# Capstone Project

# VIRTUAL EYE: A Perfect Companion for Blind

## Submitted By:


101967003 Swapnil Kumar Singh

101953014 Raju Kumar Gupta

102083026 Riya Singhal

101903527 Aditi


**BE Third Year- COE**

**CPG No. 180**

Under the Mentorship of

Maninder Singh

and

Sanmeet Kaur



**Computer Science and Engineering Department**

**Thapar Institute of Engineering and Technology, Patiala**


**Feb-Dec 2022**

# ABSTRACT

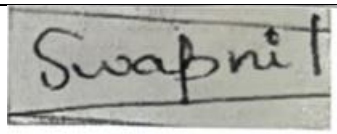
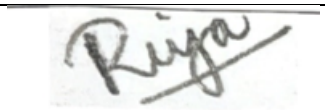
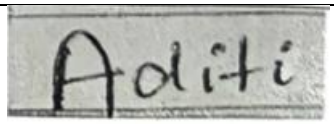
Eyesight is one of the most important sense organs needed to analyses one's surroundings and being deprived of this natural sensor fills one's life with struggles. These people have to deal with various issues regarding managing their day-to-day activities. Visually impaired people also deal with issues related to self-esteem. Unwanted favors provided to these people may also hurt their confidence as they may take that help as a form of sympathy provided to them. Empowering them through this application would not only help them obtain independence but also make them confident as they won't be looking for help every now and then. For visually challenged people it's difficult to find and recognize their belongings and things in their surrounding environment. This task may frequently require them to ask for help which may hurt their self-esteem and it's also always not possible for them to obtain some helping hand from people around them. The Virtual Eye will identify all surrounding objects for the visually impaired person. Next comes the difficulty faced in reading normal text present in the surrounding like advertisements, posters and digital text documents like pdfs. The Virtual Eye would scan these texts and convert them to audio queues to aid them by making use of hearing power.

Although, a lot many solutions exist already, but still the visually impaired people find difficulties. None of the researched applications are 100% complete in themselves. Some of the applications have difficult to handle user interface while others have issues of accuracy and speed. The applications should have decent accuracy so that the person can completely rely on the application's observations and results. Some applications are very slow in detecting objects while other are so speedy in reading the text that it becomes difficult to interpret their words. Document readers can't be reminded of the point where it was left earlier, making it difficult for especially abled person to read virtual books or other lengthy documents. In "Virtual Eye", we have tried to cover all of the loopholes of the previous works.

# DECLARATION

We hereby declare that the design principles and working prototype model of the project entitled – 'Virtual Eye: A perfect Companion for Blind' is an authentic record of our own work carried out in the Computer Science and Engineering Department, TIET, Patiala under the guidance of Maninder Singh during 6th and 7th semester 2022.

Date: 25/08/2022

| Roll No. | Name | Signature |
|---|---|---|
| 101967003 | Swapnil Kumar Singh |  |
| 101953014 | Raju Kumar Gupta |  |
| 102083026 | Riya Singhal |  |
| 101903527 | Aditi |  |

**Name: of Mentor: Maninder Singh**

**Signature of Mentor:** Maninder Singh Digitally signed by Maninder Singh Date: 2022.08.26 17:22:36 +05'30'

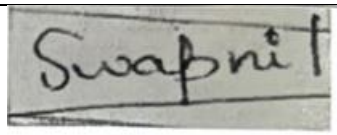
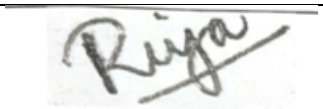
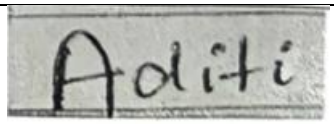
# ACKNOWLEDGEMENT

We would like to express our thanks to our mentor Maninder Singh and Sanmeet Kaur. He has been of great help in our venture, and an indispensable resource of technical knowledge. He is truly an amazing mentor to have.

We are also thankful to Shalini Batra, Head Computer Science and Engineering Department, entire faculty and staff of Computer Science and Engineering Department, and also our friends who devolved their valuable time and helped us in all possible ways towards successful completion of this project. We thank all those who have contributed either directly or indirectly towards this project.

Lastly, we would also thank our families for their unyielding love and encouragement. They always wanted the best for us and we admire their determination and sacrifice.

Date: 25/08.2022

| Roll No. | Name | Signature |
|---|---|---|
| 101967003 | Swapnil Kumar Singh |  |
| 101953014 | Raju Kumar Gupta |  |
| 102083026 | Riya Singhal |  |
| 101903527 | Aditi |  |

# TABLE OF CONTENTS

## LIST OF TABLES

# LIST OF FIGURES

## LIST OF ABBREVIATIONS

| Abbreviations | Captions |
|---|---|
| TIET | Thapar Institute of Engineering and Technology |
| OCR | Optical Character Recognition |
| WHO | World Health Organization |
| TTS | Text to Speech |
| AI | Artificial Intelligence |
| SSD | Single Shot Detector |
| HTML | Hyper Text Markup Language |
| CSS | Cascading Style Sheet |
| JS | Java-Script |
| WIFI | Wireless Fidelity |
| IEEE | Institute of Electrical and Electronics Engineers |
| ML | Machine Learning |
| TCP | Transmission Control Protocol |
| SDK | Software Development Kit |
| API | Application Programming Interface |
| MSVI | Moderate to Several Visual Impairment |
| FIG | Figure |

# INTRODUCTION

## 1.1.Project Overview

According to estimates from World Health Organization (WHO) about 285 million people are visually impaired worldwide: 39 million are blind and 246 million have low vision (severe or moderate visual impairment). We perceive up to 80 per cent of all impressions by means of our sight. This makes eyes one of the most vital sense organs thereby raising serious concerns regarding eye heath for people around the globe. This project is created by us as an attempt to develop an easy-to-use application that provides basic necessary features while overcoming some of the barriers faced by people in current technology. These features are:

Object detection
Scan text
Scan documents
Navigating through path

### 1.1.1 Problem Statement

There are three main level of difficulties when it comes to disabilities. First includes physical limitation of reading and moving around freely using eyesight for which people usually undergo Braille and mobility trailing to overcome this barrier. But these methods can only compensate them for their loss to some extent and may prove to be useless in some scenarios like reading text which are not available in Braille, moving to some unfamiliar place where they can't make use of their mobility training and so on. In today's age of digitalization, a lot of things are being digitalized which raises concerns for these people. Digital documents are now more dominant than text documents. Also, a lot of text like advertisements, posters etc. can't be all made available in Braille which creates a social barrier. Second limitation is environment oriented like inaccessible transportation, discrimination, social stigma. Third is medical limitation which includes requirement of extra support by these people. But providing unwanted support can make these people interpret some kind help as sympathy and

lead to misunderstanding raising issues like low self-confidence and low self-esteem. The best way to avoid this is to make them move towards greater self-independence which can be achieved by the use of technology. Having this application in their hand they won't have to search for helping hands every time and will become independent in executing their day-to-day activities boosting their self-confidence.

### 1.1.2 Goal

The main goal of the project, virtual eye is to provide an efficient visual aid for visually impaired people which gives a sense of artificial vision by providing information about the environmental scenario through voice output and vibration. Virtual Eye will be designed to help these people navigate through their day and get some of their daily necessities done easily. This will empower them towards greater independence and boost their self-confidence and morale. Through this application we aim to aid their abilities with technology.

### 1.1.3 Solution

This application will combine several technologies like OCR, object recognition, Flutter to provide various features as:

#### 1.1.3.1 Scanning text and documents

This feature will aid in overcoming physical limitation of reading. If a person wants to scan a document/text, then he/she would be guided through voice informing about the uncovered corners and auto focus would help detect the text with much ease. On successful scanning of the doc, the application would read it out loud. If a person wants to read a doc from device, then he/she will be guided to folders containing the docs and can get the doc from there. The application would also keep a note of the point where the person left reading the same doc earlier so that he/she can resume reading as well.

#### 1.1.3.2 Object recognition

This feature will aid in overcoming physical limitation of mobility. Since this feature will allow user to recognize objects in the surrounding the

person will be able to get a sense of his surrounding by locating objects around him and also making it easier to move around. If a person wants to find out what objects are present in his/her surroundings, then he/she can turn the camera to scan the surrounding and the application would process the data and cite out the name of objects found. Similarly, if a person is in search of a specific object, then he/she would need to provide the name of object before scanning. Once object is found, the application would generate some indication (vibration or a voice note).

#### 1.1.3.3 Navigation

This feature will also aid in overcoming physical limitation of mobility. Whenever they would require to go to some destination, the application would provide helping eyes to guide him/her throughout the path. For this he/she would turn on the navigation mode and with the help of the stick, he/she would be guided so that he/she doesn't strike any obstacle and would tell the best turn he/she should take to avoid any obstacle till maximum possible time.

#### 1.1.3.4 Voice assistant

This feature is to aid users to make use of the application conveniently with the use of voice and audio queues which will be much more convenient for them compared to prevalent touchscreen operated applications which completely rely on eyesight. During the whole stay of the person on the application, he/she would be guided to various channels/tabs of the application through voice assistant so that the person is able to operate the application easily.

All this feature collectively also contributes towards solving medical limitation as people will be able to obtain help quickly through this application whenever they want help.

## 1.2.Need Analysis

Visually challenged people already have a lot on their plates to handle. These people have to deal with various issues regarding managing their day-to-day

activities. Visually impaired people also deal with issues related to self-esteem. Unwanted favors provided to these people may also hurt their confidence as they may take that help as a form of sympathy provided to them. Empowering them through this application would not only help them obtain independence but also make them confident as they won't be looking for help every now and then. For visually challenged people it's difficult to find and recognize their belongings and things in their surrounding environment. This task may frequently require them to ask for help which may hurt their self-esteem and it's also always not possible for them to obtain some helping hand from people around them. The Virtual Eye will identify all surrounding objects for the visually impaired person. Next comes the difficulty faced in reading normal text present in the surrounding like advertisements, posters and digital text documents like pdfs. The Virtual Eye would scan these texts and convert them to audio queues to aid them by making use of hearing power.

Although, a lot many solutions exist already, but still the visually impaired people find difficulties. None of the researched applications are 100% complete in themselves. Some of the applications have difficult to handle user interface while others have issues of accuracy and speed. The applications should have decent accuracy so that the person can completely rely on the application's observations and results. Some applications are very slow in detecting objects while other are so speedy in reading the text that it becomes difficult to interpret their words. Document readers can't be reminded of the point where it was left earlier, making it difficult for especially abled person to read virtual books or other lengthy documents. In "Virtual Eye", we would try to cover all of the loopholes of the previous works

## 1.3.Research Gaps

The widespread use of mobile applications, developers and designers have led to existence of various applications and assistive technologies available for the visually impaired people to assist them with various day to day activities. Currently these applications vary in their usability. The usability of application refers to measure of time and effort required to perform a task, efficiency in

performing the required task and resultant user satisfaction. While going through the portion of literature, areas to improve the usability were noticed.

1. **Screen layout**

Most application require user to locate buttons or visual items on touch screen interfaces like smartphones and tablets to switch between various functionalities of the application. This is quite challenging for these people since they have no idea of what functionality is located where and how to enable them. They have to solely rely on the documentation of the application or trying and testing by clicking various portions of screen or based on realizations by accidentally touching a screen element and discoveries afterwards. This makes such interfaces less user friendly and inconvenient to use for new users. But these can sometimes mislead them. There is a need for developing more usable interface for blind people.

2. **Insufficient audio guidance**

Most of the mainstream applications provide almost no audio guidance or inadequate audio guidance. Some have issues in audio quality – speed, clarity. Inefficient use of non-visual output modalities like vice output and vibration feedback. Interacting with applications using visual elements is already way difficult and interacting with talkback audio which is unclear and inefficient to use does not help much.

3. **Noise cancellation**

Current TTS AI models don’t support efficient noise cancellation. It is hard to differentiate between user commands and environment noise for current TTS models. There are only few solutions existing which includes some wearable headphone that incorporates training-based noise cancellation technique. But these kinds of devices are not convenient to use on the long run and can’t be used by visually impaired people for a long time. This puts a limit for these people since it concerns usability of software.

S Rashmi et al 2018 J. Phys.: Conf. Ser. 1142 012019

**4. Object detection**

The accuracy for object detection in practical is found to be not very efficient and accurate. We have tried and tested various models available like tflite, yolo, mobile-net SSD which are said to be some of the most accurate models available to date but none of these result in good accuracy.

**5. Voice base application**

Current mainstream applications are almost completely based on touch mechanism. Completely voiceover mechanism is not yet implemented completely on applications. Some applications do support voice assistant technology but only to a very limited extent. Even the applications are are specially tailored for blind people rely on visual elements and don't support voice-based interface. Research for usability of applications using voice assistant is still lagging

## 1.4.Problem Definition and Scope

Most of the mainstream applications provide limited functionality to the user. Most applications are only capable to address one or two tasks- predominantly identifying objects or scanning things. Most of them provide bad or inefficient usability by providing a lot of visual elements on the interface to the user which is very inconvenient and not user friendly for visually impaired people.

They have to rely on their ability to discover and understand the interface and functionalities of the application through documentation, exploring the application or accidental discoveries.

Considering these limitations, the scope of this project is as follows.

1. Identifying objects in the environment to keep the user aware of its environment and help identify and locate objects easily in a unfamiliar environment. It will identify both specific and all objects found around the person.
2. Scanning text and document to make it easy for the user to go through digital text without installing any other screen reader applications.

3. Navigating through the path to help user move around freely and confidently in an unfamiliar environment by identifying and locating objects which are close to the user.

4. AI voice powered application which will increase the usability of our application and make the application user-friendly and convenient to use since they won't need to rely a lot on visual elements on the screen and will be guided through voice commands.

## 1.5.Assumptions and Constraints

Table 1 shows the assumptions considered for the “Virtual Eye” project.

Table 1: Assumptions for the “Virtual Eye” project.

| S.NO. | Assumptions |
|---|---|
| 1. | It is assumed that the user is not audibly impaired. |
| 2. | Since a major part of the functionality of the project is voice-dependent and as the model is trained only for English, hence it is assumed that the user is comfortable in speaking, reading and writing English language. |
| 3. | It is assumed that the user is well aware of the fact that there might be chances where a user may encounter different pronunciation of the same word. |
| 4. | Users must have decent internet connectivity for making use of functionalities that use Internet connection. |
| 5. | The device employed by the user must meet required hardware and software requirements for the application to work properly. For example, the device must possess a modern browser to run advanced HTML, CSS, JS scripts, sufficient memory to download the application etc. |

Table 2 shows the constraints for the project.

Table 2: Constraints for the "Virtual Eye" Project

| S.NO. | Constraints |
|---|---|
| 1. | The system will only support only English language for voice communication. |
| 2. | The system will need good internet connection for some features to execute. |
| 3. | The object detection functionality limits object distance to a specific range. |
| 4. | The text being scanned must be clear. Unclear text will negatively affect accuracy. |
| 5. | The blind stick must be controlled properly (stability of orientation and speed). |

## 1.6.Standards

The standards used in this project are:

HTTP: This is Hyper Text Transmission Protocol which specifies how data is encoded, delivered, and how Web servers and browsers will take to respond to different commands. It is the fundamental protocol employed by the World Wide Web.

Wi-Fi: Wi-Fi is a popular wireless networking system which provides high-speed wireless Internet and network connectivity using radio waves. Wi-Fi is a term labeled with IEEE 802.11x definition.

## 1.7.Approved Objectives

To create a stick which would identify the distance from obstacles in user's path.

- To create a module which would detect all objects present in the surroundings of the user.

- To create a module which would detect specific objects in the surrounding of the user.
- To create a module which would scan text from the surroundings and read it aloud to the user.
- To create a module which would scan text of the documents from the local storage of phone and read it aloud to the user.
- To integrate above all in an application using voice assistant.

## 1.8.Methodology

1. Voice assistant is built using Alan studio's voice assistant SDK for flutter.
2. Object detection module which is used to find/spot a specific object in the environment, identify all objects being captured in the environment, is built using the tflite package and a pre-trained SSD-Mobile Net model.
3. Recognition of text from image or document is done using ML kit of google firebase.
4. For the stick used for navigating through path, we would use ultrasonic sensor for detecting distance from obstacles present at a distance less than one meter and would send it to app using Wi-Fi module and the Alan voice assistant would be used to send the information to the user in voice format.

## 1.9.Project Outcomes and Deliverables

The outcomes to expect at the completion of the project is a final product.

Virtual Eye- A Perfect Companion for Blind promises the following:

1. Object Detection:

a. The user's speech in real-time will be heard by an emotionally intelligent bot in real-time.

b. User would be able to find all objects in his/her surroundings as well as he/she could focus on finding a specific object also.

2. Text to Speech Conversion:

a. The application can help the user in reading soft copies present in his/her phone's local storage.

b. The application would also help user in reading text present in the user's surroundings like: a hoarding board etc.

3. Voice Assistant:

a. The user's speech would be heard in real-time and be used as a command to navigate through various options in the application.

b. The user would get every output or instruction in audio format.

## 1.10. Novelty of Work

Most of the applications existing today are limited to one or two features and touch screen operated leading to decreased usability of using these applications and the inconvenience faced by visually impaired people. The decreased usability of applications reduces their quality and importance as well. They make the applications less user-friendly. Most of the mainstream applications today do provide solutions to basic necessities of visually impaired people, but they don't address the efficiency of how these solutions are being addressed. And this efficiency is not very good since touch-based interface creates a lot of confusion and misleading to people who can't view visual elements and is very time and effort taking to understand these applications for these people. Also, most applications provide singular features which requires the user to switch to other applications. In our application we have tried to provide all the basic necessities to the user in a single application. Also, our application incorporates voice interfacing to increase usability of application, and tries tackle and remove the dependency on visual elements in user independence. All-together our application aims to provide all the necessary features that are scan text, scan document, object detection and navigating path with more user-friendly delivery, convenience and ease of use to improve user satisfaction and experience in using our applications. Better usability, user satisfaction, bundled appropriate features are key components of an application which most of the current applications are lacking which we aim to improve and incorporate efficiently.

## 2.1. Literature Survey

The literature survey includes the current knowledge including substantive findings, as well as theoretical and methodology contributes to the topic.

### 2.1.1 Theory Associated with Problem Area

Eyesight is one of the most important sense organs needed to analyses one's surroundings and being deprived of this natural sensor fills one's life with struggles. These people have to deal with various issues regarding managing their day-to-day activities. Visually impaired people also deal with issues related to self-esteem. Unwanted favors provided to these people may also hurt their confidence as they may take that help as a form of sympathy provided to them. Empowering them through this application would not only help them obtain independence but also make them confident as they won't be looking for help every now and then. For visually challenged people it's difficult to find and recognize their belongings and things in their surrounding environment. This task may frequently require them to ask for help which may hurt their self-esteem and it's also always not possible for them to obtain some helping hand from people around them. The Virtual Eye will identify all surrounding objects for the visually impaired person. Next comes the difficulty faced in reading normal text present in the surrounding like advertisements, posters and digital text documents like pdfs. The Virtual Eye would scan these texts and convert them to audio queues to aid them by making use of hearing power.

Although, a lot many solutions exist already, but still the visually impaired people find difficulties. None of the researched applications are 100% complete in themselves. Some of the applications have difficult to handle user interface while others have issues of accuracy and speed. The applications should have decent accuracy so that the person can completely rely on the application's observations and results. Some applications are very slow in detecting objects while other are so

speedy in reading the text that it becomes difficult to interpret their words. Document readers can't be reminded of the point where it was left earlier, making it difficult for especially abled person to read virtual books or other lengthy documents. In "Virtual Eye", we have tried to cover all of the loopholes of the previous works.

### 2.1.2 Existing System and Solutions

Numerous companies are working on building a digital help for blind which can make their work easier. Example of such systems are: Envision AI, Tap Tap See, Seeing AI etc. Owing to this competition, we have a completely voice assisted application which would not only find objects in the surroundings but also help the user navigate through various paths. The major issue in previously existing works is that they are mostly dependent on text and hence are difficult for the visually impaired to use them. While, a few of them also have very low accuracy making them negligibly reliant. Also, we found nearly zero applications which have integrated hardware helping the blind to navigate through paths, with software. So, our main aim is to remove all of these shortcomings of the previously existing software.

### 2.1.3 Research Findings for Existing Literature

Table 3: Research Findings

| S. No. | Roll Number | Name | Paper Title | Tools/ Technology | Findings | Citation |
|---|---|---|---|---|---|---|
| 1 | 102083026 | Riya Singhal | A Conversational News Application Project using Artificial Intelligence based Voice Assistance | Alan AI Studio | Found that Alan AI can be used for voice over interface in an application | Mahajan, Sameer & Kulkarni, Nahush. (2020). IRJET \| A Conversational News Application Project using Artificial Intelligence based Voice Assistance. 7. 3779-3782. |
| 2 | 101903527 | Aditi | PicToText: Text Recognition using Firebase Machine Learning Kit | OCR, Firebase, ML kit | Finding one of the most used text recognition techniques with fairly accurate results | D. K. S. Neoh et al., "PicToText: Text Recognition using Firebase Machine Learning Kit," 2021 2nd International Conference on Artificial Intelligence and Data Sciences (AiDAS), 2021, pp. 1-5, doi: 10.1109/AiDAS53897.2021.9574187. |
| 3 | 101967003 | Swapnil | Object Detection Based on YOLO Network | YOLO model | Found an object detection model (YOLO model) to be one of the models giving one of the highest accuracy available. | C. Liu, Y. Tao, J. Liang, K. Li and Y. Chen, "Object Detection Based on YOLO Network," 2018 IEEE 4th Information Technology and Mechatronics Engineering Conference (ITOEC), 2018, pp. 799-803, doi: 10.1109/ITOEC.2018.8740604. |
| 4 | 101953014 | Raju | MAKING TOUCH-BASED MOBILE PHONES ACCESSIBLE FOR THE VISUALLY IMPAIRED | Touch based interfaces | Finding inconvenience and challenges faced by users in operating touch based applications | [5] S. Khusro, B. Niazi, A. Khan and I. Alam, "Evaluating Smartphone Screen Divisions for Designing Blind-Friendly Touch-Based Interfaces," 2019 International Conference on Frontiers of Information Technology (FIT), 2019, pp. 328-3285, doi: 10.1109/FIT47737.2019 |

### 2.1.4. Problem Identified

The problems identified are as follows –

- There are no such applications which have completely removed their dependence on the ability of user to read or touch the screen to locate the icons.

- There is a lack of accuracy in identifying the objects in multiple applications which were surveyed.
- There is lack of accuracy in identifying the text while converting text-to-speech as well. Also, in few applications, text is not at all identified.

- There is no such application which compiled hardware support along with software to make the travel easier for blind. Either there were software applications or hardware only applications.
- There were almost no applications in which one could navigate through voice commands only.

### 2.1.5 Survey of Tools and Technologies Used

The tools and technologies that have been used are as follows –

1. Object Detection

- Camera | Flutter package
- tflite | Flutter package
    1. SSD Mobile-net
    2. YOLOv2
    3. Mobile-Net
    4. Pose-Net

2. Text detection

- Google ML kit
- Firebase

3. Navigation through path

- Arduino UNO
- Ultrasonic Sensor
- Firebase

4. Voice Based Assistant

- Alan Voice SDK kit

## 2.2. Software Requirement Survey

### 2.2.1 Introduction

This section deals with the Software Requirements Specification of the platform.

#### 2.2.1.1 Purpose

The purpose of this document is to identify our project, "**VIRTUAL EYE:** A Perfect Companion for Blind" a model that is aimed to help visually impaired people analyze the world around them with much more ease. Our model will help in identifying objects in the surroundings and respond by voice about the object found. Our application will also help in navigating through path by detecting all obstacles in the path. Our model will also help to bring a more audio-based navigation/assistant system so that it is handy to use for the blind.

#### 2.2.1.2 Intended Audience and Reading Suggestions

This project is aimed to help visually impaired people which includes completely blind as well as people with moderate to severe visual impairment. As per research, in 2015, there were an estimated 253 million people with visual impairment worldwide. Of these, **36 million** were blind and a further 217 million had moderate to severe visual impairment (MSVI). While in India, there are over 18 million blind people. Keeping in mind the language barrier, Virtual Eye can only help the English speaking people which account for about 1.35 billion people in this world.

#### 2.2.1.3 Project Scope

The scope of this project is as follows:

1. Identifying objects in the environment to keep the user aware of its environment and help identify and locate objects easily in an unfamiliar environment. It will identify both specific and all objects found around the person.

2. Scanning text and document to make it easy for the user to go through digital text without installing any other screen reader applications.

3. Navigating through the path to help user move around freely and confidently in an unfamiliar environment by identifying and locating objects which are close to the user.

4. AI voice powered application which will increase the usability of our application and make the application user-friendly and convenient to use since they won't need to rely a lot on visual elements on the screen and will be guided through voice commands.

### 2.2.2 Overall Description

This section deals with the overall description of the platform.

#### 2.2.2.1 Product Perspective

As we are conquering big challenges using technologies, somehow, we are still lacking in solving problems faced by visually impaired people to an optimal level. We aim to provide a platform where we are not only helping people find a safe path and find objects in their surroundings, but also give them the power to understand the text present in their surroundings which can't be converted to braille, like: posters, advertisements. We aim to reduce the struggle blind people face in understanding the user interface of various applications built for them and make it easier to use. This could become one step towards making the visually impaired more independent and confident and reduce their struggle to a large extent.

#### 2.2.2.2 Product Features

Our project aims to provide the following features:

1. A system that will perform object detection (identify all objects and identify specific object).
2. A model which can recognize user's commands and work accordingly and also give information in audio form itself.
3. A system which can scan text from various documents present in phone's

local storage and read it aloud for the user.

4. A system which can identify text present in the user's environment by clicking picture and analysing it.
5. A blind stick with ultrasonic sensors which would identify the distance of the user from various obstacles and inform the user using audio commands.

### 2.2.3 External Interface Requirements

External interface requirements mention the hardware, software or the database system with which the project and its various components must interface with. It provides essential information which is required for the interfacing with the external environment

#### 2.2.3.1 User Interface

- The user interface will be in the form of a mobile application.
- The user can scan document/photo and convert the text to speech
- The user can detect all objects in the surroundings or a specific object in the surrounding.
- The user can use the application to walk in his/her environment safely by identifying the distances from various obstacles using virtual eye.

#### 2.2.3.2 Hardware Interfaces

- A smartphone with internet connectivity will be required to run the mobile application.
- The device to be used must have a microphone to record user audio and a screen/speaker to output the response to the user.
- The device needs to have a good quality camera to capture the surroundings of the user.
- The user must carry the stick to make use of the navigate through path module.

#### 2.2.3.3 Software Interfaces

- A mobile application to interact with the user.
- ALAN-AI to identify the audio input from user and generate a suitable response.
- Google ML kit is used for text recognition in scan text feature.
- Tiny YOLOv2, SSD Mobile-Net, Pose-Net are majorly used for object detection.

### 2.2.4 Other Non-functional Requirements

These are those requirements which don't have a direct effect on functionality of our project but are in some way required for building an efficient and optimized model.

#### 2.2.4.1 Performance Requirements

- The system should be efficient and reliable which would reduce the user's time and effort. In this way the user would more likely use our application if it meets his real time needs.
- The system will generate responses in a reasonable amount of time.

#### 2.2.4.2 Safety Requirements

- The System should be reliable and safe from external threats, bugs and hacks as some important documents which user scans might be hacked or leaked through online sources.
- Only authorised personnel and the users should be able to access the admin privileges.

#### 2.2.4.3 Security Requirements

- User's information should not be shared with any third party.
- A secure database should be used.

## 2.3. Cost Analysis

Table 4: Cost Analysis

| S.no | Product Name | Quantity | Price/unit | Total Price (INR) |
|---|---|---|---|---|
| **1.** | Arduino UNO Wifi Rev2 | 1 | Rs. 4299/- | Rs. 980/- |
| **2.** | Ultrasonic sensor | 2 | Rs. 130/- | Rs. 260/- |
| **3.** | Stick | 1 | Rs. 310/- | Rs. 310/- |
| **4.** | Coaxial cable | 1 | Rs. 70/- | Rs. 70/- |
| **5.** | Battery | 2 | Rs. 225/- | Rs. 450/- |
| **6.** | Breadboard | 1 | Rs. 90/- | Rs. 90/- |
| **7.** | Connecting wires + resistors+ capacitors | 1pack | Rs. 99/- | Rs. 99/- |
| | Alan studio subscription | 1500 | Rs. 1/- | 1500 |
| **Total** | - | - | - | Rs. 3759/- |

## 2.4. Risk Analysis

Though our project takes care of every casualty, certainly there can be many flaws and further risks involved if not checked every part of our project's functioning.

- Quite loud noise can be present in the room, which can hinder the functioning of the speech recognition system. Due to this the model might not be able to recognize the correct message/command of the user.
- The user is expected to give exact commands, in case of even slight variations, the model might misfunction.
- Faulty microphone can lead to bad audio recording which might hinder the model's efficiency.
- A bad camera quality of the user's phone might spoil the picture quality and make it difficult for the model to identify the text and object, reducing the

accuracy of the model.

- Any language other than English is not supported, hence the user is expected to understand and speak English comfortably. However, accent is flexible.
- Bad Bluetooth connectivity with hardware might hamper the user experience.

# METHODOLOGY ADOPTED

## 3.1. Investigative Techniques

We are creating a new system model. We haven't invented something new; we have used already developed tools and technologies, most of the system we have studied are based on specific feature only some had object detection so some had text detection they were all standalone system with only one feature there were other who had both but they had nonfunctioning feature and accuracy moreover none of the system had voice assistant integrated with them. A complete package with will help blind is still not implemented which will allow more better accuracy and satisfactory responses. Also, the accuracy of the above system was limited Hence due to limited accuracy no system was developed and deployed with complete package from object detection to text detection or be it alarming the blind person with the obstacle nearby them. So, we are a building our system after taking all these shortcomings into consideration. With our system, the object detection from finding specific object to scanning through object around us or text detection of images to scanning through documents and voicing it over to the blind person or be it navigating through paths by detecting obstacles around them. We are developing open-source system which can help the blind person with all the above features mentioned above. We will use all these four modules integrate them to develop a flutter App for the blind person to overcome day to day difficulties at no cost

## 3.2 Proposed Solution

The proposed solution is to develop an interactive voice-based assistant that will interact with the use taking into account the user's current need while walking or finding object or reading text from document. Such a system will provide a more fulfilling experience to the user of things around them. This voice-based assistant will communicate with them in English language. This solution of aiding the blind person in day-to-day life is divided into four parts, the first part being the object detection, second part being the text detection, third part being the blind stick, fourth part integrating all the above 3 part into one app which will be voice over by the voice-based assistant.

Object detection is divided into two part finding specific object user will give input to voice assistant then rotate the camera of the phone it that specific object comes in camera frame the it will be voice overed by the assistant found! found! found! Second part is detecting all object present in the surrounding for this will be using tfflite package and a pre trained SSD mobile net model or yolo depending on the accuracy factor.

Second part i.e., Text detection involving detection text in images and can scroll the document and all the scan text captured is converted to speech to recognize Text in Images we will be using ML kit of google firebase API google-ml-kit with fairly good accuracy

Third part Navigating through path which will be featured by the stick whenever there is an obstacle or something is near the corresponding blind the sticks fires out an alarm of 'beep' 'beep' so that he can be made aware of the obstacles and take action accordingly The Arduino of the stick will be activated from the app only using NodeMCU TCP connection and the stick will be simply an Arduino based beeping stick which beeps when something is near its vicinity.

Last part and the most important voice-based assistant as visually impaired people cannot see so it will be useless app if there is no voice assistant which will be taking input from the user and performing action accordingly for this, we will be using Alan’s voice assistant SDK for Flutter will be integrating all the above 3 models with the Alan's voice assistant

## 3.3 Work Breakdown Structure

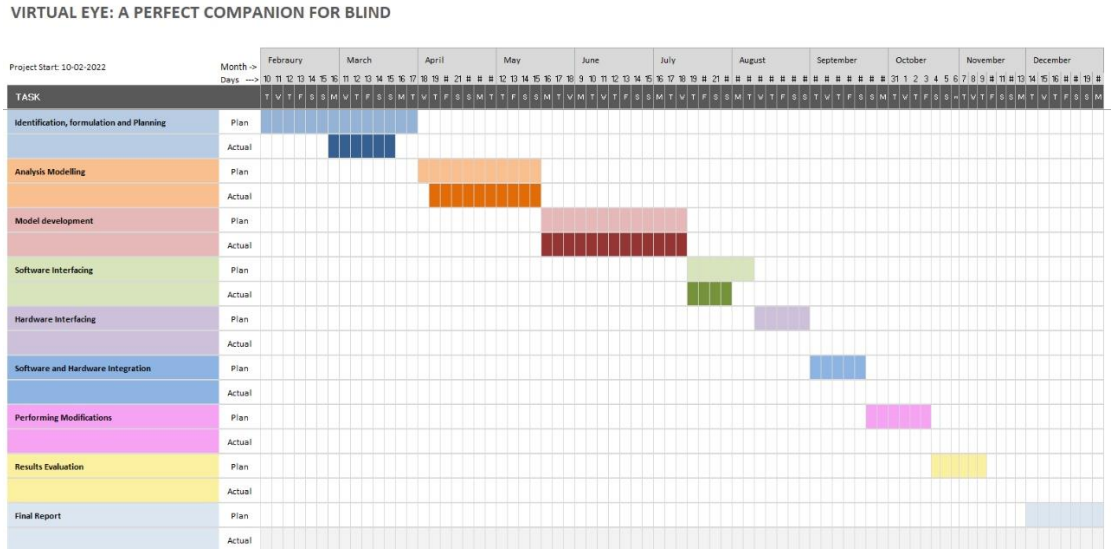



Fig 1: Work Breakdown Structure

## 3.4 Tools and Technology:

These are the following tools used:

Flutter: Application development (Base)

Google Ml kit API: To recognize text in images

tfflite package: To detect object in images and image classification

NodeMCU TCP connection: To send signal from base flutter app for the Arduino activation

Alan SDK kit: Voice-Based Assistant

# DESIGN SPECIFICATIONS

## 4.1 System Architecture

### 4.1.1 Architecture Diagram

Figure 2 is the architecture diagram for the “Virtual Eye” project. Architecture diagram provides an overall view of the physical deployment of system and its evolution roadmap. The figure shows the architecture diagram for “Virtual Eye”.

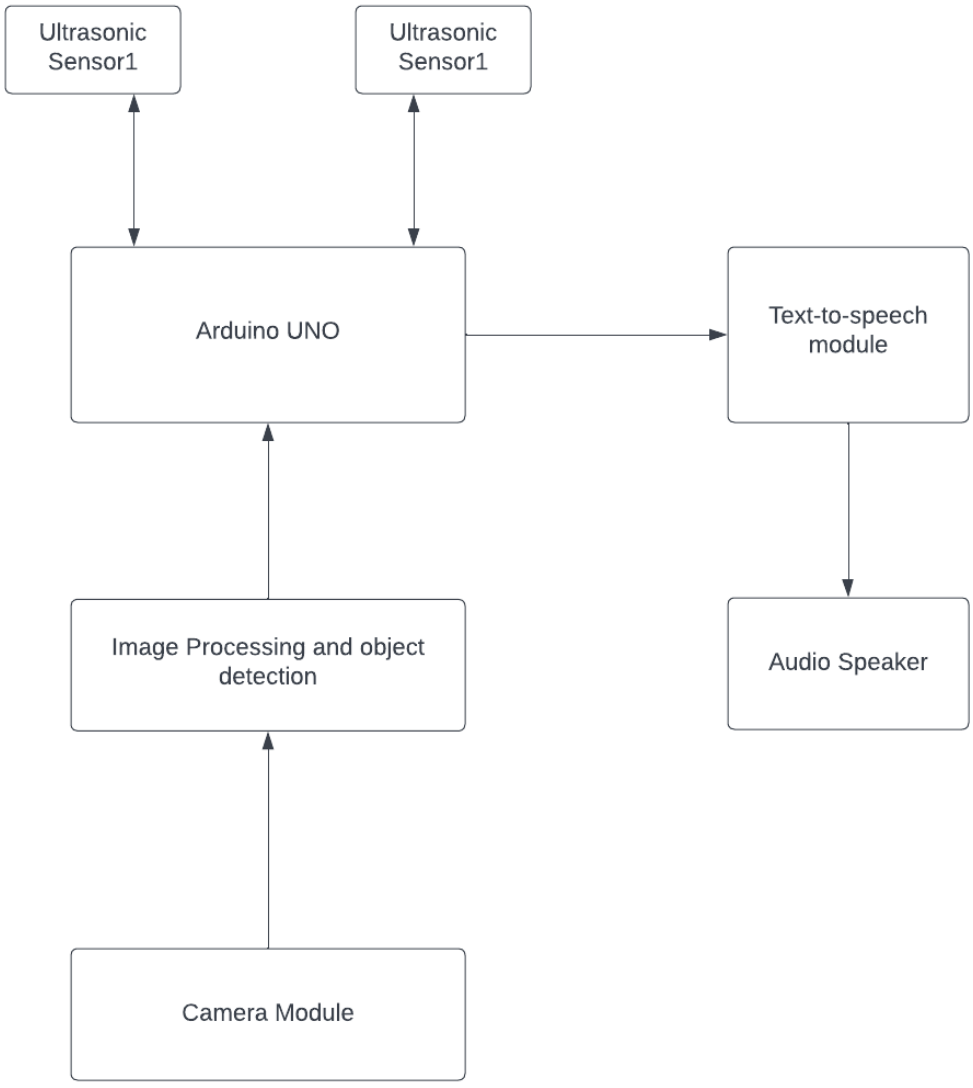



Fig 2: Architecture Diagram

### 4.1.2 Product Perspective:

Figure 3 is the product perspective diagram for the “Virtual Eye” project. It describes overall view of the product, showing all the essential features and components. The figure shows the product perspective diagram for “Virtual Eye”.

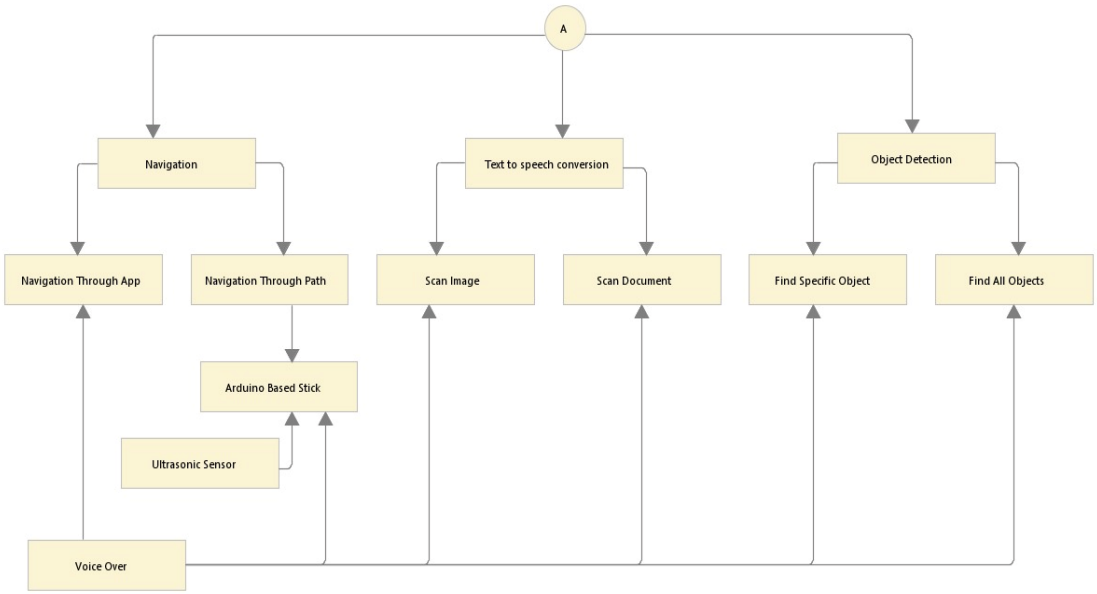



Fig 3: Product perspective Diagram

## 4.2 Design Level Diagrams

### 4.2.1 Activity Diagram

Activity diagram are describing business workflows (business processes), workflows within use-cases, workflows between use-cases and workflows for complex operations. Activity diagram consist of activities, states and transitions between activities and states.

We have created 2 separate activity diagram one for “*USER*” and one for the “*ADMIN*”

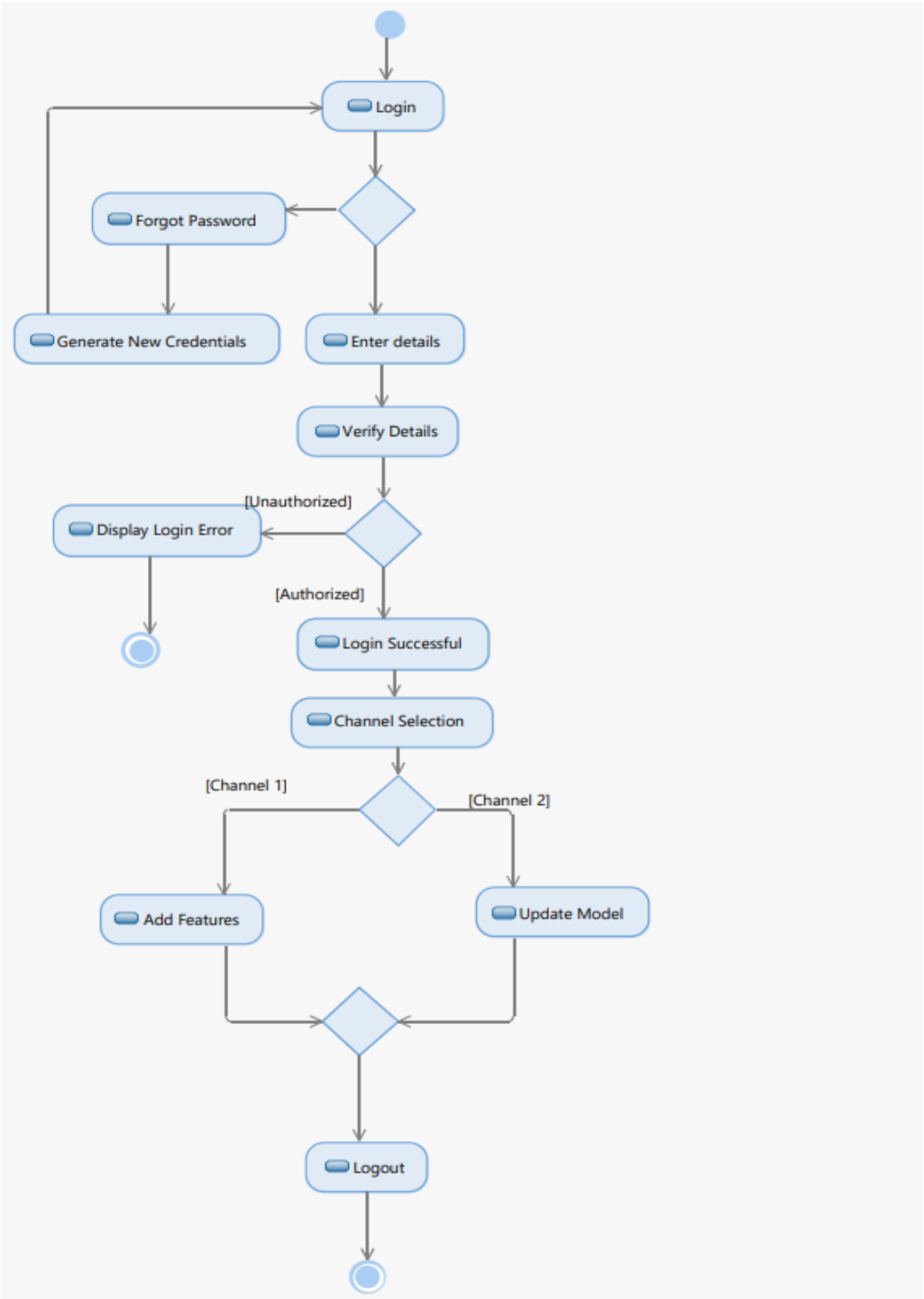



Fig 4: Activity Diagram for Admin

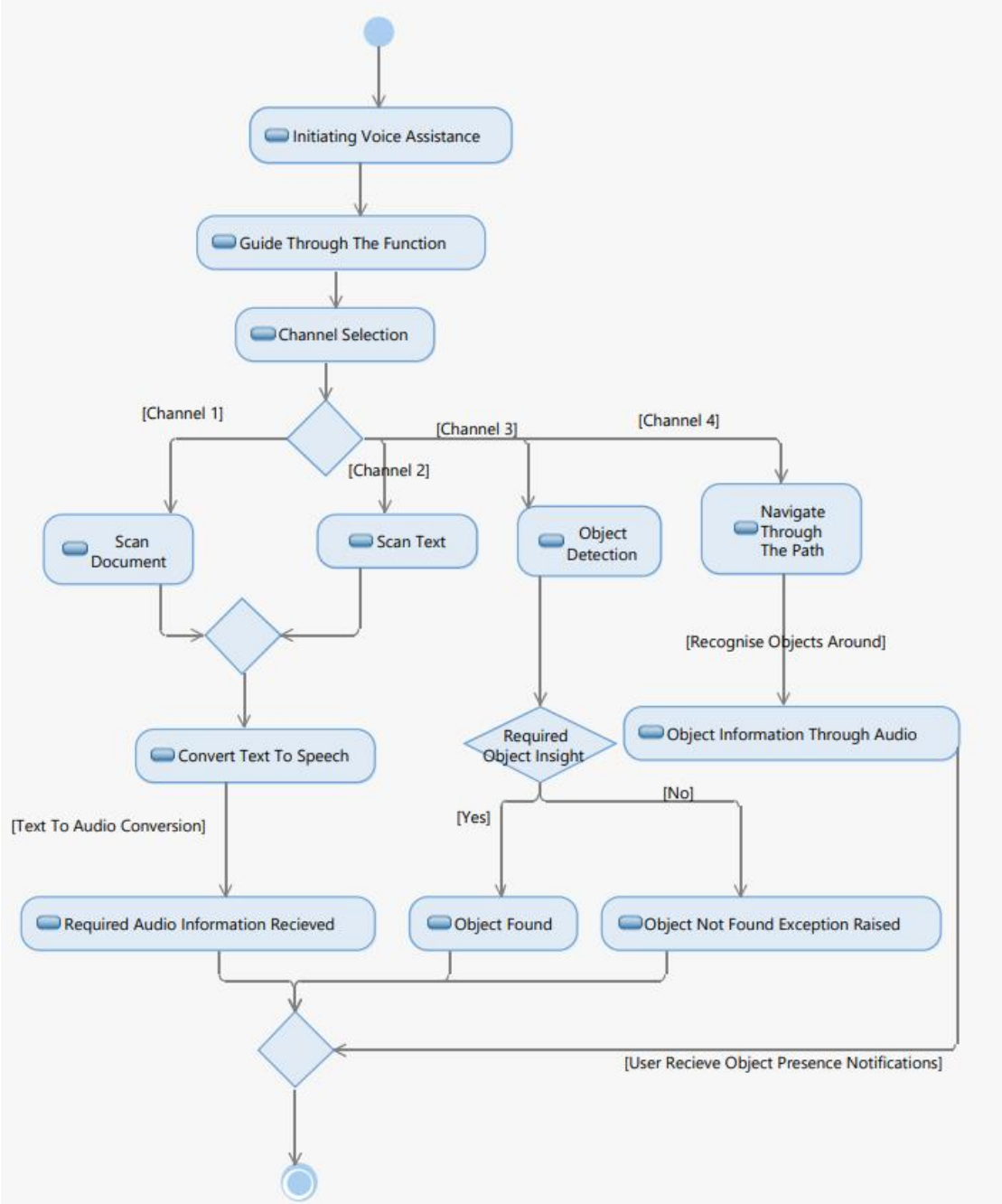



Fig 5: Activity Diagram for User

### 4.2.2 Class Diagram

Figure 6 is the class diagram of "Virtual Eye" project. A Class diagram models is the static structure of a system. It shows relationship between classes, objects, and operations.

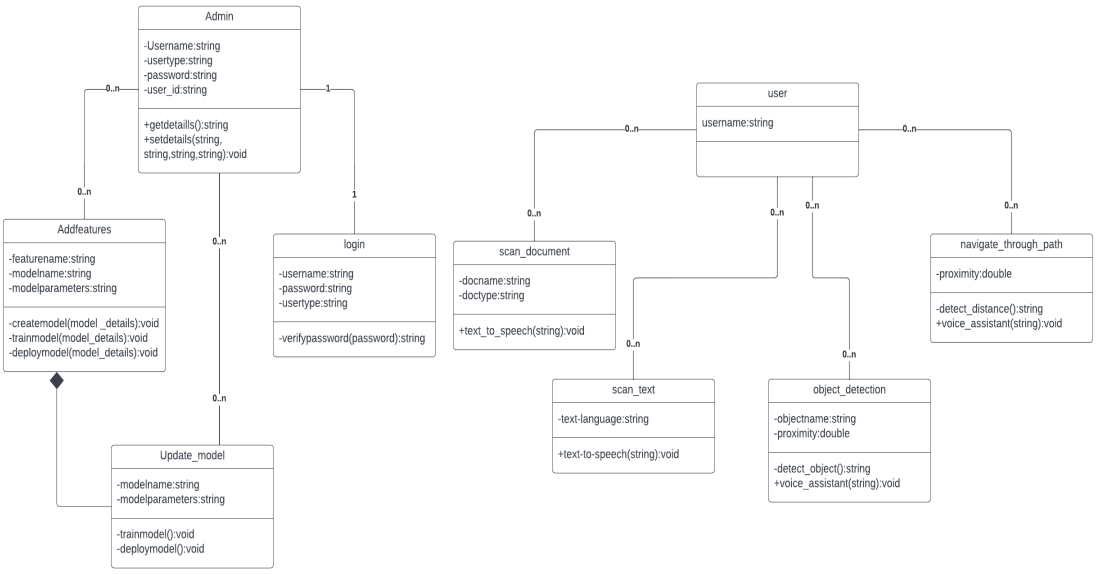



Fig 6: Class Diagram

### 4.2.3 DFD Diagram

Data flow diagram are used to graphically represent the flow of data in a system. We have created three level of data flow diagram.

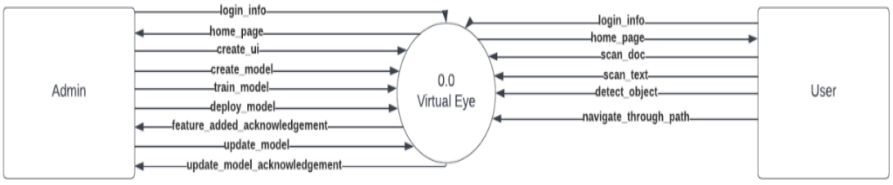



Fig 7: DFD Level 0

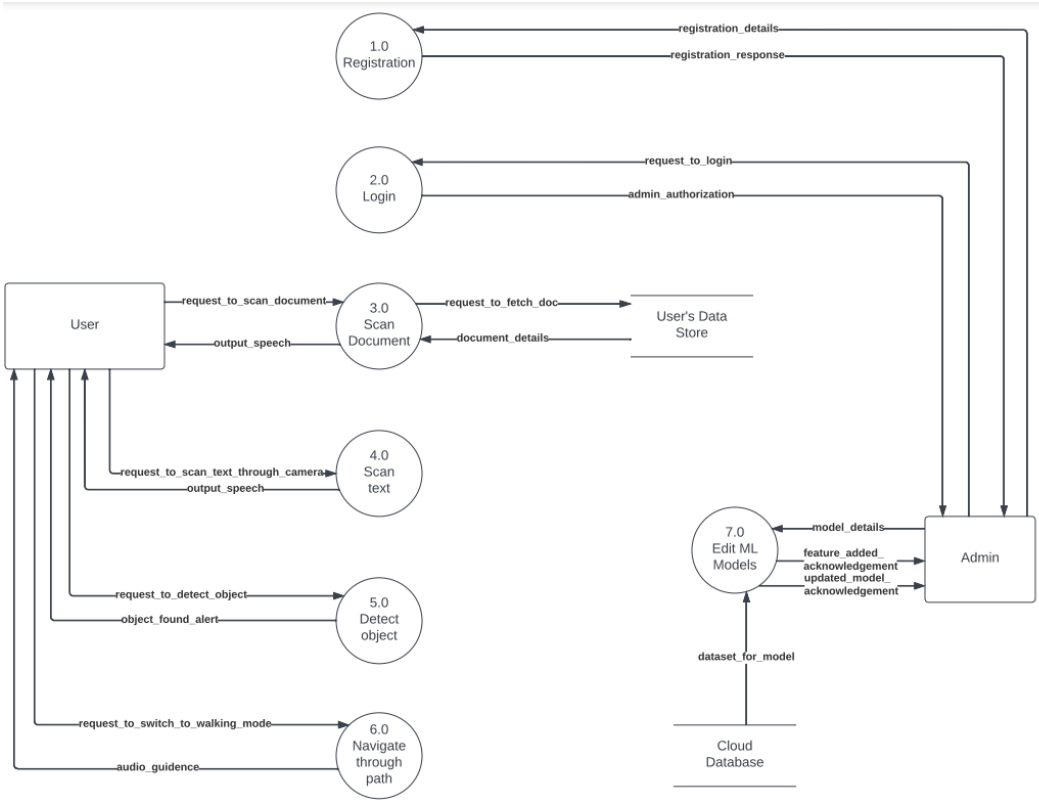



Fig 8: DFD Level 1

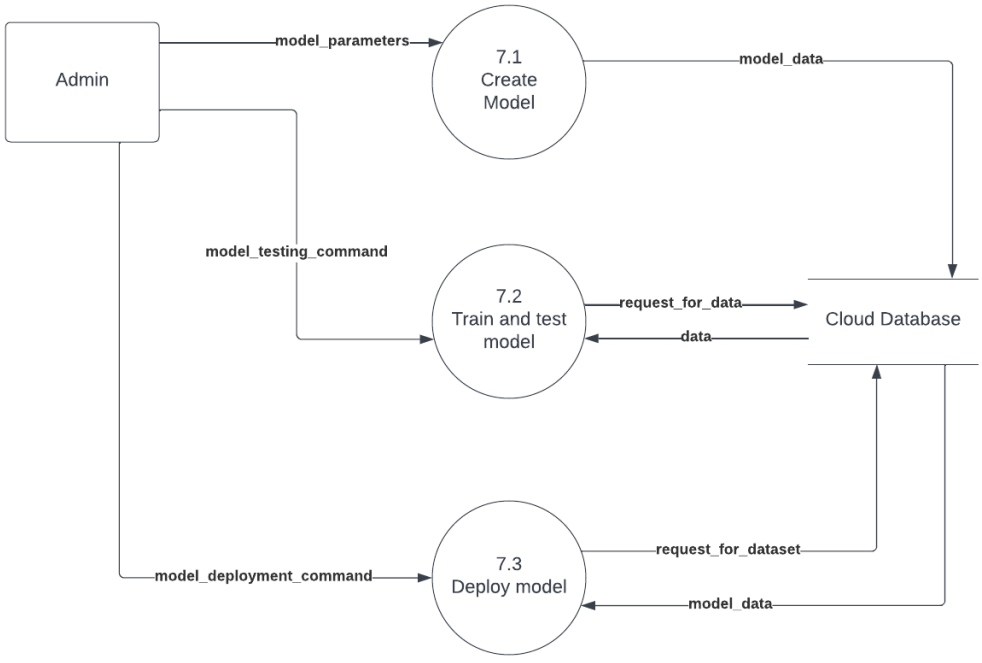



Fig 9: DFD Level 2

### 4.2.4 E-R Diagram

ER-modeling is a data modeling method used in software engineering to produce a conceptual data model of an information system

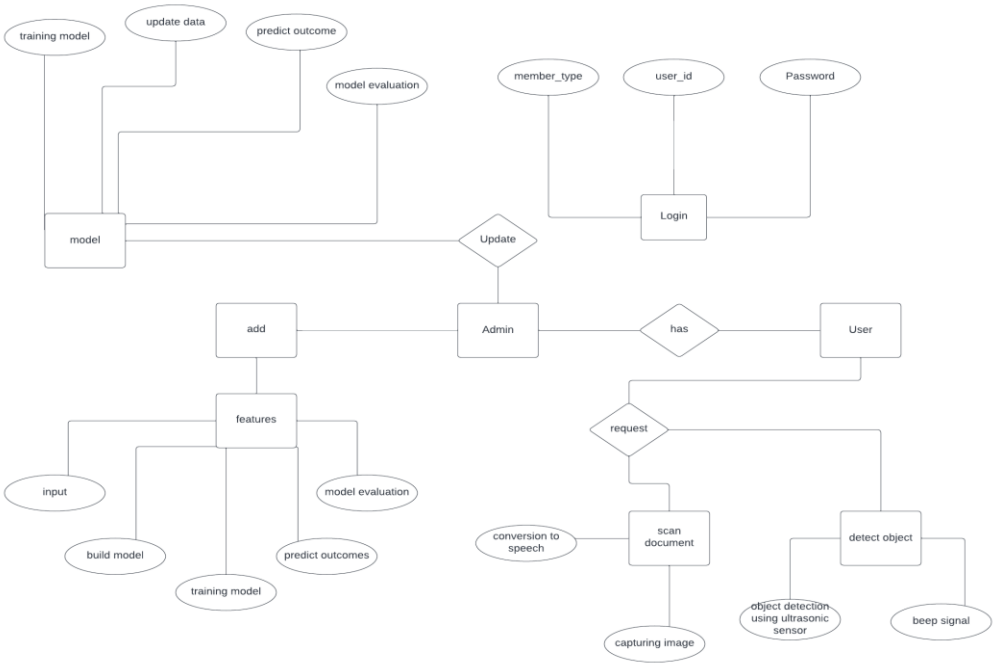



Fig 10: E-R Diagram

### 4.2.5 State Chart Diagram

A state chart diagram describes the behavioral of a single object in response to a series of events. These diagram model the dynamic flow of control from one state to another of a particular object within a system.

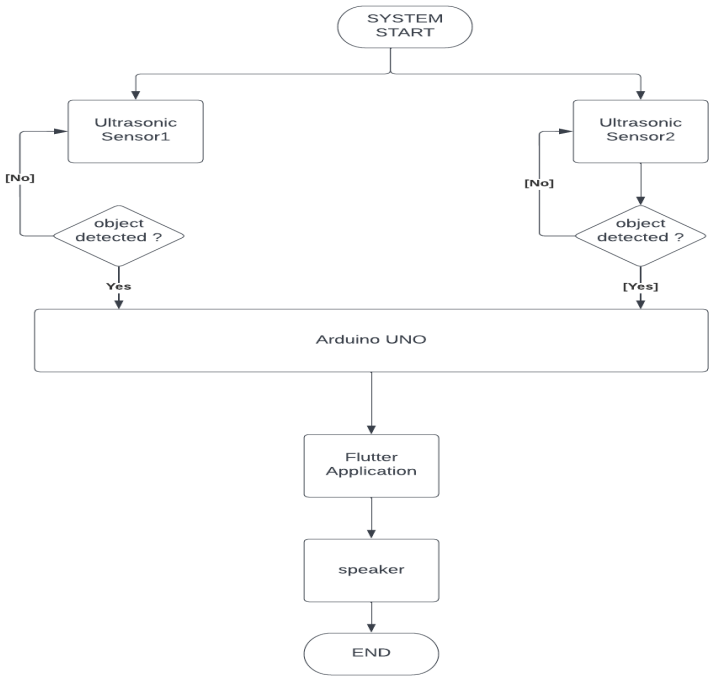


Fig 11: State Chart Diagram

## 4.3 User Interface Diagrams

### 4.3.1 Use case Diagram

Figure 12 is the use case diagram for the “VIRTUAL EYE” project. It helps depict the software system requirements for a system under development. The figure shows the use case diagram for “VIRTUAL EYE”. The primary actors are the user and admin, secondary actors are database.

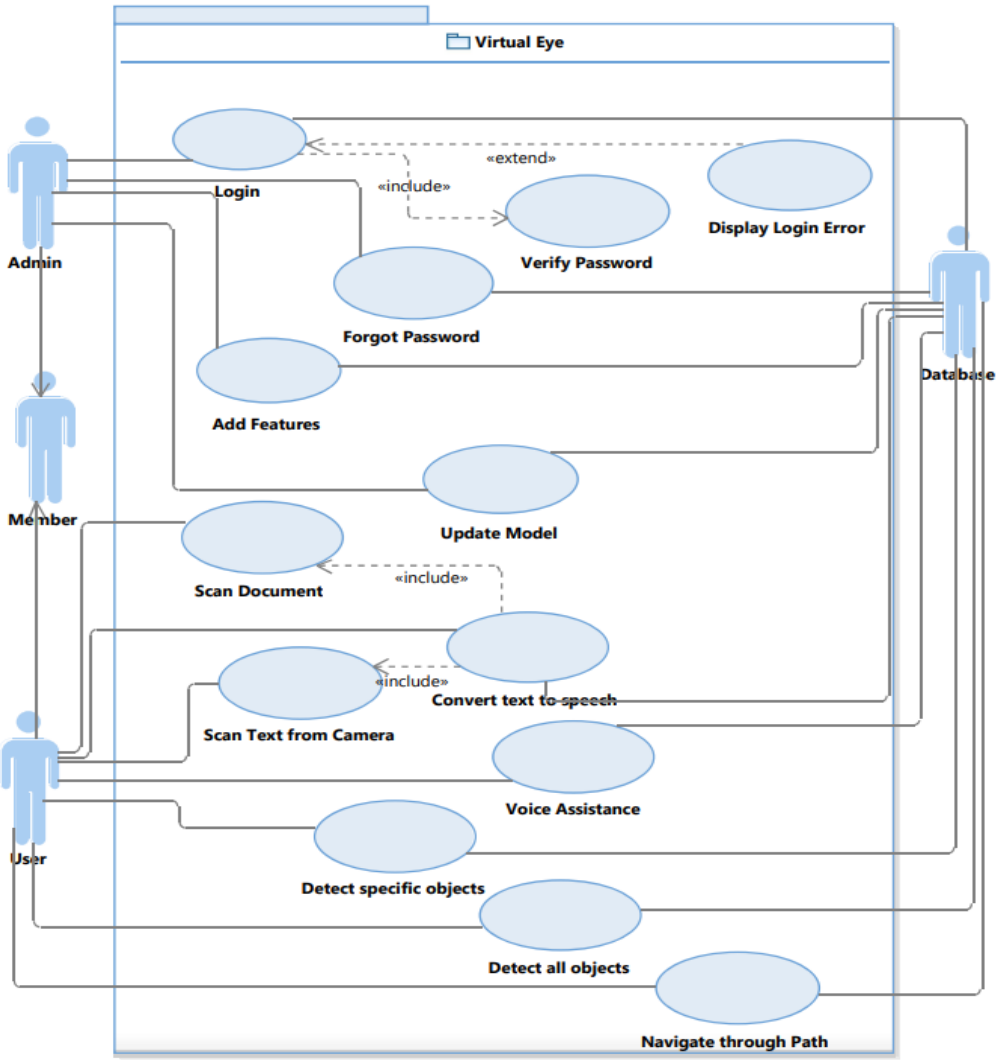



Fig 12: Use Case Diagram

#### 4.3.1.1 Use case templates

**Use Case Scenario 1**

Table 5 shows use case scenario 1. This depicts login mechanism of our application which is available for only admins. The actor has been restricted to Admin considering it will add on to one more task for our users to perform for using our application decreasing ease of using our application. The admin will require to login through his correct credentials recorded in the database to enjoy required admin rights.

Table 5: Use Case Scenario 1

| | |
|---|---|
| 1. Use Case ID | 1 |
| 2. Use Case Title | Login |
| 3. Actors | Admin |
| 4. Purpose: Logging into app as admin | |
| 5. Description:<br>Used for logging into app as admin | |
| 6. Pre-conditions:<br>1 Admin account should exist. | |
| 7. Task Sequence:<br>1. The admin needs to enter the password and click on submit.<br>2. The admin will be granted access to Virtual Eye as a admin. | |
| 8. Alternate flow: The Admin has change of mind and hits the “cancel”. | |
| 9. Exceptional flow: If the admin enters incorrect login credentials, the system will issue an error to prompt the user to re-enter the details. | |
| 10. post-conditions:<br>1. The login credentials will be verified.<br>2. Login error will be displayed if the details don’t match.<br>3. The admin will be given administrative rights for the application “Virtual Eye” | |
| Modification History: 02-May-2022 | |
| Author: Raju Kumar Gupta, Swapnil Kumar Singh, Riya Singhal, Aditi | |

**Use case scenario 2**

Table 6 shows use case scenario 2. This depicts the forgot password use case where if an admin forgets his correct credentials recorded in the database, then he can use this to reset his credentials.

Table 6: Use Case Scenario 2

| | |
|---|---|
| 1. Use Case ID | 2 |
| 2. Use Case Title | Forgot Password |

| | |
|---|---|
| 3. Actors | Admin |
| 4. Purpose: The Admin forgets the password | |
| 5. Description:<br>This feature provides the admin to change their password in case they forget it. | |
| 6. Pre-conditions:<br>1. Login failed | |
| 7. Task Sequence:<br>1. Admin forgets the password.<br>2. Then the admin can click on the 'Forgot Password' button. | |
| 8. Alternate flow:<br>1. The admin is able to retrieve the previous password. | |
| Exceptional flow: The admin enters the initial password. | |
| 9. post-conditions:<br>1. The admin resets the password. | |
| Modification History: 02-May-2022 | |
| Author: Raju Kumar Gupta, Swapnil Kumar Singh, Riya Singhal, Aditi | |

### Use case scenario 3

Table 7 shows use case scenario 3. This depicts adding features use case which comes under amin privilege. The admin can add new features to the application to keep it up to date according to new requirements of the user and keep up with changing technology.

Table 7: Use Case Scenario 3

| | |
|---|---|
| 1. Use Case ID | 3 |
| 2. Use Case Title | Add Features |
| 3. Actors | Admin |
| 4. Purpose: To add new features to the software. | |
| 5. Description:<br>This feature allows the admin to add new features to the software. | |
| 6. Pre-conditions:<br>1. Amin must be logged in with admin privilege. | |
| 7. Task sequence:<br>1.The admin logs in.<br>2.The admin adds new features.<br>3.Changes would be displayed on the screen. | |
| 8. Alternate flow: No alternate flow. | |
| 9. Exceptional flow: No exceptional flow. | |
| 10. post-conditions:<br>1.The admin can check the new features. | |
| Modification History: 02-May-2022 | |
| Author: Raju Kumar Gupta, Swapnil Kumar Singh, Riya Singhal, Aditi | |

**Use case scenario 4**

Table 8 shows use case scenario 4. This depicts update models use case which also comes under admin privilege. The admin can use this to update predictions model use to keep up with new research and technological advances in machine learning models. This will help in ensuring our prediction models can predict as accurately as possible according to currently existing ideologies.

Table 8: Use Case Scenario 4

| | |
|---|---|
| 1. Use Case ID | 4 |
| 2. Use Case Title | Update Models |
| 3. Actors | Admin |
| 4. Purpose: To update the ML and DL models. | |
| 5. Description:<br>This feature allows the admin to update the ML and DL models used for predictions. | |
| 6. Pre-conditions:<br>1. Amin must be logged in with admin privilege. | |
| 7. Task sequence:<br>1.The admin logs in.<br>2.The admin updates the existing models.<br>3.The admin logs out of his account. | |
| 8. Alternate flow: No alternate flow. | |
| 9. Exceptional flow: No exceptional flow. | |
| 10. post-conditions: No post-conditions. | |
| Modification History: 02-May-2022 | |
| Author: Raju Kumar Gupta, Swapnil Kumar Singh, Riya Singhal, Aditi | |

**Use case scenario 5**

Table 9 shows use case scenario 5. This depicts the first feature which is provided to the user. The scanning document feature makes it easier for the user to realize the contents of the provided document through audio when combined with text to speech conversion mechanism.

Table 9: Use Case Scenario 5

| | |
|---|---|
| 1. Use Case ID | 5 |
| 2. Use Case Title | Scan Document |
| 3. Actors | User |
| 4. Purpose: To scan documents from device's local storage. | |
| 5. Description:<br>This feature allows the user to scan documents from device's local storage for reading text from those documents. | |

| |
|---|
| 6. Pre-conditions:<br>1.Document must be available on device's local storage.<br>2.Document must only contain English characters. |
| 7. Task sequence:<br>1.The user clicks on "Scan Document".<br>2.The user selects file from his device to upload.<br>3. The document would be uploaded and scanned by the software. |
| 8. Alternate flow: The user changes his mind and leaves the page. |
| 9. Exceptional flow: No exceptional flow. |
| 10. post-conditions:<br>1.The document can now undergo text to speech conversion and the user would be able to listen the content of the document. |
| Modification History: 02-May-2022 |
| Author: Raju Kumar Gupta, Swapnil Kumar Singh, Riya Singhal, Aditi |

### Use case scenario 6

Table 10 shows use case scenario 6. This depicts the second feature which is provided to the user. The scanning text feature makes it easier for the user to realize the contents of the provided text through audio when combined with text to speech conversion mechanism. This will make it easier for the user to realize contents of print media like flyers, advertisements, posters and other text present in the environment like text inscription on monuments.

Table 10: Use Case Scenario 6

| | |
|---|---|
| 1. Use Case ID | 6 |
| 2. Use Case Title | Scan Text from Camera |
| 3. Actors | User |
| 4. Purpose: To scan text from camera. | |
| 5. Description:<br>This feature allows the user to scan text from the user's environment using the device's camera. | |
| 6. Pre-conditions:<br>1.The device should have a working camera.<br>2.The software should be allowed to access the camera. | |
| 7. Task sequence:<br>1.The user clicks on "Scan Text from Camera".<br>2.The user clicks an image which contains the text to be read.<br>3. The image would be uploaded and scanned by the software for text recognition. | |
| 8. Alternate flow: The user changes his mind and leaves the page. | |
| 9. Exceptional flow: No exceptional flow. | |
| 10. post-conditions:<br>1.The image can now undergo text to speech conversion and the user would be able to listen to the readable text of the image. | |
| Modification History: 02-May-2022 | |
| Author: Raju Kumar Gupta, Swapnil Kumar Singh, Riya Singhal, Aditi | |

**Use case scenario 7**

Table 11 shows use case scenario 7. This text to speech conversion feature will make it possible for the user to realize the contents of the scanned document and scanned text through audio commands.

Table 11: Use Case Scenario 7

| | |
|---|---|
| 1. Use Case ID | 7 |
| 2. Use Case Title | Text to speech Conversion |
| 3. Actors | User |
| 4. Purpose: To convert text to speech. | |
| 5. Description:<br>This feature allows converts the text written in the document or on the image to speech so that the user can listen to it. | |
| 6. Pre-conditions:<br>1.The document or image should be scanned and available for processing.<br>2.The document or image should contain readable English characters.<br>3.The text should not be handwritten. | |
| 7. Task sequence:<br>1.The user clicks on “Convert text to speech”.<br>2.The software performs natural language processing and then speech generation<br>3. The user would be able to listen to the text written in the document or image. | |
| 8. Alternate flow: No alternate flow. | |
| 9. Exceptional flow: If there is no readable text present then the software would inform that “no text is found”. | |
| 10. post-conditions:<br>1.The user understood the text written in document and image. | |
| Modification History: 02-May-2022 | |
| Author: Raju Kumar Gupta, Swapnil Kumar Singh, Riya Singhal, Aditi | |

**Use case scenario 8**

Table 12 shows use case scenario 8. This depicts the third feature which is provided to the user. The voice assistant feature will provide tremendous ease of use to our user so that they can access and control our application using audio commands.

Table 12: Use Case Scenario 8

| | |
|---|---|
| 1. Use Case ID | 8 |
| 2. Use Case Title | Voice Assistant |
| 3.Actors | User |
| 4. Purpose: To guide the user to different modes of the software through speech. | |
| 5. Description:<br>This feature guides the blind user through different modes of the software. | |
| 6. Pre-conditions:<br>1.Device's speakers should be in working conditions and the software should be given access to it. | |
| 7. Task sequence:<br>1.The user clicks on "Convert text to speech". | |
| 8. Alternate flow: No alternate flow. | |
| 9. Exceptional flow: No exceptional flow. | |
| 10. post-conditions:<br>1.The user can easily work with the software. | |
| Modification History: 02-May-2022 | |
| Author: Raju Kumar Gupta, Swapnil Kumar Singh, Riya Singhal, Aditi | |

## Use case scenario 9

Table 13 shows use case scenario 9. This depicts the fourth feature which is provided to the user. This feature will help the user to identify specific objects in his surrounding according to his requirement.

Table 13: Use Case Scenario 9

| | |
|---|---|
| 1. Use Case ID | 9 |
| 2. Use Case Title | Detect specific objects |
| 3.Actors | User |
| 4. Purpose: To detect a specific object in the user's surroundings. | |
| 5. Description:<br>This feature allows the user to detect a specific object in his/her surroundings. | |
| 6. Pre-conditions:<br>1.The device should have a working camera.<br>2.The software should be allowed to access the camera.<br>3.The camera should be shown the environment with slow speed. | |
| 7. Task sequence:<br>1.The user clicks on "Detect specific object".<br>2.The user selects which object is to be found in the environment from the drop-down menu.<br>2.The user turns the camera towards the surroundings to scan the objects nearby.<br>3.As soon as the software finds the object, it indicates by saying "Found! Found! Found!". | |
| 8. Alternate flow: No alternate flow. | |
| 9. Exceptional flow: If the camera is rotated with high speed, then the software won't be able to detect the object. | |
| 10. post-conditions:<br>1.The user will be able to find the required object. | |
| Modification History: 02-May-2022 | |
| Author: Raju Kumar Gupta, Swapnil Kumar Singh, Riya Singhal, Aditi | |

**Use case scenario 10**

Table 14 shows use case scenario 10. This depicts the fifth feature provided to the user. This will enable the user to recognize its environment by acting as a virtual eye for the user.

Table 14: Use Case Scenario 10

| | |
|---|---|
| 1. Use Case ID | 10 |
| 2. Use Case Title | Detect all objects |
| 3.Actors | User |
| 4. Purpose: To detect all objects in the user's surroundings. | |
| 5. Description:<br>This feature allows the user to detect all objects in his/her surroundings. | |
| 6. Pre-conditions:<br>1.The device should have a working camera.<br>2.The software should be allowed to access the camera.<br>3.The camera should be shown the environment with slow speed. | |
| 7. Task sequence:<br>1.The user clicks on "Detect all objects".<br>2.The user turns the camera towards the surroundings to scan the objects nearby.<br>3.As soon as the software finds the object, it indicates by saying the object name which is detected. | |
| 8. Alternate flow: No alternate flow. | |

| |
|---|
| 9. Exceptional flow: If the camera is rotated with high speed, then the software won't be able to detect the object. |
| 10. post-conditions:<br>1.The user will be able to find all objects of the surroundings. |
| Modification History: 02-May-2022 |
| Author: Raju Kumar Gupta, Swapnil Kumar Singh, Riya Singhal, Aditi |

## User case scenario 11

Table 15 shows use case scenario 11. This depicts the sixth feature provided to the user. This will enable the user to navigate through the paths easily by recognizing their environment in a better way.

Table 15: Use Case Scenario 11

| | |
|---|---|
| 1. Use Case ID | 11 |
| 2. Use Case Title | Navigate through paths |
| 3.Actors | User |
| 4. Purpose: To help the user navigate through his/her path. | |
| 5. Description:<br>This feature allows the user to navigate through his/her path. | |
| 6. Pre-conditions:<br>1.The person should be equipped with the hand stick.<br>2.The sensors present on the stick should be in working conditions. | |
| 7. Task sequence:<br>1.The user clicks on "Navigate through paths" which activates the hand stick.<br>2.The user now rotates his/her stick to search for obstacles.<br>3.As soon as the software detects some obstacle, it would send a beep signal. | |
| 8. Alternate flow: No alternate flow. | |
| 9. Exceptional flow: If the stick is rotated with high speed, then the software won't be able to detect the obstacle. | |
| 10. post-conditions:<br>1.The user will be able to move through his/her path safely. | |
| Modification History: 02-May-2022 | |
| Author: Raju Kumar Gupta, Swapnil Kumar Singh, Riya Singhal, Aditi | |

### 4.3.2 Sequence Diagram

These diagrams describe the interaction between objects in a sequential order i.e., the order in which these interactions take place.

We have created 2 separate sequence diagram one for "*USER*" and one for the "*ADMIN*"

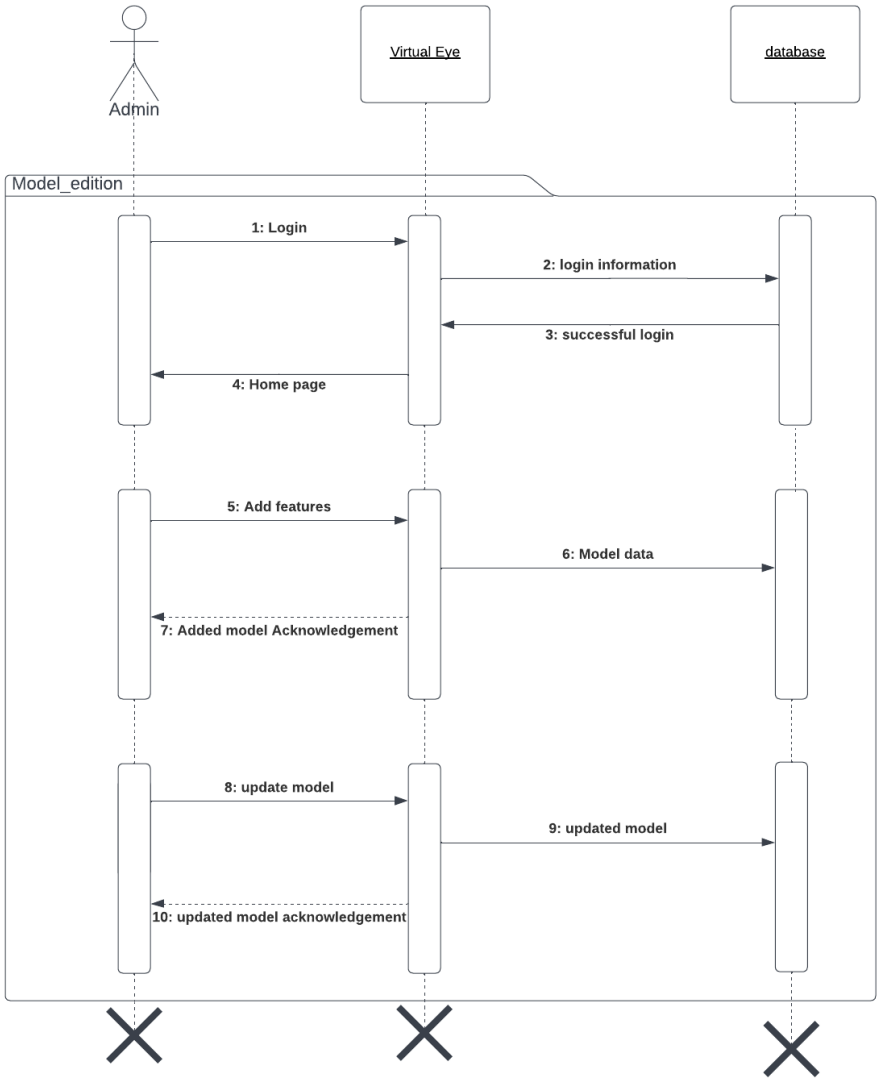



Fig 13: Sequence Diagram for Admin

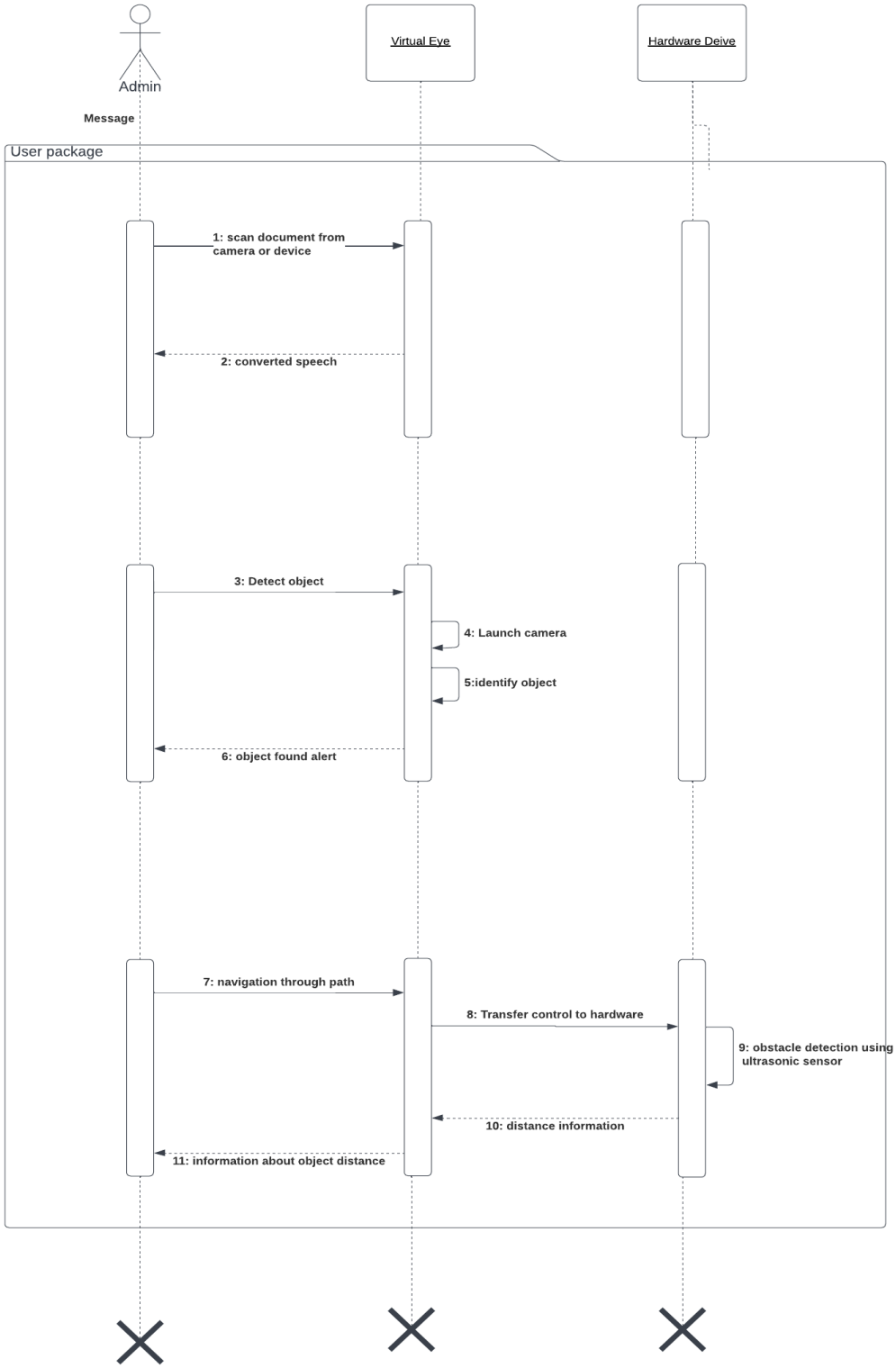



Fig 14: Sequence Diagram for User

## 4.4 Snapshot of Working Prototype

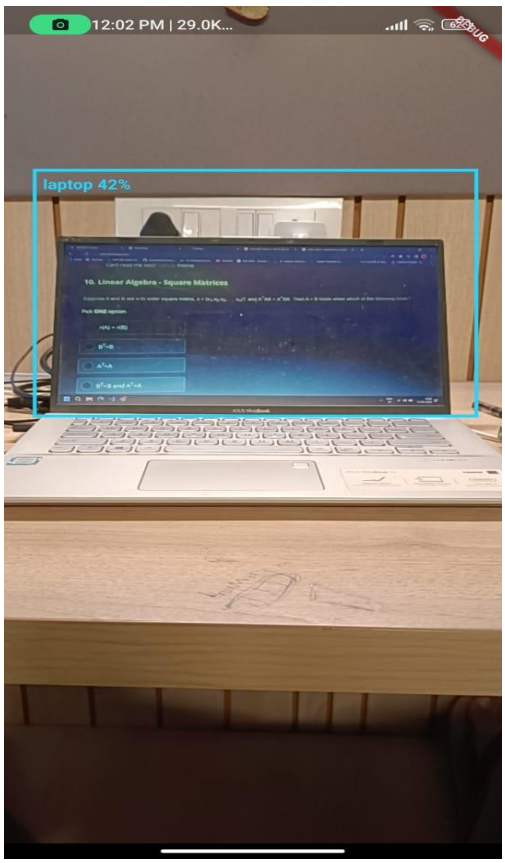



Fig 15: Object detection of specific object

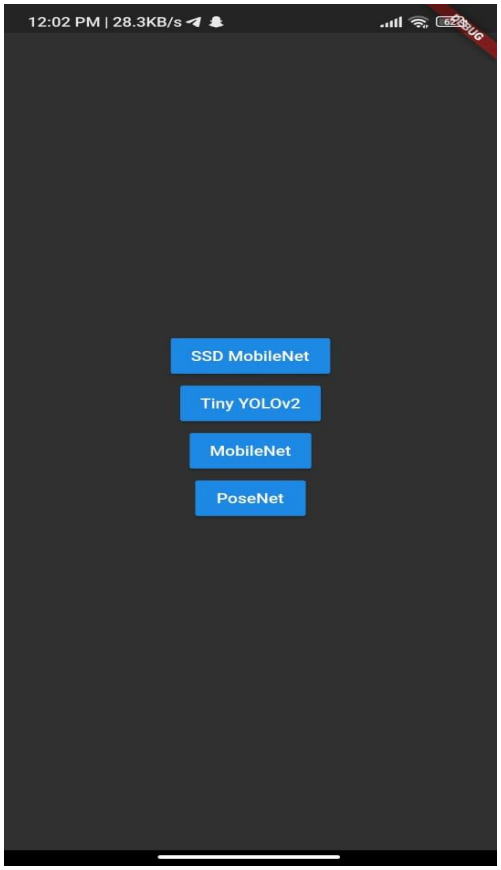
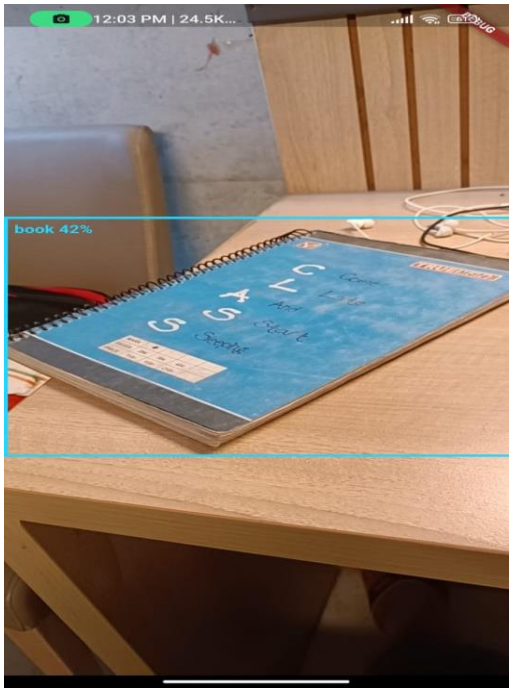



Fig 16: Used and compared the accuracy Of all the above model



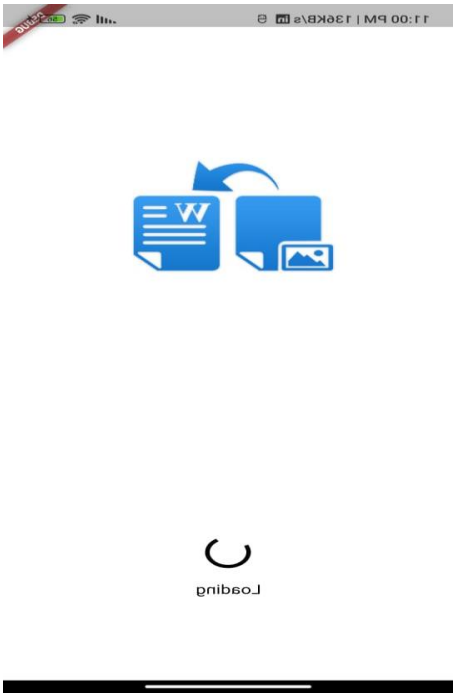



Fig 17: Object detection

Fig 18: Loder page for text detection

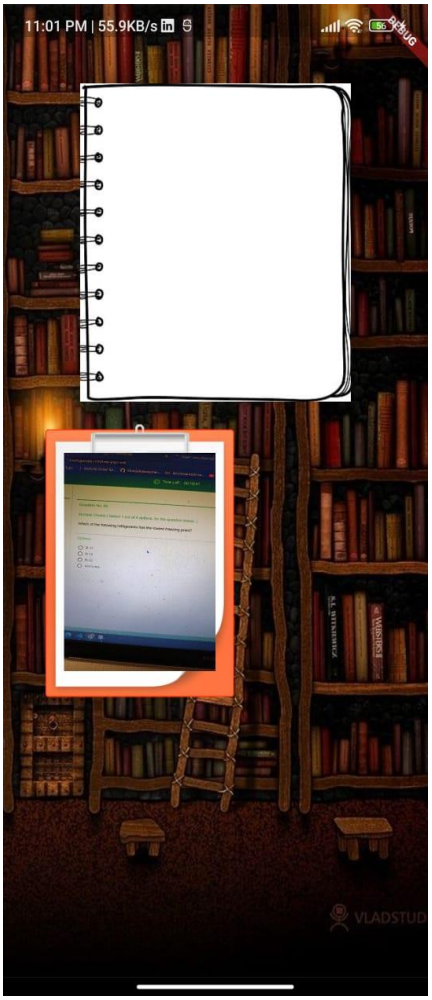




Fig 19: Scan Images

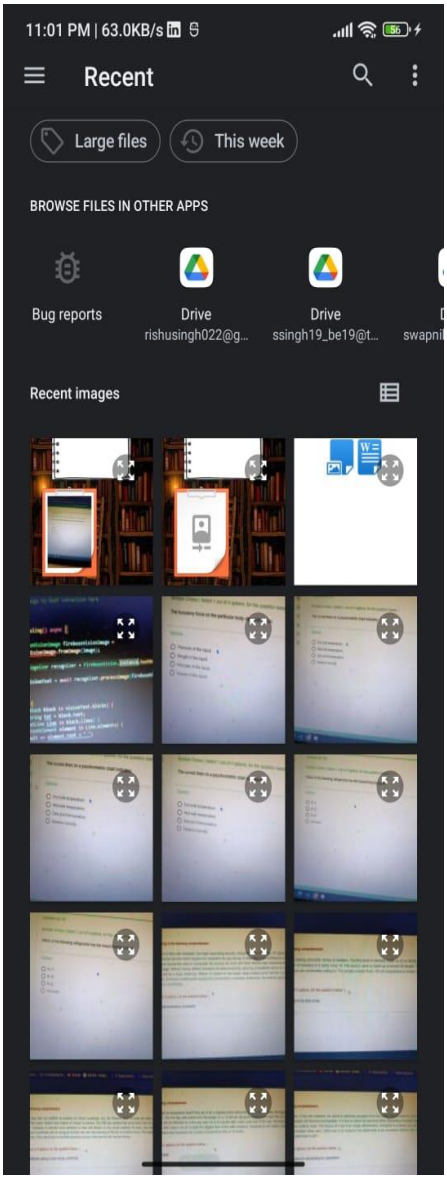



Fig 20: Upload images from gallery

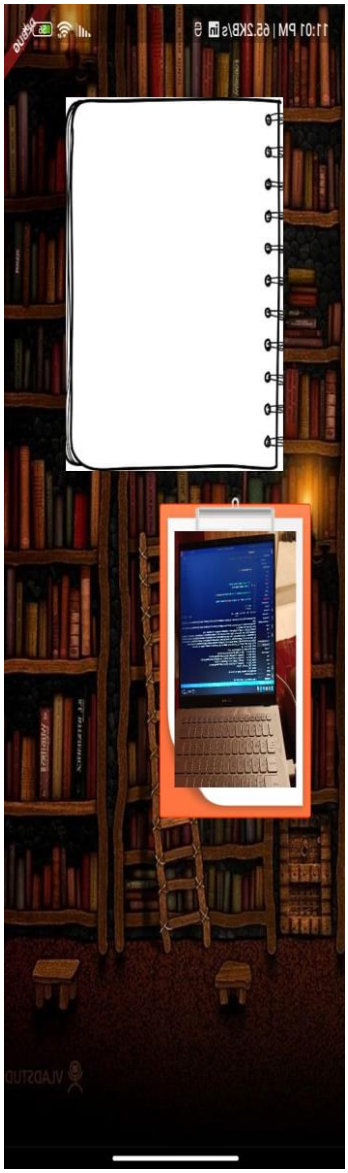


Fig 21: Scan report

# IMPLEMENTATION AND EXPERIMENTAL RESULTS

## 5.1 Experimental Setup

For our project virtual eye, we have majorly divided it into 2 parts. first the software and second is the hardware part.

Now for the software part, it is basically a flutter application, which requires its APK to be built on smartphones with USB connection with the laptop and developer mode ON for continuous build and test of the APK.

The integration of the hardware part and base flutter app requires NodeMCU Tcp connection for that we would require.

## 5.2 Working of the project: In our flutter-based app there is 3 modules.

1. **Object Detection:**
   i. Find specific object: if user want to find any specific object so he has to give the input to the app and start rotating the phone if object is present in the surrounding then app will simply speak out FOUND! FOUND!
   ii. identify all object being captured: when we will use this feature, after opening the app we have to Start rotating our phone if object is present in pre-trained SSD mobile net model and also present in surrounding then it will speak out name of that particular object.
2. **Text to speech Conversion:**
   i. scan text from the surroundings: We are capturing Real time images and all the words present in the images are Scanned and voice overed the app.
   ii. Scanning through Documents: Document present in our local storage and all the words present in the Document are scanned and voice overed the app page by page.
3. **Navigation through path**:
   In this feature we are providing an Arduino based stick to the blind person which is activated through flutter-based app using NodeMCU TCP connection. The H/W module is fixed on the upper part of the

stick near the handle so as to get full outer coverage of the environment. After processing, the distance of the closest obstacle is detected and measured. If it's proximity is much closer to the user, the distance is then converted and sent to the app.

### 5.2.1 Procedural Work Flow:

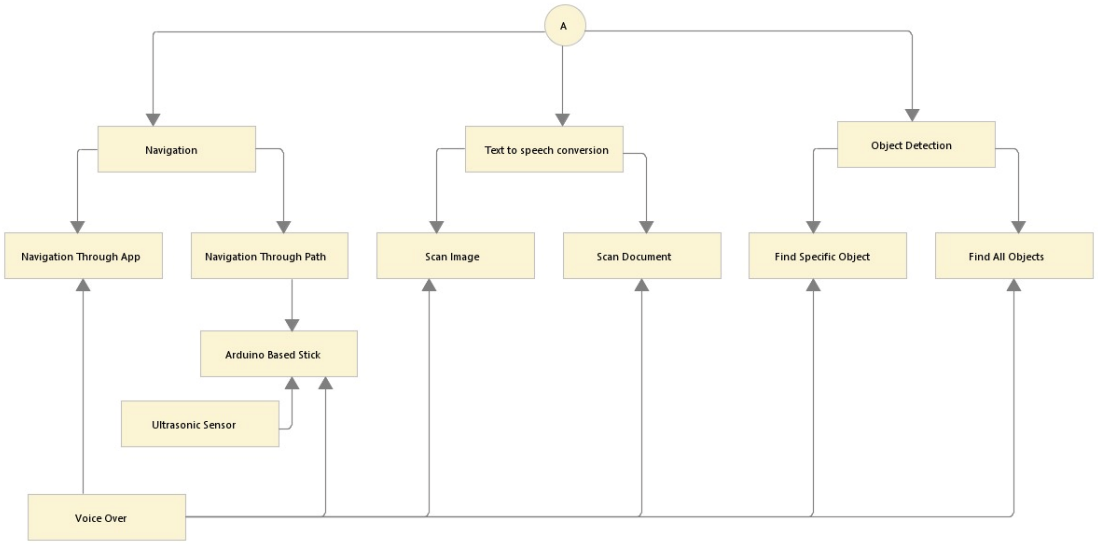
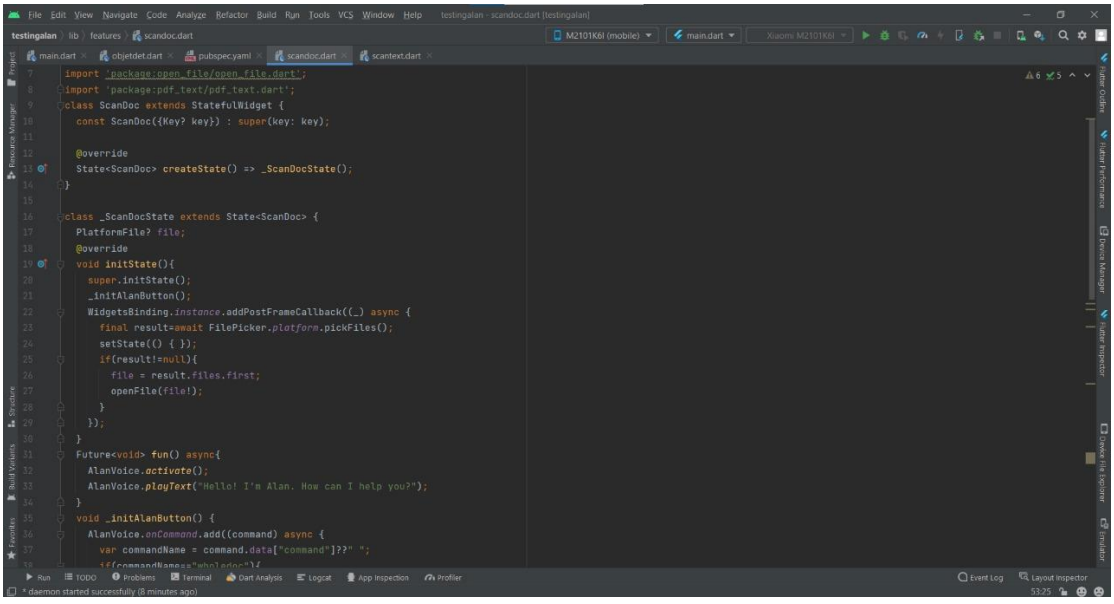



Fig 22: Work Flow Diagram

### 5.2.2 Algorithm Approaches Used:



```
import 'package:open_file/open_file.dart';
import 'package:pdf_text/pdf_text.dart';
class ScanDoc extends StatefulWidget {
  const ScanDoc({Key? key}) : super(key: key);

  @override
  State<ScanDoc> createState() => _ScanDocState();
}

class _ScanDocState extends State<ScanDoc> {
  PlatformFile? file;
  @override
  void initState(){
    super.initState();
    _initAlanButton();
    WidgetsBinding.instance.addPostFrameCallback((_) async {
      final result=await FilePicker.platform.pickFiles();
      setState(() { });
      if(result!=null){
        file = result.files.first;
        openFile(file!);
      }
    });
  }
  Future<void> fun() async{
    AlanVoice.activate();
    AlanVoice.playText("Hello! I'm Alan. How can I help you?");
  }
  void _initAlanButton() {
    AlanVoice.onCommand.add((command) async {
      var commandName = command.data["command"]?? '';
```

Fig 23: Code

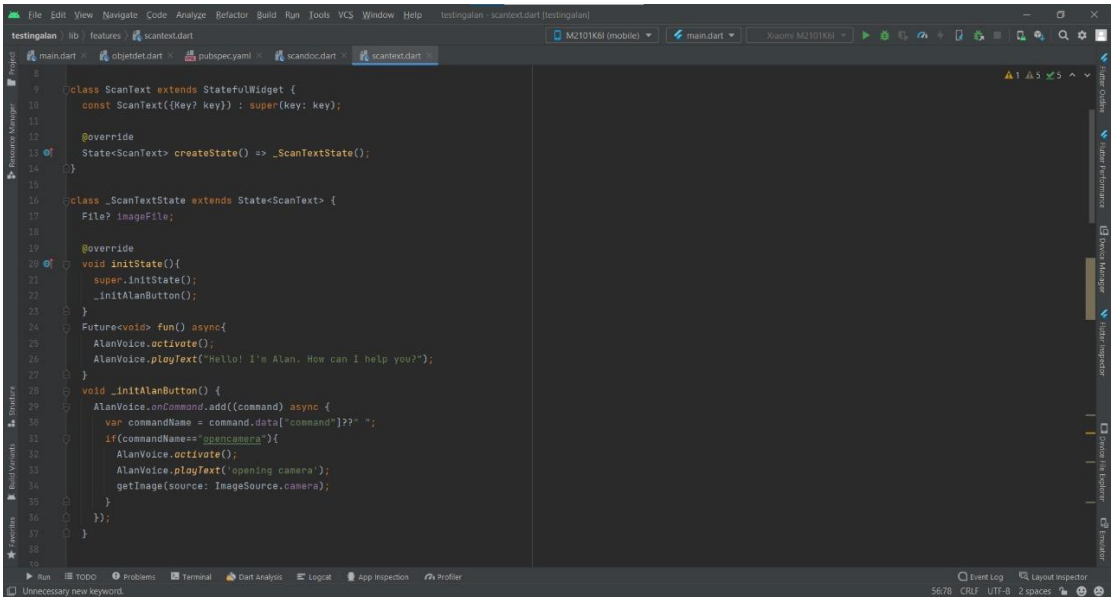



Fig 24: Code

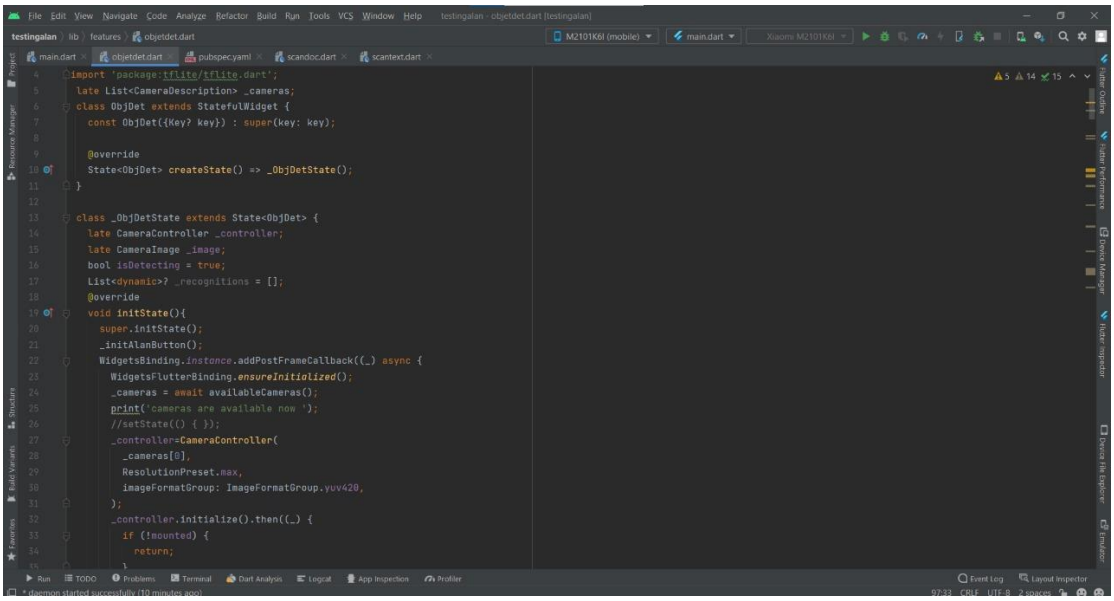



Fig 25: Code

## 5.3Testing Process:

A testing process based on the following parameters:

### 5.3.1 Test Plan

A test plan consists of the scope, approach, resources and schedule of intended test activities.

**Scope:** The scope of testing is to provide accuracy of the model while Detecting the objects from the surroundings as well as detecting text from the images and documents to speech conversion using voice assistant. Testing is important as it tells us the accuracy as well as the further need to improve the model by adding data and labels.

**Approaches & Resources:** The approach for testing the hardware part as well as the software part is done by undertaking various scenarios that the user can face. Keeping these scenarios and implementing a proper solution to the problem faced basically tells us the importance of testing

### 5.3.2 Features to be tested

The features tested were:

1. Detecting all objects from the surroundings.
2. To find specific objects from the surroundings.
3. Text to speech conversion from the surroundings images.
4. Text to speech conversion from the documents.
5. Obstacle detection in path
6. Voice assistant

### 5.3.3 Test Strategy

After building the app from scratch, the user was able to hear out all the objects in his\her surrounding as well as if he\she wants to find any specific objects in the surroundings he\she can do that as well, the user was also able to hear out the text present in the images or in the document, the user could also detect the obstacle present in his\her path.

### 5.3.4 Test Cases and Results:

We tested the app for various test cases, some of them are as shown below with parameters being the case description, expected and actual output:

Table 16: Test Case 1

| | |
|---|---|
| Test Case name | Detect all objects |
| Description | This feature allows the user to detect all objects in his/her surroundings. |
| Expected Output | Successful |

| | |
|---|---|
| Actual Output | Successful |

Table 17: Test Case 2

| | |
|---|---|
| Test Case name | Detect specific objects |
| Description | This feature allows the user to detect a specific object present in his/her surroundings |
| Expected Output | Successful |
| Actual Output | Successful |

Table 18: Test Case 3

| | |
|---|---|
| Test Case name | Voice assistant |
| Description | This feature guides the blind user through different modes of the software |
| Expected Output | Successful |
| Actual Output | Successful |

Table 19: Test Case 4

| | |
|---|---|
| Test Case name | Text to speech conversion |
| Description | This feature allows converts the text written in the document or on the image to speech so that the user can listen to it. |
| Expected Output | Successful |
| Actual Output | Successful |

Table 20: Test Case 5

| | |
|---|---|
| Test Case name | Scan text from image |
| Description | This feature allows the user to scan text from the user's environment using the device's camera |
| Expected Output | Successful |
| Actual Output | Successful |

Table 21: Test Case 6

| Test Case name | Scan document |
|---|---|
| Description | This feature allows the user to scan documents from device's local storage for reading text from those documents. |
| Expected Output | Successful |
| Actual Output | Successful |

Table 22: Test Case 7

| Test Case name | Navigate through path |
|---|---|
| Description | This feature allows the user to navigate through his/her path. speech so that the user can listen to it. |
| Expected Output | Successful |
| Actual Output | Successful |

## 5.4 Validation of Objectives

Table 23: Validation of objective

| S. No. | Objectives | Status |
|---|---|---|
| 1 | Detect all objects | Successful |
| 2 | Detect specific objects | Successful |
| 3 | scan text from the surroundings | Successful |
| 4 | scan text of the documents | Successful |
| 5 | To integrate all features in an application using voice assistant. | Successful |
| 6 | Navigation through path | Successful |

# CONCLUSION AND FUTURE SCOPE

## 6.1 Work Accomplished

We have developed and integrated basic features required by visually impaired people to go through most of their day-to-day necessities. These features are scan text where the user can capture any image in its surrounding and convert all the text in that image to speech, scan document where user can scroll the document and all the text captured is converted to speech, object detection where user can identify specific object as well as objects present in the surrounding in general, navigating through path where the user will be guided about obstacles near him to help him move around freely even in an unfamiliar environment.

We have tried and tested various models for object detection like yolo, mobile-net SSD to see which provided better results and which model to incorporate in our application.

We have tried and incorporated text recognition API's like google-ml-kit for text recognition in captured images which gives a fairly nice accuracy.

We have integrated all our features using Alan AI voice assistant in flutter to increase the usability and provide the user with a more user-friendly and convenient application. The speech commands are have good clarity and speed. Also, it is flexible regarding the accent, gender, age of user and has a good accuracy in recognizing voice commands. This will greatly reduce the user's dependence on visual elements of user interface of the application improving the user experience and ease of use. This proves Alan AI to be a fairly good TTS technology.

## 6.2 Conclusion

A large body of work is presented in this report. There were some challenges initially in the integration and accuracy part but then our team worked together to develop the application

We have tried our best that the visually impaired people feel comfortable while using this App by providing them with basic features of Object detection, Text

detection, and navigation through paths. These features mostly cover the necessities required in daily lifestyle

During the research, we found out that the life of a visually impaired person is full of struggles including both physical as well as mental challenges. Although we can't completely mitigate these challenges, we can still contribute to improving their life and decreasing their challenges. In our project, the biggest obstacle we have faced concerns accuracy in object detection model since the current best available model can also be misleading sometimes, and the popularity of touch-based applications which decreases the usability of these applications for visually impaired people popular applications like lookout by google envision Ai all have the visual element in their interfaces which creates a major challenge due to absence of non-visual output modalities Therefore we have tried to move from touch-based application towards complete voice automated to greatly reduce the usability of application on visual elements Our voice assistant Alan Ai does this job quite well and enable the user to operate the applications through voice commands. There is still limitation though like noise cancellation.

Our application is a basic prototype of combining the basic necessary features automated through voice to make these applications more user-friendly for visually impaired people and make them more convenient to use and increase their usability.

## 6.3 Environmental

- Economic: - Once the visually impaired people become able to live their life with more ease, they would be able to contribute more to the economy as they would be able to move to independence. A blind person originally had very limited resources to increase his/her knowledge as only 7 per cent of published books are available in formats that help visually impaired readers. Now, by the help of technology they would be able to read and explore more. Having an application which is completely working by audio commands, would reduce the dependence on vision and could prove to be a boon for the blind as now they could carry out their work without any support. An adult whose vision gets treated/blindness revered can take care of his family in a better way and also

contribute to the economy. Also, working in any organization won't be a hurdle for them as they could do most of their work with no dependence on anyone for any help. Some study suggests, every $1 invested in blindness elimination generates approximately $4 in economic gain in low-resource countries. The system is not very costly and can bring economic benefit by avoiding costly mistakes.

- Social: - There are three main level of difficulties when it comes to disabilities. First includes physical limitation of reading and moving around freely using eyesight for which people usually undergo Braille and mobility trailing to overcome this barrier. But these methods can only compensate them for their loss to some extent and may prove to be useless in some scenarios like reading text which are not available in Braille, moving to some unfamiliar place where they can't make use of their mobility training and so on. In today's age of digitalization, a lot of things are being digitalized which raises concerns for these people. Digital documents are now more dominant than text documents. Also, a lot of text like advertisements, posters etc. can't be all made available in Braille which creates a social barrier. Second limitation is environment oriented like inaccessible transportation, discrimination, social stigma. Third is medical limitation which includes requirement of extra support by these people. But providing unwanted support can make these people interpret some kind help as sympathy and lead to misunderstanding raising issues like low self-confidence and low self-esteem. The best way to avoid this is to make them move towards greater self-independence which can be achieved by the use of technology. Having this application in their hand they won't have to search for helping hands every time and will become independent in executing their day-to-day activities boosting their self-confidence. Also, a more confident person would become a major contributor in his/her society.

## 6.4 Future Work Plan

Our next step will be to improve the accuracy of the object detection and text detection by comparing and using different other model available over the internet Once the model gives satisfactory accuracy, research will be continued

to decide the best model for our response generation step. We also plan to add object detection in the stick using raspberry pi3 Model B for finding the obstacle and to find distance from the object Pi camera is used to capture the video frames and feed each frame for processing. For real time object detection, neural network will be used to train the images.

# PROJECT METRICES

## 7.1 Challenges Faced

During the entire process, starting from project planning to the final testing of our software and hardware model, there were various challenges we faced.

1. The first and major challenge was to create a setup for flutter and android studio as it required VPN setup to be made.

2. In the navigation module, we faced challenge in choosing such an Arduino board which would be best suitable for transferring the distance data to firebase, and also to be affordable.

3. We faced challenge in adapting to each new Arduino board which we tried to work with.

4. We faced issue in the accuracy of the object detection as well as the number of objects it can identify.

5. we faced issue in the accuracy of the text detection as well as voice assistance.

But with the help of our mentors and the process of literature survey, we were able to identify the novelty of our project. with regular testing and modifications in our code, we were able to achieve a consistent and accurate working model.

## 7.2 Relevant Subjects

Table 24: Subject Name and Subject Code

| S.no. | Course Name | Course Number |
|---|---|---|
| 1 | Natural language processing | UML602 |
| 2 | Machine learning | UML501 |
| 3 | Deep Learning | UCS742 |
| 4 | Software Engineering | UCS503 |
| 5 | Computer Vision | UCS532 |
| 6 | Flutter (Dart) | N/A |
| 7 | Electronic Engineering | UEC001 |
| 8 | Engineering Design Project-I | UTA012 |
| 9 | Engineering Design Project-II | UTA014 |

## 7.3 Interdisciplinary Knowledge Sharing

In our project, we have tried to cover various parameters. It required in-depth knowledge of Flutter (Dart Programming language), Deep Learning, machine learning, computer vision, natural language processing, for the software part and for the hardware part it required in depth knowledge of Arduino programming and different sensors for the proper functioning of the model developed. Hardware device has been activated through our mobile app only using NodeMcu TCP connection.

## 7.4 Peer Assessment Matrix

Table 25: Assessment Matrix

| | | **Evaluation of** | | | |
|---|---|---|---|---|---|
| | | Raju Kumar Gupta | Swapnil Kumar Singh | Riya Singhal | Aditi |
| **Evaluation By** | Raju Kumar Gupta | 5 | 5 | 5 | 5 |
| | Swapnil Kumar Singh | 5 | 5 | 5 | 5 |
| | Riya Singhal | 5 | 5 | 5 | 5 |
| | Aditi | 5 | 5 | 5 | 5 |

## 7.5 Role Playing and Work Schedule

Below are the roles of members:

1. Riya Singhal- Project analysis and design, hardware analysis and design, Research on how to make the application more and more user friendly by adding voice assistant.

2. Raju Kumar Gupta- Project Analysis, document formatting and report work, Hardware analysis and design.

3. Aditi- Preparation of base flutter app and modules, integration of all the modules and Voice Assistant with the base flutter app and

Compare various models and find the best model / API for each of text detection and object identification.

4. Swapnil kumar Singh- Preparation of modules, Project idea and design, Poster design, compare various models and find the best model for object identification.

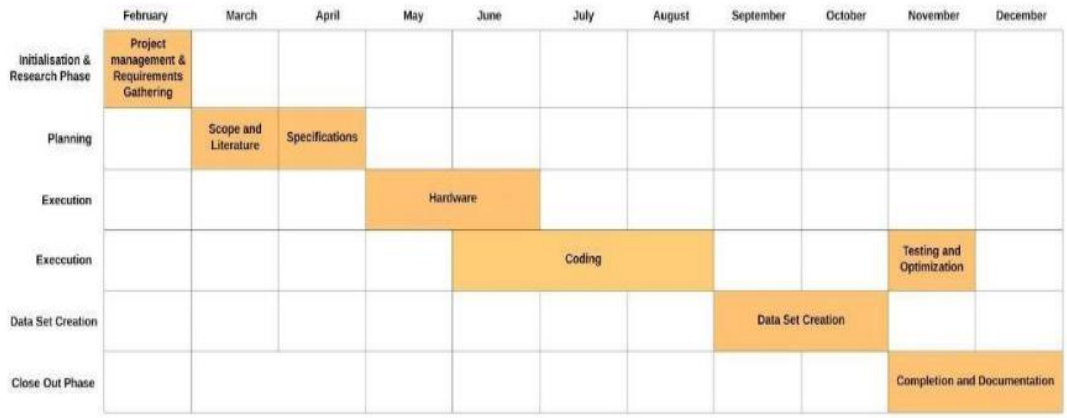



Fig 26: Gantt Chart

## 7.6 Student Outcomes Description and Performance Indicators

Table 26: Students outcomes and performance indicators

| S0 | Description | Outcome |
|---|---|---|
| B1 | Identify the constraints, restrictions and models for the problems | Various restrictions were identified by us in software as well as hardware. |
| B2 | Use tested methods, tools and techniques for object and text detection. | Various module and APIs were tried by us to ensure which ones fit our needs for our software. |
| B3 | Analyze and interpret results with respect to assumptions, restrictions and theory. | Integration of hardware and software were tested multiple times. |
| D2 | Can play different roles as a team player. | All the team members worked on both, the working of device as well as the documentations' part. |
| E3 | Compared various models and APIs. | Implemented and tested various models/APIs to get the desired accuracy. |
| F1 | Produce a variety of documents such as laboratory or project reports using appropriate formats. | All of the team members worked on documentation process and ensured the value of time is maintained. |
| F2 | Deliver a clear and effective oral presentation. | The team delivered PPTs on various occasions to show the progress to our mentor and the panel. |
| G1 | Apply various codes and algorithmic processes. | Various codes and conversions were used while writing the code. |

## 7.7 Brief Analytical Assessment

**Q1. What sources of information did your team explore to arrive at the list of possible Project Problems?**

**Ans:** We studied various Research papers, visited different websites and consulted our mentor to arrive at the list of possible project problems.

**Q2. What analytical, computational and/or experimental methods did your project team use to obtain solutions to the problems in the project?**

**Ans:** we did literature survey and analyze their shortcomings, we also looked at different resources like research papers, journals, websites etc. and consulted our mentor to obtain solutions to the problems in the project.

**Q3. Did the project demand demonstration of knowledge of fundamentals, scientific and/or engineering principles? If yes, how did you apply?**

**Ans:** Yes, our project demanded demonstration of various principles. It includes development of different modules, integration of them with base flutter app and testing of our approved objectives, code debugging, hardware - software synchronization and many more engineering principles.

**Q4. How did your team shares responsibility and communicate the information of schedule with others in team to coordinate design and manufacturing dependencies?**

**Ans:** All the work was divided equally among the team members. As all the members were from the same branch group and had already worked together previously on various projects, there was no communication gap between us.

**Q5. What resources did you use to learn new materials not taught in class for the course of the project?**

**Ans:** Our mentor helped us a lot during the project. Whenever we reached a dead end, he was always there to rescue us. Various research papers were also used to learn new concept that were not taught previously in the class. And also, we learned new courses on various platform like YouTube, Udemy.

**Q6. Does the project make you appreciate the need to solve problems in real life using engineering and could the project development make you proficient with software development tools and environments?**

**Ans:** Engineering has evolved and gained importance as a career option only because of it's applicability in solving the real world problems. We got a chance and platform to make use of our engineering principles to solve the problems faced by visually impaired people. Using the concepts of machine learning, IOT and application development concepts we were able to solve a real-world problem.

The project development skills helped us a lot in planning our work plan and successfully executing it on time. It kept reminding us of the deadlines and helped us in easily distributing the work amongst the whole team.

# APPENDIX A: References

[1] Al-Razgan M, Almoaiqel S, Alrajhi N, Alhumegani A, Alshehri A, Alnefaie B, AlKhamiss R, Rushdi S. A systematic literature review on the usability of mobile applications for visually impaired users. PeerJ Comput Sci. 2021 Nov 22;7:e771. doi: 10.7717/peerj-cs.771. PMID: 34901428; PMCID: PMC8627222.

[2] S. Khusro, B. Niazi, A. Khan and I. Alam, "Evaluating Smartphone Screen Divisions for Designing Blind-Friendly Touch-Based Interfaces," 2019 International Conference on Frontiers of Information Technology (FIT), 2019, pp. 328-3285, doi: 10.1109/FIT47737.2019.00068.

[3] E. Ghidini, W. D. L. Almeida, I. H. Manssour and M. S. Silveira, "Developing Apps for Visually Impaired People: Lessons Learned from Practice," 2016 49th Hawaii International Conference on System Sciences (HICSS), 2016, pp. 5691-5700, doi: 10.1109/HICSS.2016.704.

[4] Mahajan, Sameer & Kulkarni, Nahush. (2020). IRJET | A Conversational News Application Project using Artificial Intelligence based Voice Assistance. 7. 3779-3782.

[5] D. K. S. Neoh et al., "PicToText: Text Recognition using Firebase Machine Learning Kit," 2021 2nd International Conference on Artificial Intelligence and Data Sciences (AiDAS), 2021, pp. 1-5, doi: 10.1109/AiDAS53897.2021.9574187.

[6] C. Liu, Y. Tao, J. Liang, K. Li and Y. Chen, "Object Detection Based on YOLO Network," 2018 IEEE 4th Information Technology and Mechatronics Engineering Conference (ITOEC), 2018, pp. 799-803, doi: 10.1109/ITOEC.2018.8740604.

# PLAGRISM REPORT

VIRTUAL EYE: A Perfect Companion for Blind

ORIGINALITY REPORT

| 12% | 9% | 2% | 9% |
|---|---|---|---|
| SIMILARITY INDEX | INTERNET SOURCES | PUBLICATIONS | STUDENT PAPERS |

PRIMARY SOURCES

| | | |
|---|---|---|
| 1 | www.coursehero.com<br>Internet Source | 4% |
| 2 | Submitted to Netaji Subhas Institute of Technology<br>Student Paper | 1% |
| 3 | archive.org<br>Internet Source | 1% |
| 4 | Submitted to Fort Valley State Univeristy<br>Student Paper | 1% |
| 5 | Submitted to University of Mauritius<br>Student Paper | 1% |
| 6 | www.engr.sjsu.edu<br>Internet Source | <1% |
| 7 | www.ijitee.org<br>Internet Source | <1% |
| 8 | www.rguhs.ac.in<br>Internet Source | <1% |
| 9 | transmitter.ieee.org<br>Internet Source | <1% |