++, C#, BOOTSTRAP, REACT, MYSQL, JQUERY, EXCEL, XML, DJANGO, and NODE. On the left, a sidebar lists categories like Python Arrays, Python Classes/Objects, Python Inheritance, Python Iterators, Python Polymorphism, Python Scope, Python Modules, Python Dates, Python Math, Python JSON, Python RegEx, Python PIP, Python Try...Except, Python User Input, Python String Formatting, File Handling, Python File Handling, Python Read Files (highlighted), Python Write/Create Files, Python Delete Files, and Python Modules. The main content area is titled 'Open a File on the Server' and contains the following text: 'Assume we have the following file, located in the same folder as Python:'. Below this is a code block for 'demofile.txt' containing: 'Hello! Welcome to demofile.txt', 'This file is for testing purposes.', and 'Good Luck!'. The text continues: 'To open the file, use the built-in open() function.' and 'The open() function returns a file object, which has a read() method for reading the content of the file:'. An 'Example' section shows the code: 'f = open("demofile.txt", "r")' and 'print(f.read())'. A 'Run Example »' button is below the code. A 'Get your own Python Server' button is also present. On the right, there are two promotional banners: one for 'with ultra low spreads' with a 'Trade Now' button, and another for 'Ready for the next step in your career? Get Python Certified!'.

Figure 22-W3 School