AsSee : ASSISTIVE MOBILE APPLICATION FOR VISUAL IMPAIRED INDIVIDUAL

LIM WEI HAO

UNIVERSITI TUNKU ABDUL RAHMAN

AsSee: ASSISTIVE MOBILE APPLICATION FOR VISUAL IMPAIRED INDIVIDUAL

LIM WEI HAO

A project report submitted in partial fulfilment of the requirements for the award of Bachelor of Science (Honours) Software Engineering

Lee Kong Chian Faculty of Engineering and Science
Universiti Tunku Abdul Rahman

Octorber 2024

DECLARATION

I hereby declare that this project report is based on my original work except for citations and quotations which have been duly acknowledged. I also declare that it has not been previously and concurrently submitted for any other degree or award at UTAR or other institutions.

Signature : 

Name : LIM WEI HAO

ID No. : 2200446

Date : 2/10/2024

APPROVAL FOR SUBMISSION

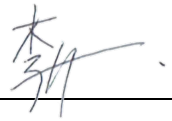
I certify that this project report entitled **"AsSee : Assitive Mobile Application for Visual Impaired Individual"** was prepared by **LIM WEI HAO** has met the required standard for submission in partial fulfilment of the requirements for the award of Bachelor of Software Engineering with Honours at Universiti Tunku Abdul Rahman.

Approved by,

Signature : 

Supervisor : Dr. Faranak Nejati

Date : 2/10/2024

Signature : 

Co-Supervisor : Dr. Lee Ming Jie

Date : 2 October 2024

ABSTRACT

The numbers of visually impaired individuals increase year by year but in this era of developing rapidly, the mobility and safety of visual impaired individuals issue during travel raised due to unfriendly environments and limited information about the surrounding of the environment. There are numbers of assistive technology ready to assist them in different perspectives, but the price which is not affordable for majority of the visual impaired individual and required extra equipment to take advantage of the assistive technology. Therefore, there is a need of assistive mobile application for visual impaired individuals to assist them in gather information of the environment and feedback to them accurately for better understanding of the environment and decisions making. Furthermore, the usage of the paper and documentation in most of the procedures in daily lifes where is a big challenges for them to understand written documents and texts in not in braille without any assistance which lead to need of application that helps user to capture written texts and convert into speech to let them gain understanding and information on the paper is important where could assist them in daily life activities and in some important procedures. In conclusion, this project aims to provide a mobile application to assist user in real-time environment object detection to assist them gather information of the environment and the text recognition and convert to speech for them accessible to the information of written text.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF SYMBOLS / ABBREVIATIONS

WHO	World Health Organization
API	Application Programming Interface
SDLC	Softwaere Development Life Cycle
GPS	Global Positioning System
AI	Artificial Intelligence
ML	Machine Learning
SDK	Software Development Kit
IoT	Internet of Things
JVM	Java Virtual Machine
UC	Use Case
UI	User Interface
COCO	Common Objects in Context
SUS	Software Usability Scale
UAT	User Acceptance Test

LIST OF APPENDICES

CHAPTER 1

INTRODUCTION

1.1 General Introduction

Vision is one of the important senses in our daily life, without vision we will be more tougher during every important stages of life. We own our vision since our birth, we use our vision to learn, identify, gather information and make decisions. Vision play a critical role in our life, its help a lot in our daily lives, help us to identify risk, alphabets, color and appearance of an object. Without vision, we would find it difficult to recognize objects and truly experience the color of color. Not everyone granted with perfect vision, there are 2.2 billion of people was vision impairment based on WHO analysis (WHO, 2023) .

There are many difficulties in the lives of people with visual impairments that we cannot imagine. In their lives, the colors are not so bright, the shapes of objects cannot be fully recognized, and even in the most ordinary walking process, they encounter more difficulties than we imagine. Even though,in daily life there are a lot of design was intended to help them like braille,tacticle paving, screen reader and talking book to help them to be more comfortable and convenient in their life, but there are still a lot of challenges that visually impaired people will face in daily life (Battle For Blindness, n.d.)

The challenges that may faced by visual impaired people like:

- Access to Information : The information on printed materials, digital content, braille-less content or any format that is inaccessisble by them.
- Mobility and Navigation : Navigating to a unfamiliar space is a challenge for them, especially those space without signage, tactile paving, and full of obstacles.
- Transportation : The platforms that is not maintained well may lead to inaccessible to public transport, and the public transport without audio notifications will also affect the transportation experience of visually impaired people.

- Technology Accessibility : Inaccessible technology, websites, applications, and devices may affect the experience of the visual impaired people to technology and limited their access to technology.
- Daily Life Activities : The daily life activities might be a big challenge faced by visually impaired people where they need longer time to identify daily necessities like kitchen utensils, personal care products and etc to complete their daily activities (Envision, 2024).

Therefore, a mobile application that provide assistant to visually impaired community inspired. This technology aim to provide a assistive mobile application for visually impaired make it easily accessible and to use. This technology can help visual impaired individuals to overcome many problems and challenges in daily life and ensure their convenience and independence in common activities.

1.2 Importance of the Study

AsSee which is the combination of Assistant and See, can easily tell by the name which aim to assist in visual. The introduction of AsSee important because existing assistive technology for visually impaired individuals is very expensive and not affordable for everyone, therefore, with this study, we can have a technology to help those who needs assistance during their daily life with affordable price.

Besides, the lacking of technology in helping visually impaired communities also make this study important. With the introduction of this study can raise the attention of public toward technology for visually impaired. Furthermore, there are a lot of challenges that faced by visually impaired individuals in real life where they required assistance to help them to overcome these challenges or reduce the difficulty of their daily task.

The introduction of the AsSee System will increase end user safety and mobility with provide assistance for them to gather information surrounding where could help them to understand the surrounding environment and take corrective action to deal with object nearby. In addition, with the text recognition and text to speech function, the user was able to access and understand the written text and help them accessible to the information.

Therefore to ensure the safety issues which is the main concern when developing technology for visually impaired individuals. There are potential risk surrounding them which cause them hard to move even an inch. Existing aids like Smart Cane with limited coverage of sensors where couldn't identify the dangerous from above waist since Smart Cane sensor integrated within the white cane where helps user to identify object on the ground. Besides, current assistive technology could increase the mobility and safety of the visually impaired individuals but relying on precise sensor and multiple sensor integrated make the technology is unaffordable and inaccessible like WeWalk Smart Cane.

In conclusion, to arise the awareness of the needs of application for visually impaired individuals and to assist them in different situations in daily life and improve their safety, mobility and accessibility to information make the development of the AsSee application is important.

1.3 Problem Statement

There are problems to overcome in order to create a assistive technology that could assist the visual impaired individual effectively and in certain situations that required assistance. Therefore, there a problems that identified and need to resolve in this project to ensure that the technology enables targeted users to be more convenient and assisted in daily travel process.

Gap in Current Solutions:

a. Current assistive mobile applications require users to take pictures to recognize objects and lack real-time object detection
b. Current assistive mobile applications require users to take pictures to recognize written text, rather than providing real-time text recognition and feedback.
c. Current assistive mobile applications are designed with complex interfaces, requiring too many steps and interactions.
d. Current assistive mobile applications offer only one mode of interaction, which creates challenges for users when engaging with the app.

To mitigate the problem statement and the development of assistive mobile applications, AsSee to enhance the mobility and independence of the visual impaired individuals in daily life, by helping them in detect objects surrounding them, which by providing the real-time object detection which existing assistive technology could’nt provide such services.

Besides, the current assistive technology where required user to perform take photos for object detection and text recognition, where it does not provide the real-time assistance to the user. Furthermore, the take pictures action might not be able successfully perform by the user and to understand the surroundings, there are a lot of photo needed to be captured which a lot of actions need to be perform by the user and this indicates the needs of real-time object detection for visually impaired individuals to understand their surrounding better and without too much actions needed.

Furthermore, the current assistvie technology also required to perform take photo function to recognize written text, therefore, with the real-time text recognition implementation where reduced the steps needed to be taken to perform text recognition function and make visually impaired individuals to access information from the written text. Therefore, to implement the real-time text recognition function which does not require complex and many actions to access information from written text.

In addition, the current assistive technology where provide only one method to interact with the application, which its limited the user forced to interact with all those interface like buttons to enjoy the functionality and features, therefore with the speech recognition implemented, it provide alternative method for the user to interact with the application and enjoy the features and function without complex interface needed to interact with.

In conclusion, AsSee was motivated by to enhance the mobility and independence of visually impaired individuals in daily life by addressing the gap and shortcoming of the current assistive to provide a better and easier ways for the user to interact with the applications and enjoy the AI powered services to help them in different situations. With all the real-time functions

provided by the application and simplified interface with speech recognition support to interact with the application. AsSee was delivered to assist them in more efficient and effective way.

1.4 Aim and Objectives

I. To implement AI-driven real-time object detection to detect object for user.
II. To implement AI-driven real-time text recognition to recognize text.
III. To design interface for user to interact with application.
IV. To implement AI-driven speech recognition to accept command from user.
V. To perform testing and evaluation of the integrated AI-driven object detection, text recognition, and speech recognition within the application to evaluate it meet user expectations and satisfaction.

There are different methods used to evaluate whether the application meets its objectives. First and foremost, unit testing is conducted to ensure that the application's functionality is working as expected and that the individual components perform well. Additionally, the System Usability Scale (SUS) is applied to evaluate the usability of the system, ensuring it aligns with the objectives and works effectively in a real-world environment. Finally, User Acceptance Testing (UAT) is performed to ensure the project meets user expectations and satisfies the project scope, confirming that the software is ready for release.

1.5 Proposed Solution

This system will aid the visual impaired in few situations:

I. Object Detection: This system will help the user to identify object real-time and feedback to user.

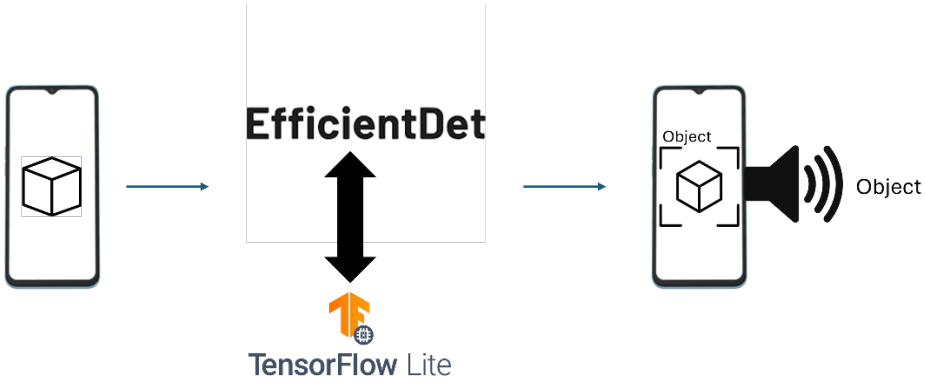


Figure 1.1 Object Detection Work Flow

II. Text Recogntion: Real-time identify text and convert it into speech so that the user can understand text is not written in braille.

Figure 1.2 Text Recognition Work Flow

III. Speech Recognition: Accept command from user to interact with mobile application.

Figure 1.3 Speech Recognition Work Flow

IV. Interface: Allow user to interact with application with interface.

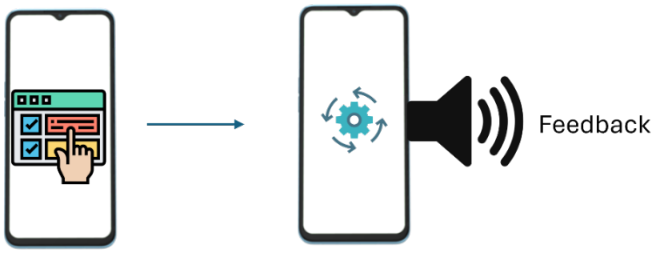


Figure 1.4 Interface Mechanism

1.6 Proposed Approach

Prototyping methodology applied to AsSee system development. Prototyping is one of the famous SDLC that widely use. There are 6 phases in this methodology which at the start of the development, it will gather all the requirements and after all requirements gathered and analyzed, will proceed to quick design phases which it will design a prototype that contain main features or functions of the system. After the quick design complete, the first version of prototype will started to draft and pass to stakeholders to testing and the feedback will be gather and for improvement for the prototype. After the user evaluate the prototype, the development team will refine and review the prototype of the system based on the feedback given and then these two processes will repeat until the system is good to deploy for develop a true product for implementation and the maintenance will always done periodically to minimize the downtime and prevent large-scale of error occur in the system. The prototyping able to ensure the quality of the outcome fulfilled customer requirements and always adapt to the customer requirements based on the involvement of stakeholders in the development process.

1.7 Project Scope

I. Develop an application that provide object detection function.
II. Develop an application that provide text recognition function.
III. Developp an application that integrates speech recognition function.
IV. Develop an application that provides interface for user to interact with application.

This section discusses the scope of the application where it aim to develop an application where occupied with object detection function to help user to detection object. Furthermore, there is implementation of text recognition to help users to recognize text which is not written in braille. Furthermore, the integration of speech reocgition that allows users to interact with the application which an alternative method to communicate with the application. Instead only rely on the speech recognition to interact with the application, interface also design and prepare for the user to interact with the application.

1.7.1 Assumption

i. User is using Android devices.

ii. User is visually impaired individuals but not deaf-blind.

1.7.2 Targeted User

Visual impaired individuals who needs the aid in daily life and activities. AsSee system aim to provide different assistance in daily life to improve the quality of life and the independence of visual impaired individuals in daily activities. Besides, to increase the mobility of the end user when discovering unfamiliar area and identifying potential risk surrounding them.

1.7.3 Main Features

1.7.3.1 Object Detection

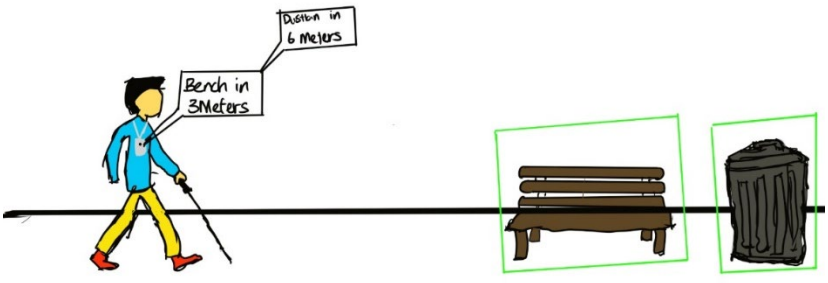


Figure 1.5 Object Detection Features of AsSee Prototype

This features aimed to help user to identify the object nearby the user and notify them with objects details in real-time and without any other additional actions needed to be taken to detect the object surrounding them where it enhance the usability of the application and increase the independence in their daily life. This module detect the object from the live-stream frame that capture by the camera which it detect the object and return the information to the user in real-time with AI-powered model, it enhance the mobility of the visual impaired individual.

1.7.3.2 Text Recognition

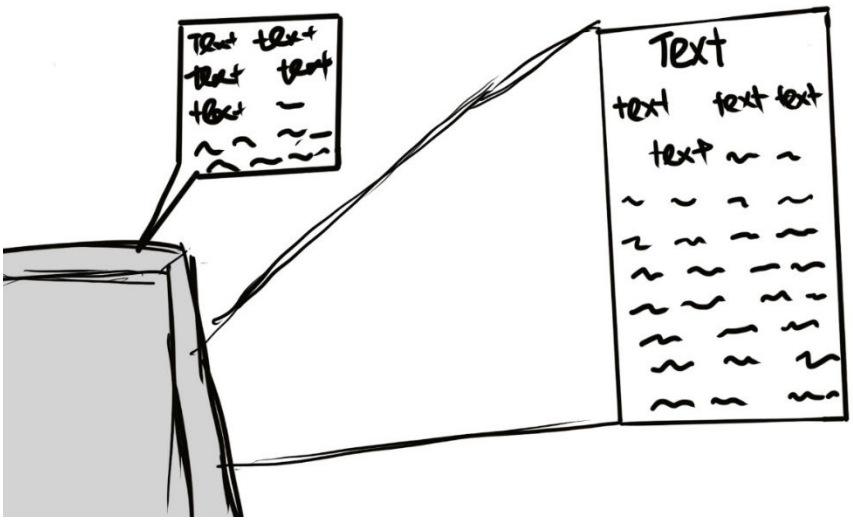


Figure 1.6 Text Recognition Features of AsSee Prototype

This features is to help user to understand written text and convert to speech to let user understand and access information from the text that is not written in braille in real-time and doesn't required complex take photo action to recognizing text.

1.7.3.3 Speech Recognition

This features receive command from the user and do corresponding action to ensure the user could interact with the application in an alternative ways instead of relying on the interface like buttons etc. With the command received, the AI-driven model will recognize the speech and react to the speech recognized.

1.7.3.4 Text-to-speech

This module was cooperating with other modules to convert the information from the AI models into speech and provide feedback to the user for object detected, text recognized, and feedback based on the command made from the user.

1.7.4 Limitations

I. Enhance the accuracy of the Object Detection AI Model.
II. Enhance accuracy of Speech Recognition AI Model coverage of languages and dialects.
III. Establish connections between different smart devices to integrate more functions.
IV. Integrate speech-based navigator and location tracking systems to enhance mobility and navigation capabilities.
V. Integrate real-time scene description to feedback surrounding to user.

CHAPTER 2

LITERATURE REVIEW

2.1 Introduction

This section reviewed the existing technology and tools to have more understanding of this project potential shortcomings to overcome and select the most appropriate tools and methodology to applied in this project.

2.2 Literature Review

Literature Review is conducted to have deep insight into the existing technologies and methodologies that could be used or referred to within the development process of AsSee. An appropriate methodology and lesson learned from existing technologies is very important to produce a better outcome.

2.2.1 Existing Assistive Products

This section will review the existing assistive product, the assistive product reviewed is aimed to have better insight and understanding or the needs of the potential user and try to figure out a way to integrate them together to produce a harmony technology that could provide assistance in different perspectives.

2.2.1.1 Computer Vision Based Assistive Technology

Computer Vision is a field of Artificial Intelligence (AI) that uses machine learning and neural networks to teach the computer and system understand digital content like images, videos and other visual input and translate to information.The computer vision enables computer to see, observe and understand and it also works as human vision (IBM, n.d.). Therefore assistive technology with computer vision based can work as eyes for visual impaired person. (Sivan and Darsan, n.d.)

2.2.1.1.1 Travel Aid for Visually Impaired

This is to increase the safety during the travel process of visual impaired individual.Besides, the mobility and operationcan be increase by this assistive technology. It is a common challenges that faced by visually impaired individual during their travel process,there are a lot of product invented to

increase the mobility and convenient during the travel process (Tian et al., 2013). There are two different travel aid for visually impaired:

i. Indoor

The indoor travel aid is to help the visually impaired person to understand the unfamiliar indoor environments. This help the user to access information to gather information of the surrounding and this was the guide for the user to the place that user intended. Floor, door, toilet, lab, and etc will be recognize and feedback to the user using voice to increase their mobility and operation indoor. There are few main features of Indoor Travel Aid

- Sign : To assist user to understand sign in indoor like Exit, Laboratory, Escalator and etc.

Figure 2.1 Travel Aid for Visually Impaired - Sign Recognition (Tian, Y., Yi, C., & Arditi, A.,2010)

Door : To assist the user to understand doors to help them identify door.

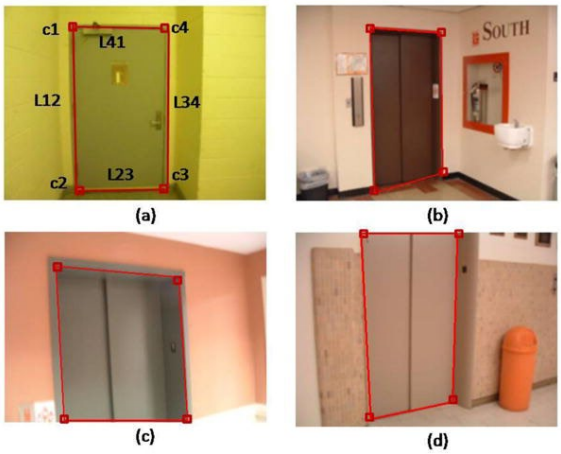


Figure 2.2 Travel Aid for Visually Impaired - Door Recognition (Tian, Y., Yi, C., & Arditi, A.,2010)

- Inset and protrude : There are a lot of object in indoor that have inset or protrude, like door, escalator, cabinet and etc.

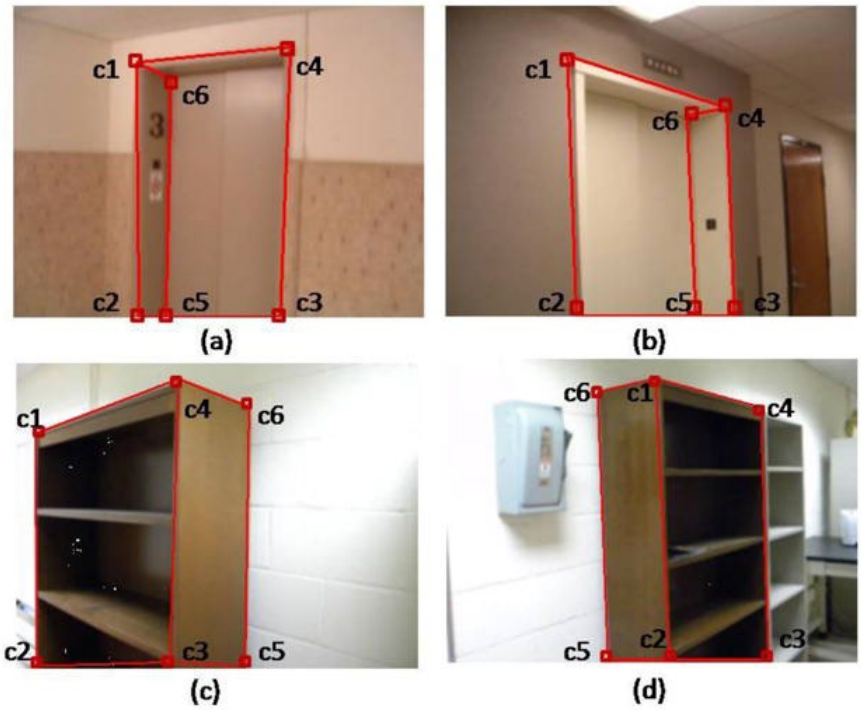


Figure 2.3 Travel Aid for Visually Impaired - Inset and protrude recognition (Tian, Y., Yi, C., & Arditi, A.,2010)

- Text Recognition : The text in indoor also recognizeable to convert it into information for user. The information like floor numbers, men and women, and etc.

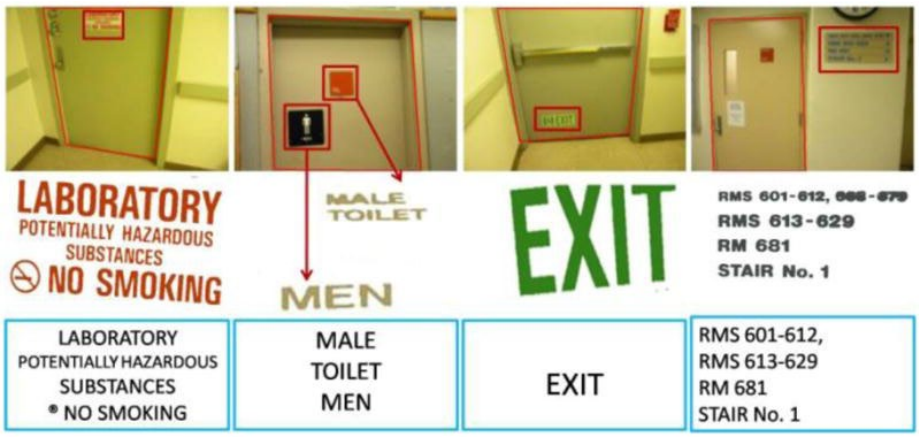


Figure 2.4 Travel Aid for Visually Impaired - Text Recognition (Tian, Y., Yi, C., & Arditi, A.,2010)

ii. Outdoor

The Outdoor travel aid is aim to assist visual impaired individual during their outdoor travel, the are a lot of challenges and problems may encounter by them, thefore the outdoor travel aid assistive technology will help them to increase their mobility and safety at outdoor. The assistance will provide in different type of mode, like smart hat, smart belt, smart shoe and etc.

- Obstacles detection : Help user to identify the obstacles at infront to avoid any accident happends.
- Path detection : Help user to identify incoming path is there any holes, branches, and other potential dangerous on path.
- Information gathering : Help user to gather information nearby to help user make corresponding actions based on information collected and feedback return to the user.

2.2.1.1.2 The vOICe

Figure 2.5 The vOICe (1996,Peter B.L. Meijer)

“See with your Ears” which is the aim of The vOICe technology slogan, which it aim to process the image and convert it into voice format and let the user to see the world by using their ear. Other than only simply convert the image to sound, the vOICe technology also relate the brightness to loudness, which make the user have better understanding the surrounding.

The vOICe technology was able to installed in most of the Android Based Smart Glasses, therefore it is much more affordable and doesn’t required additional purchasing to have more function on smart glasses. With the introduction of the vOICe technology, there are some activities that previously was impossible was also enjoyable for visual impaired individuals

like photography, it let user to understand the surrounding better even the brightness and details it enables user to produce better photo with this technology. (The vOICe, 2024)

2.2.1.1.3 Finger Reader

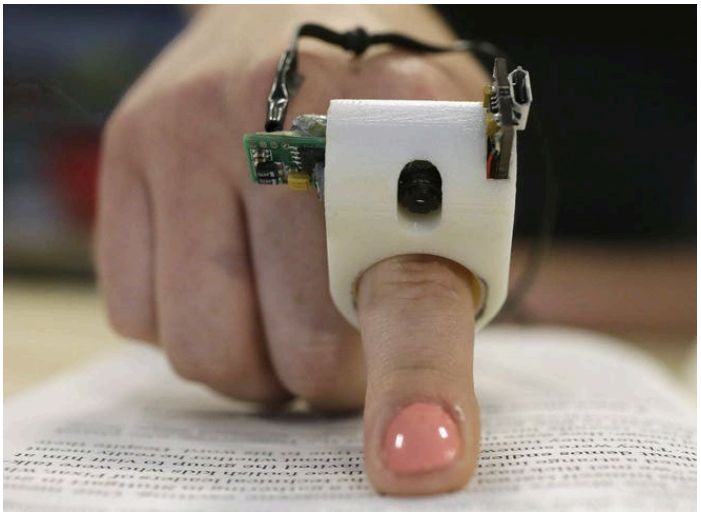


Figure 2.6 Finger Reader (Boldu et al., 2018)

Finger reader which is a ring that wearable and integrated with camera, speaker (MIT Fluid Interfaces Media Lab, n.d.). When the user points at the content of books, menu and etc. the camera will capture the text and read the identified text. With the finger moves along the page, the finger reader able to speak out a whole complete sentences and paragraph. The Finger reader will notify user when user deviates from the current line or reached end of the line throught the integrated vibrator or sound notification.Besides, not only book and menu can be read, but words on packaging, money notes can also be read.(Boldu et al., 2018).

Figure 2.7 Finger Reader - Packaging Recognition (Boldu et al., 2018)

Figure 2.8 Finger Reader - Bank Notes Recognition (Boldu et al., 2018)

2.2.1.2 Braille Technology

Braille is a system or writing that uses dots or bumps which also knows as cell on material to represent letters and numbers to get visual impaired person to read, identify and understand the book, or any other informative stuff. Besides, with special arrangement of the dots, it represent different words and with the combination, it can be words, sentences or paragraph.This system helps visual impaired person to read and write independently (InFocus, 2022).

Braille technology was aim to convert information to Braille form which able visual impaired person to understand the information and even though it is digital content like website, blogs and etc.

2.2.1.2.1 Braille Note Takers

Braille Notetakers is a small and portable device to store the information (Fable, n.d.). Every different Notetakers will run on different operating system, processor, memory and etc. The notes stored was accessible through the built-in speech synthesizer, braille display and etc. The braille notetakes normally use to advanced word processing, web browsing, record phone number, names, track appointment and taking notes.It also knows as Personal Digital Assistants before the widely use of smartphone or tablet (AFB, n.d.).

Figure 2.9 Braille Note Takers

2.2.1.2.2 Braille Printer

Braille Printer which also known as Embosser receive data from computers and emboss it into braille and print onto paper (Carmen Willings, 2024). With the control of solenoids, the embossing pins was able to emboss the data onto the paper. The printer work slower and noiser than a regular printer and the paper needed to print the outcome also heavyweight than regular. Mostly printer only print one side, but there are interpoint printer emboss braille on both sides of pages with more expensive price.

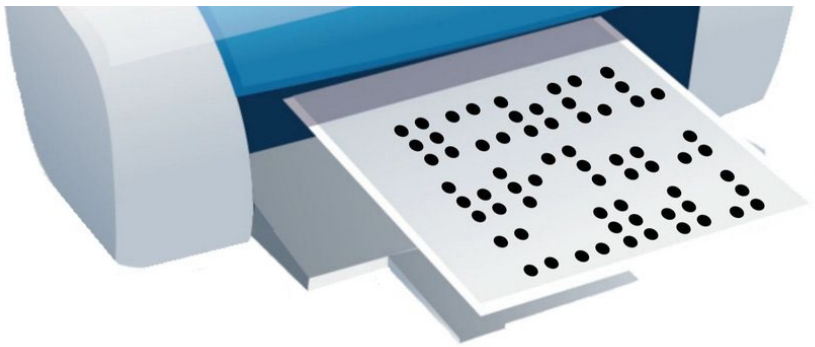


Figure 2.10 Braille Printer

2.2.1.2.3 Refreshable Braille Display

Figure 2.11 Refreshable Braille Display

Braille Display is design for the visual impaired person to read and understand the content in the digital format like webpages, documents, social media, blogs and etc. Braille Display sync and read the connected device screen by collaborate with screen reader.The Braille Display was consist of few parts (TSBVI, n.d.):

i. Refreshable braille cells

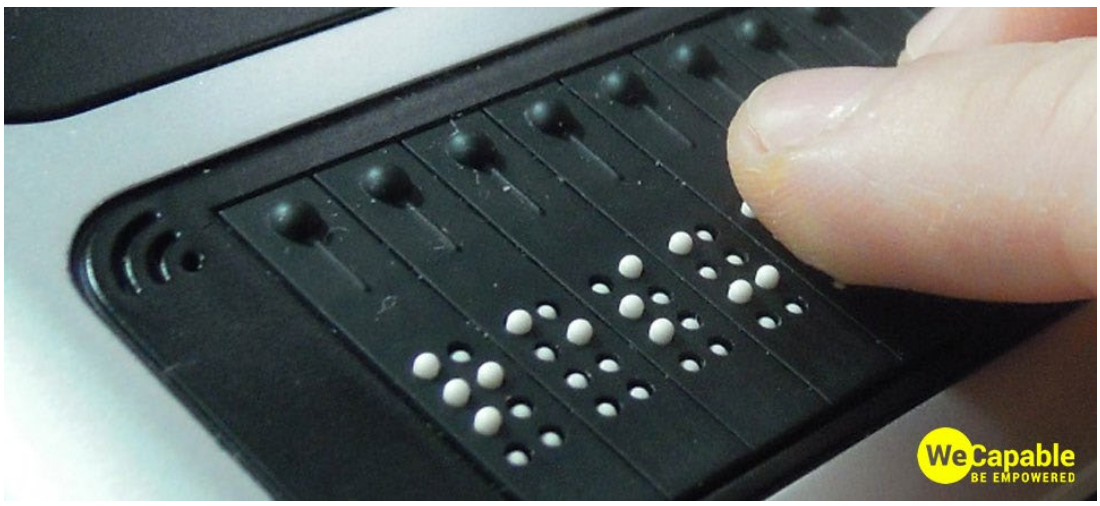


Figure 2.12 Refreshable braille cells

These cells display braille characters when the screenreader read the screen and raise or lower the small pins to form Braille characters and make user able to understand the sceen of electronic display. The pins will be real-time to

output the text, which able the user to understand the digital content and read email, browsing website and etc.

ii. Cursor routing buttons

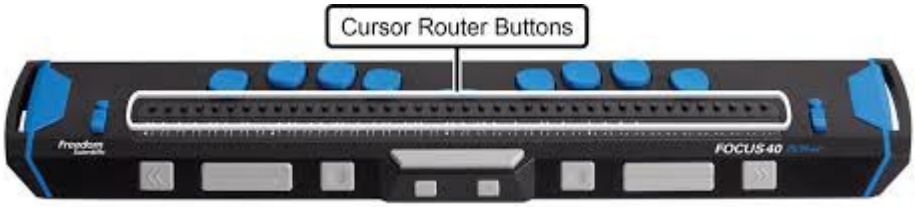


Figure 2.13 Cursor routing buttons

The cursor routing buttons allow the user to move or change the position of the cursor as user intended to complete task more effectively and efficiently.

iii. Perkins-style keyboard used for braille entry

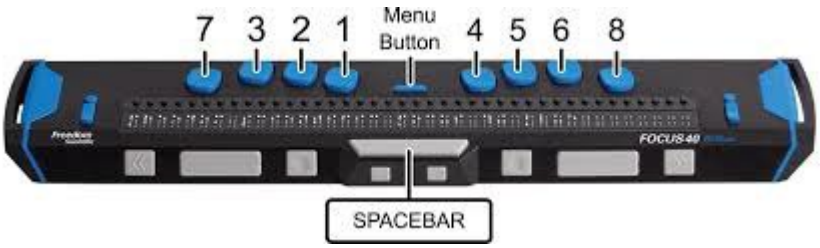


Figure 2.14 Perkins-style keyboard used for braille entry

The Perkins-style keyboard helps user to input into the devices with the collaboration with Dot 7 Key which is Backspace and Dot 8 Key which is Enter Key. Besides, there are spacebar key not only works as standard space bar but use during entering commands.

2.2.1.3 Smart Cane

Smart Cane is an assistive technology to help visually impaired person. There are different type of function that provided by a Smart Cane with different sensor integrated. There are several function that mainly focused by Smart Cane which is:

i. Mobility and Operation

The Smart Cane was aim to increase the mobility and operation of end user. The walking process was a difficult task for visual impaired person, and there are a lot of challenges even a short distance walking process. Therefore, the smart cane was invented to increase the mobility and operation of end user to make sure they more secured and convenient during the process.

ii. GPS

The Smart Cane integrated Global Positioning System could help other to identify the location of the smart cane to always monitor the smart cane user. Besides, it also can help others find the lost smart cane by using the GPS integrated. But GPS was mainly used to collect the information nearby and help user to understand unfamiliar spaces.

iii. Gather Information and Feedback

The Smart Cane normally was designed to help user to gather the surrounding information like obstacles on ground and return feedback to raise user attention. Besides, GPS based to collect nearby information and return to user to let user know shops nearby.

iv. Obstacles Detection

Every Smart Cane was designed to help end users to detect obstacles on ground like holes, tactile ground, trash and etc. This is to ensure the safety of the mobility of the user, even though white cane can help user to detect the obstacles, but with the sensor integrated smart cane, it could identify the obstacles earlier and more accurate.

2.2.1.3.1 SmartCane Device

The SmartCane Device helps a visual impaired person to detect obstacles on ground like holes, surface and etc. The SmartCane Device integrated ultrasonic to make sure the detection is accurate and return distance information to the user so that the user can make corresponding action to the obstacles detected. In order to notify the user, it also integrated vibrator to raise user attention. (Indian Institute of Technology Delhi, n.d.)

Figure 2.15 SmartCane Device

2.2.1.3.2 WeWalk Smart Cane

This is a product that aim to enhancing the mobility of visually impaired person. The WeWalk provided few revolutionary features through WeWALK smart cane and application. Besides, WeWALK smart cane let user to interact with WeWALK smart cane features using built-in touchpad. With WeWALK u can also get clsoser with public transport, it helps to navigate to public transport platforms, help user to check timetable. Besides,.it also help user to discover unfamiliar spaces by returning full list of nearby shops to user and help user to identify the shops with automatic voice feedback (WeWALK, n.d.)

Figure 2.16 WeWalk Smart Cane

2.2.1.3.3 CAN Go

The smart can that aim to keeps user in safe, active and connected (CAN, n.d.).There are multiple features provided by CAN Go that benefit visual impaired person.

i. GPS

   CAN Go integrate GPS into their smart can which able to let the others to track the user's location if he is missing, besides, when the smart cane lost, the GPS also can lead the others to find it accurately.

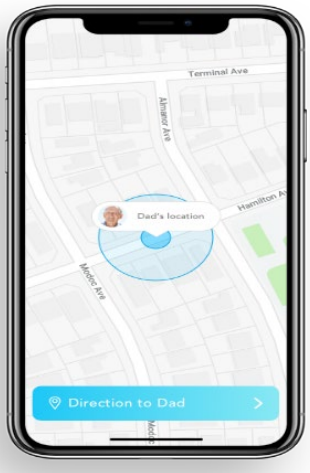


Figure 2.17 CAN Go- GPS Tracking Interface

ii. Cellular voice calling

CAN Go has built-in phone for emergency calling. The embedded speakers and microphone is high quality and it also integrated reliable and stable LTE technology. Wherever and whenever the user want to make a call for help, only a button to click for call. With this features, the user can make calls without the phone.

Figure 2.18 CAN Go Smart Cane

2.2.1.3.4 Summary

The summary of existing assistive technology shows that, there are a lot of technology ready to help visual impaired individuals but most of them is very pricy where majority of the visual impaired individuals couldn't afford it. Besides, there are diversity of technologies which requried external hardware to accept the assistance of the technology.

Table 1 Summary of Existing Assistive Technology

Feature	**Travel Aid for Visually Impaired**	**The vOICe**	**Finger Reader**	**Braille Note Takers**	**Braille Printer**	**Refreshable Braille Display**	**SmartCane Device**	**WeWalk Smart Cane**	**CAN Go**
Text Recognition	✓	NA	✓	NA	NA	✓	NA	NA	NA
Obstacle Detection	✓	NA	NA	NA	NA	NA	✓	✓	✓
Navigation Assistance	✓	NA	NA	NA	NA	NA	NA	✓	NA
Integration with Smart Devices	NA	Android-based Smart glasses	NA	NA	NA	Computers,phones and etc.	NA	touchpad	NA
Portability	✓	NA	✓	✓	NA	✓	✓	✓	✓
Feedback Mechanism	Voice	Converts image to sound	Sound	Sound	Printed Braille Paper	Braille	Sound/Vibration	Sound/Voice Feedback	Sound
Cost	Affordable	Affordable	Affordable	Pricy	Pricy	Pricy	Pricy	Pricy	Pricy
Ease of Use	Moderate	Moderate	Moderate	Easy	Moderate	Moderate	Moderate	Moderate	Moderate
GPS Tracking	NA	NA	NA	NA	NA	NA	NA	✓	✓
Cellular Voice Calling	NA	NA	NA	NA	NA	NA	NA	NA	✓

2.2.2 SDLC

Software development lifecycle (SDLC) is process that to make sure the development teams was cost-effective and time-efficient use to design and build high quality software. Therefore, an appropriate SDLC could help to make sure the software meets customer expectations and needs by minimizing the risk at the mean time. The SDLC is important because it helps every stakeholder to understand the development process, and the well-planned, estimation and schedule could better arrange the project. Besides,SDLC help to manage the risk and cost in the project and to make sure the quality of the deliverables which is systematic could better satisfy the customer expectations (AWS, n.d.).

2.2.2.1 V-Shaped Model

V-Model is an extension of Waterfall Model which it conducts testing during the end of each stage of development to make sure that there is no error (Alexandra, 2024), therefore the possibility process will stop due to unsolvable problem encounter which cannot proceed to next stages unless the problems solve, and this model also known as verification or validation model. (GeeksforGeeks, 2024)

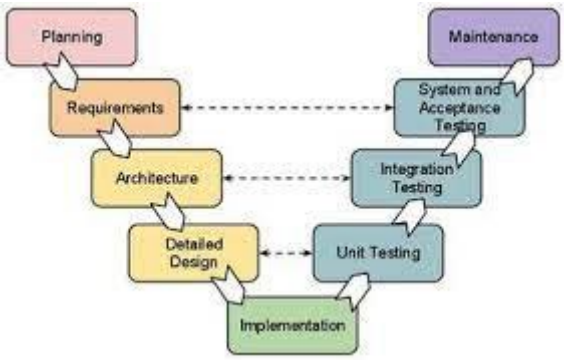


Figure 2.19 V-Shaped Model

2.2.2.2 Iterative Model

The iterative model suggests that to develop the software subset which required and enhance the versions iteratively over time until the complete software was developed, new versions of the software will produce at the end of each iteration and until it was ready to deploy to production (AWS, n.d.).

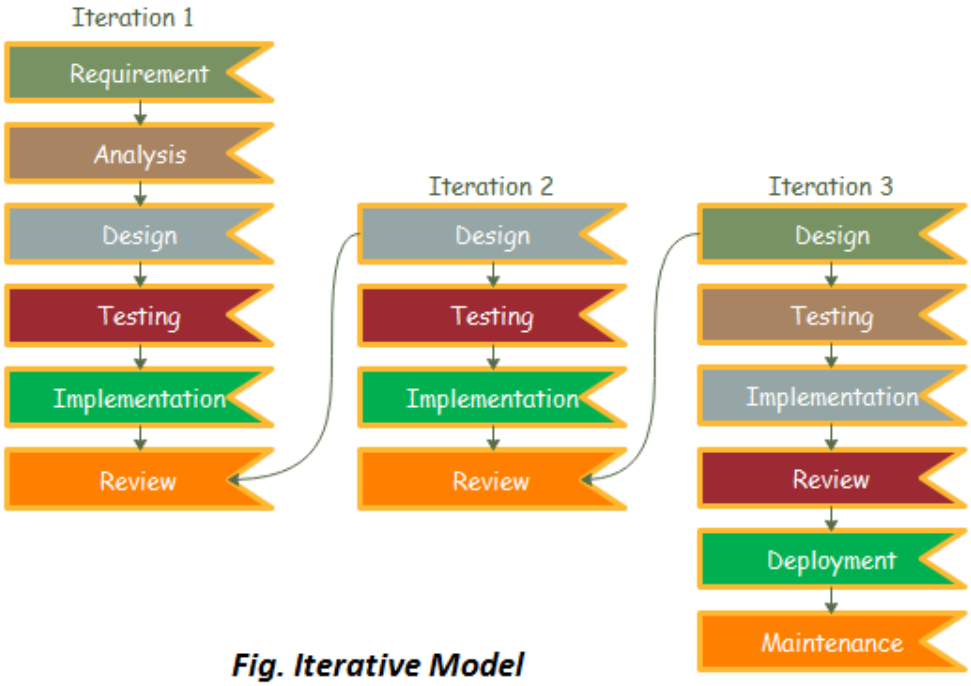


Figure 2.20 Iterative Model

2.2.2.3 Agile Model

The agile model will divide the lifecycle into few development cycles which the team aims to iterates through each phase and with different small deliverables by applying changes in each different cycle. The requirement will evaluate and be responsive to the result by applying changes to it, agile model is iterative and incremental which make it more efficient (AWS, n.d.)

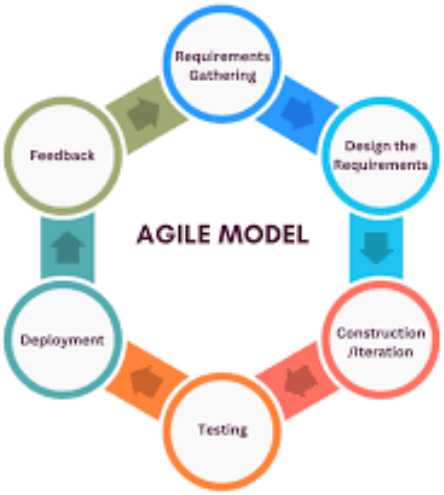


Figure 2.21 Agile Model

2.2.2.4 Waterfall Model

Waterfall Model will arrange the phases sequentially and it make sure every phase starting components will be based on last phases outcome.Therefore,the process will force to stop if last phases couldn't be complete on-time and the whole process will be delay (Alexandra, 2024).

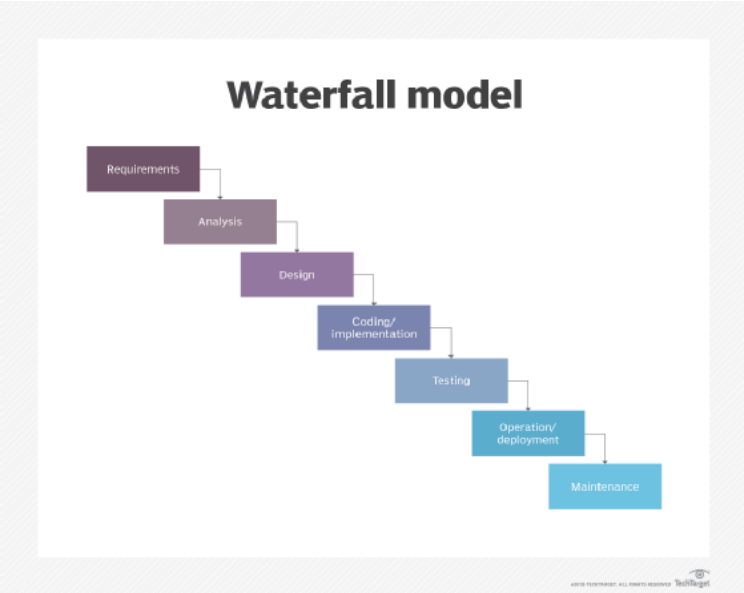


Figure 2.22 Waterfall Model

2.2.2.5 Spiral Model

The spiral model combines both waterfall model and iterative model which will prioritize risk analysis in the whole lifecycle. Development team obey spiral model to ensure software's gradual release and improvement by prototyping the software in different phases. (Alexandra, 2024)

Figure 2.23 Spiral Model

2.2.2.6 Summary of SDLC

Table 2 Summary of SDLC

Aspect / SDLC	Advantages	Disadvantages
Waterfall Model	- Simple and easy to understand. - Clear project milestones. - Well-suited for stable requirements. - Documentation-focused.	- Limited flexibility. - High risk of customer dissatisfaction. - Late detection of defects. - Long delivery time.
Agile Model	- Encourages collaboration between development teams, testers, and stakeholders. - Allows for rapid adaptation to changing requirements or safety standards. - Facilitates early and continuous delivery of valuable AsSee system features.	- Requires active involvement and commitment from all team members. - May be challenging to maintain documentation. - Can be less suitable for projects with strict regulatory or compliance requirements.
Iterative Model	- Allows for rapid prototyping and refinement of AsSee system features. - Enables quick feedback and adaptation to change requirements. - Facilitates early validation of AsSee system functionality.	- May require additional effort to manage multiple iterations simultaneously. - Increased complexity due to frequent iterations and changes. - Potential for scope creep if not managed effectively.
V-Shaped Model	- Ensures thorough testing and validation of AsSee system features. - Clearly defines testing requirements for each development phase. - Helps in identifying defects early in the development process.	- Can be rigid and inflexible, making it challenging to accommodate changes. - Testing phases may become time-consuming and costly. - Limited adaptability to evolving requirements or safety standards.

Spiral Model	- Allows for early identification and mitigation of safety risks. - Enables iterative development and refinement of AsSee system features. - Provides flexibility to accommodate changes in requirements or safety standards.	- Can be resource-intensive due to frequent prototyping and risk analysis. - May require extensive documentation and management of project risks. - Potential for project delays if risk analysis is not conducted effectively.

After reviewing the SDLC, the SDLC chosen for this project is Agile Model. Because the prototyping methodology under the Agile Model will be used for the development of the AsSee system due to its outcome will fulfil the end user and stakeholders satisfaction, the assistive technology which should be acceptable by the end user due to they are the one who accepting the assistance and use the technology, therefore their involvement in the development process is very important. Besides, with the process of the prototyping by refining the prototype and review the feedback which could improve the technology to ensure it fulfil and can really benefit the end user.

2.2.3 Development Framework

2.2.3.1 Programming

Programming enables developer to develop and realize ideas, there are different strengths and weaknesses for each different programming language, therefore select an appropriate programming language to develop the project is important and it will affect the project in many aspects like performance, speed, availability and etc.

2.2.3.1.1Java

Figure 2.24 Java

Java is an object-oriented programming language that runs on billions of devices. Java syntax and rules are based on C and C++ languages (IBM, n.d.). Java normally used for web development, mobile applications development, enterprise application development, Internet of Things (IoT), game development, AI and Machine Learning (ML).

The Java is widely used due to its independent, which Java code was allowed to runs on different devices which with different operating systems where does not required any modifications to make it adapt to different envinroment. Besides, it is an object-oriented programming language where it was able to handle complex project and with its own objects which contain data and methods. Furthermore, the robustness of Java language also an advantage where it can handle errors and exceptions and with its own mechanism, it can manage its own memore to reduce the occurrence of crash.

In addition, the security that provided by Java which the encryption, authentication and access control and the ability of the Java to handle complex and large project. Last but not least, the compability which enable the older version code still able to run on newer platform. Therefore Java is a language that could not be denied when considering to develop an Android application due to its flexibility, and wish to integrate AI into it. (Shittu Olumide, 2023)

2.2.3.1.2 Kotlin

Figure 2.25 Kotlin

Kotlin which an object-oriented programming language that interoperable with Java virtual Machine (JVM). It aim to improve the java programming with its more simple syntax and rules. Kotlin not only work as Android Programming Language but like Java which provides developers with different applications develop functions.

Other that Android development, Kotlin commonly used for Server-side development where to help the Java on web applications development where Java normally used for back end where Kotlin develop the server. Besides, the Kotlin also supports for full-stack web development where ability of Kotlin for Javascript to translate it into Javascript code for front-end code and enables to applied both of them in both back ends and front ends. Furthermore, the ability to develop for different OS where Linux, Android and iOS also an advantage. Furthermore, Kotlin also used for data science purpose like building data pipelines, and deploy machine learning models into production.

The advantages of the Kotlin is the interoperability where it interoperates with Java and can be compiled into Javascript. Besides, the safety also an advantage where it helps to avoid common coding erros. Furthermore, the clarity of Kotlin where it helps to minimize the complexity of syntax and helps developer to produce more concise code which can improve their productivity and save their time. In addition, the Kotlin support from Adnroid like Android Studio, Android KTX and Android SDK.The Kotlin selected by developers normally due to it is more simpler syntax and reduce the length of code with the support from Java and Android (Ben Lutkevich, n.d.)

2.2.3.1.3 Python

Figure 2.26 Python

Python commonly used to develop project due to itssimplicity and readability which helps developer for quick development which to keep competitiveness in the iterations and fast-paced automotive industry(Chris Raroque, n.d.). Besides, the libraries and the framework provided by Python was effective for data analysis which project required system to process huge amounts of data in short time which project need to be responsive and effective and the machine learning library for recognize the symbol and object which these tools could help the development of the project for visual impaired individual.

The Python extensive libraries could effectively help the development of the project and the machine learning ability like object recognition and decision-making which also important for the function of project like obstacles recognition and action taken for certain situation.

2.2.3.1.4Flutter

Figure 2.27 Flutter

Flutter is an open source framework that used to develop mobile applications and desktop applications. The flutter is a cross-platform language where operateable for different OS with only one codebase. Besides, the performance close-to-native where it use Dart programming language where to ensure the performance of the code and its effectiveness. Futheremore, the unique rendering technique that used by Flutter which is Google's open source UI rendering library, Skia graphic library which provide fast, consistent and customizable rendering which enables to render the code in different platforms. In addition, the tools which built-in by Google like hot reload that enable to check the change in code but without reloading and with real-time to investigate the changes and widget inspector for UI configuration (Appify, n.d.).

The null safety provided by Dart Language which mainly used by Flutter helps to handle null errors that reduce maintenance time taken. Besides, the Flutter is a cross-platform language which with lesser time consume and cost but lost access to native device functionality(AWS, n.d.).

2.2.3.1.5 React Native

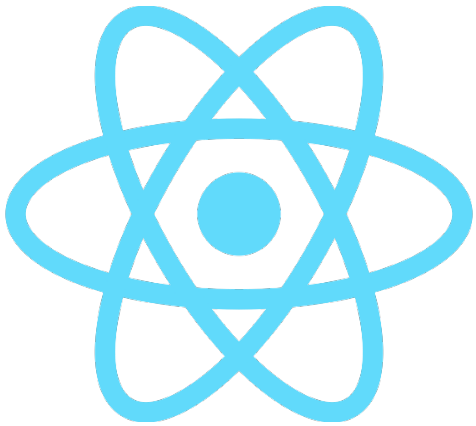


Figure 2.28 React Native

React Native is a development framework that enables developers to develop mobile applications which use JavaScript. React Native supports hot reload which its does not require recompiling to view changes made to the application where it had already enhance the productivity of the development process. Besides, with minor changes, the code was able to run on both Android and Ios platform where save a lot of time and cost and doesn't required to have two set of codes that running on different platforms.

React Native helps to develop complex mobile applications by simplify the complexity of algorithms. Besides, the library supports React Native make it more complete and more features that can be done. Therefore, with all the advantages had make the React Native popular and widely used to develop cross-platform applications and with satisfy output.(Bhagavatiprasad Vaghela, 2023)

2.2.3.1.6Summary of Programming

Table 3 Summary of Programming

Aspect / Programming Language	Advantages	Disadvantages
Java	-Platform independence -Object-Oriented Programming -Robust error handling -Strong security -Compability	-Complex Syntax -Slower Performance -JVM Required
Kotlin	-Interoperability with Java -Simpler Java Syntax -Safety to prevent null error -Readability -Support from Android Ecosystem	-Limited adoption outside of Android development -Smaller community
Python	-Simplicity and readability -Extensive Libraries -Effective for machine learning	-Slower execution speed
Flutter	-One Code Base for multiple platforms -Close-to-native performance -Fast and customizable UI rendering -Built-in tools like hot reload and widget inspector	-Lack of access to native device function -Limited community support and resources
React Native	-Cross-platform development with JavaScript -Hot reload for faster development -Simplified algorithms	-Limited access to native devices features -Security Matters -Device related issues

2.2.4 Existing Mobile Application

2.2.4.1 LookOut

Figure 2.29 Lookout

LookOut by Google is an application that helps visually impaired individuals with help them to read documents and identify objects(Abrar Al-Heeti, 2023). “There are so much visual media thas’s pervasive, and it’s so often inaccessible (Adams, n.d)” where Lookout Product Manager Scott Adams tries to make the LookOut a tools that could assist visual impaired individuals to interact with the visual world which is pervasive.

There are different features that provided by LookOut:

I. Text: Helps user to read text
II. Explore (Beta): Help user to gather information of the environment.
III. Food Labels: Help user to recognize the food by scanning barcode or front view of the food.
IV. Documents: Help user to read whole pages of document
V. Currency: Help user to identify currency value
VI. Images (Beta): Help user to identify uploaded image and return description of the image to the user.

Therefore with the computer vision provided by LookOut, the user able to gain information and understand object or environment situations with only the involvement of camera or any extra sensors on Android devices. (Google, n.d.)

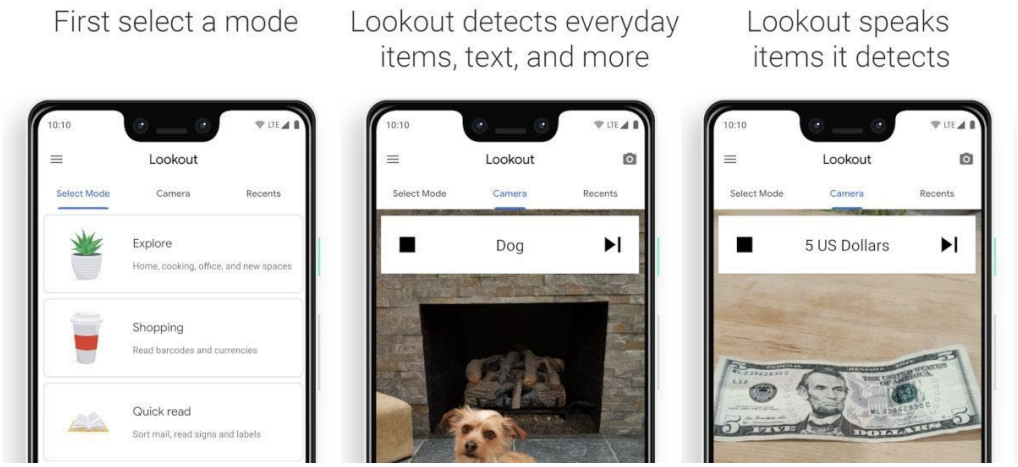


Figure 2.30 Lookout Interface

Kies uit 6 verschillende standen, afhankelijk van wat je doet.

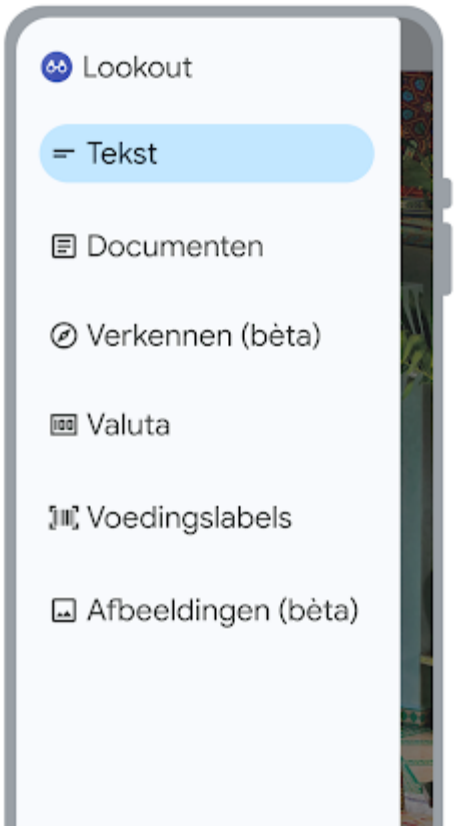


Figure 2.31 Lookout Sidebar

As we can see, there are a lot of interface that required user to interact with, therefore sometimes it may cause the user takes time to interact with the functionality that user want. Besides, the interface was bit small where it may

hard to interact with it. Furthermore, there are settings that required user to configure, therefore it may be not user friendly for those who are not familiar with technology devices.

2.2.4.2 TapTapSee

Figure 2.32 TapTapSee

TapTapSee is an application that designed for visual impaired individuals which it provide general identification of any image or video to assist user to understand and access information in the materials.

There are two main modes in TapTapSee where u tap twice on the right side of the screen, it will take a picture and recognize it, but when u double-tap on left side of the screen, it will start recording video which up to 10 seconds, to stop the video recording under 10 seconds u need to re-tap twice on left side of the screen again.

The TapTapSee offered different features which could help user to maximize the usage of the application. The application supports user to repeat the last identification result if the user missed the first feedback. Besides, it allowed users to upload image and video from gallery to the application that user want the application to identify. Furtthermore, after the identification the user can share the result. In addition, the TapTapSee allowed users to store the result to the gallery.

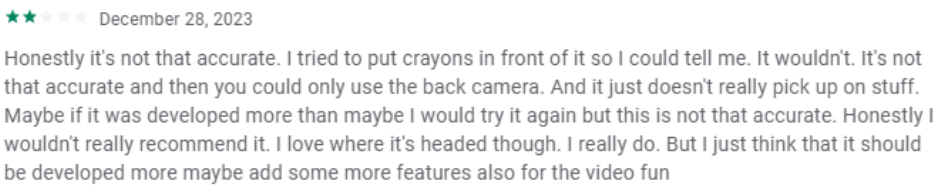


Figure 2.33 Rating for TapTapSee 1- Google Play

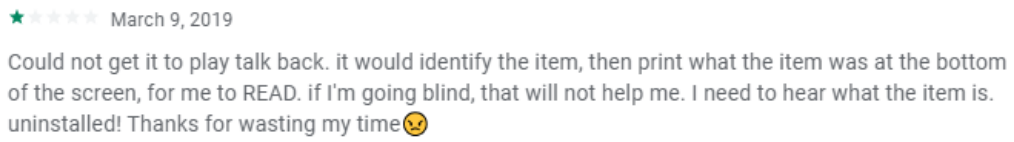


Figure 2.34 Rating for TapTapSee 2- Google Play

Based on the review from the exact users, the TapTapSee result was not satisfied due to its inaccurate result and it doesn't work with the TalkBack which is Android native assistive technology. Furthermore, there are users feedback that the required permission is more than necessary where TapTapSee required location and microphone permission where TapTapSee doesn't have features that require these two access.

In conclusion, TapTapSee is mobile application powered by CloudSight Image Recognition API where it benefits blind and visually impaired user with recognize image and video and convert to speech. (TapTapSee, n.d.)

2.2.4.3 Supersense

Figure 2.35 Supersense

Supersense is a computer vision for the blind where it integrated novel computer vision and machine learning models into it. There are different function that provided by Supersense which is convert the information of the environment and convert into hearable information for the user. Besides, the smart scan function which helps user to detect the letters, documents, books, currency, barcodes and etc with its automatically scan the object and convert to speech immediately. (Supersense, n.d.)

Other than only providing smart scan features, where the Smartsense also provides guidance to the user to make sure them how to place the camera to get better outcome. Furthermore, its allows scan multiple page in one scan and with object explorer features provided to helps user to identify

surrounding. Last but not least, it also allows users to upload pdf and image to identify text inside to convert to speech.

"With the power of AI, supersense automatically figures out what you are trying to scan, guides you on how to point the camera, and reads the content in the right format. Its unique design minimizes the time and frustrations of scanning and reading text for a blind and visually impaired." (Laura Medcalf, 2021). Which is the description written in AppStore where showed the Supersense main focus features, if the user want to access to full features that provided by Supersense, they need to subscribe to premium package when free version of Supersense enables quick read, import files function, recognition history and magnifier only. For all those, smart scanner, document scanning, multipage scanning, currecy scanning, barcode scanning, explore environment and find object requried user to pay for the services at monthly fees 4.99 USD.

2.2.4.4 Intelligent Eye: A Mobile Application for Assisting Blind People

Intelligent Eye is an android mobile application which founded by Milios Awad, Jad El Haddad, Edgar Khneisser, Tarek Mahmoud, Elias Yaacoub, Mohammad Mali in 2018 which aimed to provide few function for the blind people which is banknote recognition, color detection, light detection, and object recognition. Here is the description of the Intelligent Eye :

I. Light Detection : Make use of the light sensor on native android devices to detect the light intensity and describe the light detected to the end user.
II. Color Detection : Its help user to identify color of the image, the colour identified will convert into speech and feedback to user.
III. Object Recognition : Helps user to recognize object without internet connection with associated with database objects with guaranteed response speed.
IV. Banknotes Recognition: Identify the banknotes with associated with database to make it function without internet connection and with response speed guaranteed.

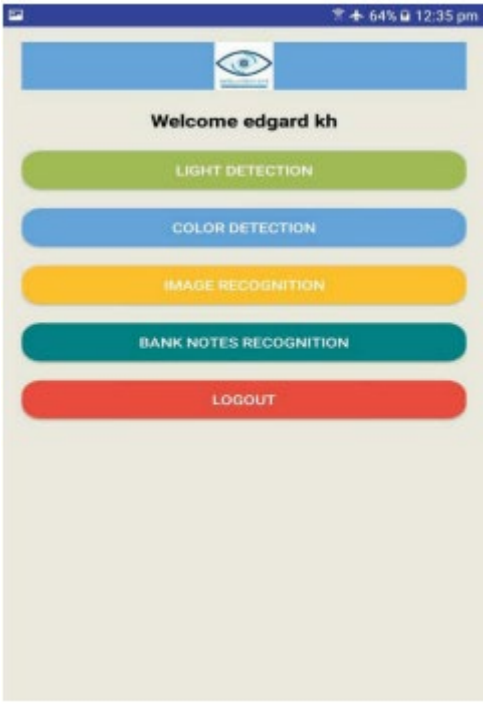


Figure 2.36 Intelligient Eye Interface

This is the interface of the Intelligient Eye, where the interface was not user-friendly for those visually impaired individuals due to small button and a lot of others information like the logo, welcome text and logout button that may cause longer time needed for the user to get familiar with the system and to use function that user intended.

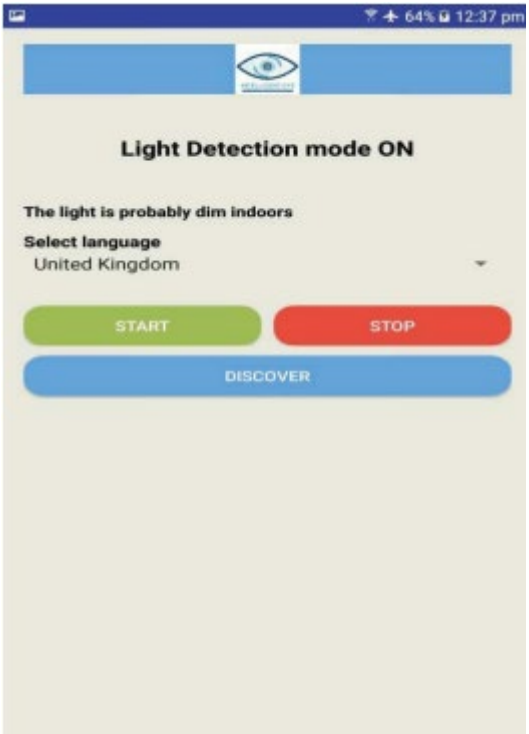


Figure 2.37 Intelligient Eye - Light Detection Interface

This is the light detection mode, where the interface was also not user-friendly where it required input like language, and there are 3 different button in one pages which may lead to confusion of user.

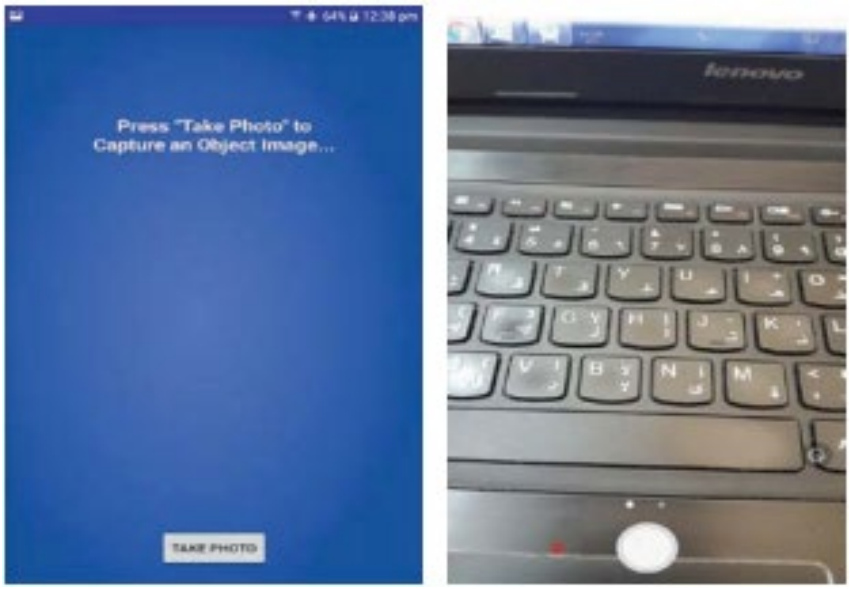


Figure 2.38 Intelligient Eye- Object Detection Interface 1

Figure 2.39 Intelligient Eye- Object Detection Interface 2

This is interface of the object detection, the take photo button is way too small for blind people where it is very difficult for them to locate the button quickly and precisely. Besides, the results showed that the result accuracy also not precise and accurate.

In conclusion, Intelligent Eye aimed to provide different functionality to users with the user-friendly and accurate result feedback to user. (Khneisser et al., 2018). But after reviewing, the project may not friendly to blind person

where the design of the system was too complex and required a lot of interaction to interact with the system and the result was not completely accurate and sometimes couldn't provide precise feedback to user.

2.2.4.5 Slightless Helper: An Interactive Mobile Application for Blind Assistance and Safe Navigation

Slightless Helper which is an application that designed to assist blind person. This system uses footstep counting and GPS for indoor and outdoor navigation. There are different sections that provided by Slightless Helper:

I. Blind Assistant Section (BAS) : There are few modules under this sections aimed to provide assistance to movement and safety to user :
   a. Contact Module : Helps user to store contact numbers.
   b. Permission Module : Make sure application granted all permission needed.
   c. Pe2dometer Module: Use smart phones built-in accelerometer sensor to count the footsteps and store inside the application.
   d. Area Module : Store locations which is safe and unsafe, and will notify user when user entered unsafe area.

II. Blind People Section (BAS) : This section is to help user set command that wish the mobile application respond corresponding.
   a. Jolt Module : The mobile application will send location information to the contact number stored and call the stored numbers.
   b. Reqquired Module: This module will notify user about the battery level, date, time and news.
   c. Indication Module: This module will notify user when they entered unsafe area that preset.
   d. The speaking scanning module will actions corresponding to the command given by user.
   e. Tap Module that respond to the tap actions perform by user.

III. Indoor and Outdoor Navigation of the Proposed System : Its help user to navigate to location that user intended.

The Slightless Helper function where preset the footstep counting to preset location like toilet which it required same starting point else the stored footstep may confusing the user. Besides, the shaking module which is very easy to trigger even though user are not intended to do so. In addition, there are a lot of input required user to input where like footstep destination name, store phone number, preset area safety level and command to react with, therefore it required a lot of interacting with the mobile application where it will be a challenge to end user.

In conclusion, it is a mobile application that could benefits user in certain situations but required plenty of configuration to fully ultilize the applications..With all the functionality provided by Slightless Helper where in different way to interact with but could improve the quality of daily life of end users and it is very reliable and usable in end user daily rountines. (Hossain et al., 2020)

2.2.4.6 Object Recognition mobile app for visually impaired user

This project aimed to produce an Android-based mobile application that helps user to gain information of the surrounding by detects and recognize the objects in real-time. Besides, it helps user to identify the objects location based on the camera location.

Figure 2.40 Object Recognition mobile app for visually impaired user Interface

This is the mobile application that is the outcome of the project, where it is simple and easy to use.With the adjustable speech rate, where can help user to modify the talking speed to make the user can set preferable talking speed. The object recognition function for the application with notify user through the three section divided,

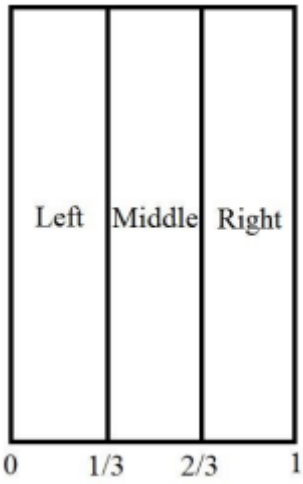


Figure 2.41 Object Recognition mobile app for visually impaired user - Three Section

like “Fan on the right ”.With the introduction of this project, the daily task of visual impaired individual could be improve and secured by notification from mobile application to let user access to information surrounding and location of the object.(WON WEI CHENG, 2021)

2.2.4.7 Summary of Existing Mobile Applications

Table 4 Sumamry of Existing Mobile Applications

Application Name	Features	Advantages	Disadvantages
Lookout by Google	- Text reading -Environment exploration -Food label recognition -Document reading -Currency identification -Image description	-Wide range of features -Integration with Android devices	-Complex interface -Small interface elements -Configuration settings can be overwhelming for some users
TapTapSee	-Image and video recognition -Repeat last identification -Upload from gallery - Share results - Save results to gallery	-Simple interface -Various identification modes	-Inaccurate results reported by some users - Incompatibility with Android's TalkBack
Supersense	- Smart scan -Object exploration -Document and image upload -Guided scanning	-AI-powered features -Guidance for better scanning	-Premium features require subscription -Limited features in free version
Intelligent Eye	- Light detection - Color detection -Object recognition -Banknote recognition	-Offline functionality -Various detection modes	-Complex interface - Small buttons -Inaccurate results reported
Slightless Helper	-Footstep counting - GPS navigation -Command-based interaction -Contact management	-Navigation features -Customizable commands	-Requires extensive configuration -Shaking module may trigger unintentionally
Object Recognition App	-Real-time object recognition -Adjustable speech rate	-Simple and easy to use -Adjustable speech rate	-Limited functionality compared to others
AsSee application	-Real-Time Object Detection -Real-Time Text Recognition -Command-based Interaction	- AI powered - Simple and ease to use - Real-time capabilities -Voice command interaction	- Limited functionality

The summary of the existing assistive mobile application for visually impaired which concluded a lot of different features and functionality provided. But mostly the application user interface designed was not user-friendly enough for the end user to interact with, in the sense of user-friendly for visually impaired, current assistive software there are too much interface and interaction required to perform some actions and configuration for the application. For example, require user to take photo before detecting object around the user and require photo taking action from user to perform recognize text action. Besides, the accuracy of the data also important concerns to deal with. In addition,there are too many configurations that required user to setup which may exhaust the user and it is challenging to them who are visually impaired. Therefore, it is glad to have these mobile applications ready to assist the user, but there are improvements that could be made to make the user benefits more and more convenient to interact with.

In conclusion, the AsSee application was designed to address some of the shortcomings of other apps while incorporating their advantages, resulting in a more effective tool to assist visually impaired individuals in daily life. With the real-time object detection and text-recognition functionality which doesn’t require too complex action from the user to perform such action, it does provide user in better way and easier way to get assistance. Additionally, the integration of multiple interaction methods simplifies communication with the app, giving users alternative ways to interact and choose how they prefer to engage with the application.

2.3 Summary

In conclusion, with the selected methodology after conducted literature review, the quality of the system could be improved and to make sure that the most suitable platform or methodology was chosen to develop the system. Besides, with the appropriate methodology chosen, it can make the development process smoother and faster. With all the insight and information gathered, the most suitable and appropriate methodology will helps in developing the AsSee System and provide a guideline for identifying problem statement and objectives.

CHAPTER 3

METHODOLOGY AND WORK PLAN

3.1 Introduction

From Table , based on the SDLC analysis, the Agile Model also practices iteravite development which for every upgrade will have to develop step-by-step. For each outcome which designed to small part and expected to complete within few weeks. For each iteration will only focus on one part to plan, develop and deploy to customer. Besides, the long-term plan are not made for Agile Model.

3.1.1 Agile Model

Agile Model is SDLC that breaks task into few parts which aim to complete it step by step which aim to complete it in few weeks and doesn't involve long-term planning. In Agile Model, it every iteration also known as frame which aim to complete the frame within one to four weeks. The purpose of Agile Model to break tasks into small part is to aim to minimize the risk that may arise in the project and the project delivery time required. (Javatpoint, n.d.)

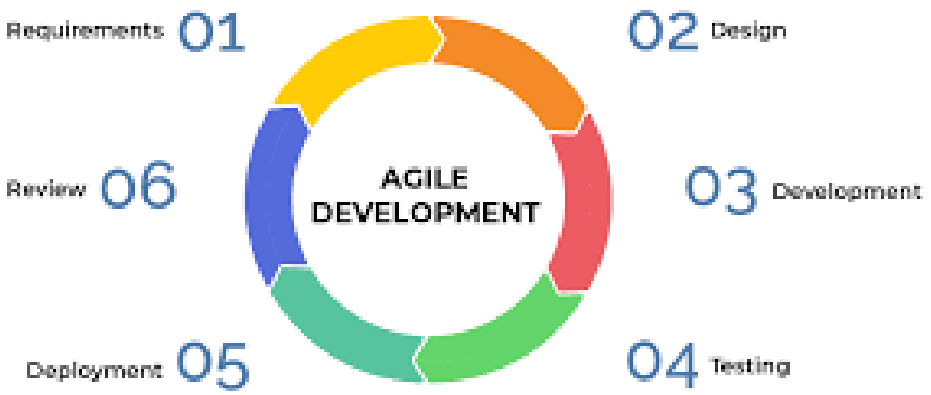


Figure 3.1 Agile Model Phases

3.1.1.1 Requirements Gathering

In requirements gathering phase, the requirements should be well-defined. By identifying requirements, the business opportunities can be identified and the estimation time, effort and resources needed for the project can well-planned during this phase. With all these information, the evaluation of the technical and economic feasibility.

3.1.1.2 Design the Requirements

After the requirements defined after the discussion with the stakeholders, the requirements could be visualized by UML diagram and user flow diagram to make sure the requirements understandable by every stakeholders, and how the system will interact with users and how to integrate new features onto it.

3.1.1.3 Construction / Iteration

After the visualization and design of the requirements, development will start and aim to deploy the product. The development will aim to develop the small part of the project which is simple and handle minor functionality.

3.1.1.4 Testing / Quality Assusrance

In this Testing phase, the outcome from previous phase will be examine the product performance and looking for the bugs and error mislook during development process.

3.1.1.5 Deployment

After the testing was completed, the product was deploy into production which let user to use and help user to complete specific and expected task.

3.1.1.6 Feedback

After the deployment, and the team will ready to receive feedback from the user of the product.

3.1.2 Agile Model Testing Methods

3.1.2.1 SCRUM

SCRUM is one of the development process that focuses on way to manage tasks in team-based development.

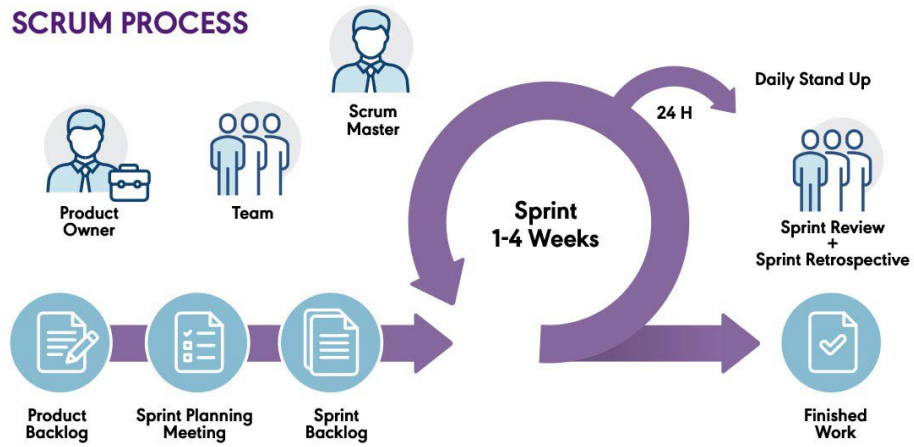


Figure 3.2 SCRUM PROCESS

There are three roles in SCRUM which is Scrum Master, Scrum Team and Product Owner. The master will set up the team, schedule and arrange meeting within the team, and remove the obstacles encounter during the process. Besides, the product owner will make the product backlog, prioritizes the delay and determine the function of the each iteration of project. Furthermore, scrum team will work for complete the sprint or cycle that scheduled and planned.

3.1.2.2 Crystal

There are three concept in Crystal, which is Chartering, Cyclic Delivery and Wrap Up.

3.1.2.2.1 Chartering

The Chareting will mainly focus on the set up development team, perform feasibility analysis and developing plans, etc.

3.1.2.2.2 Cyclic Delivery

There are two cycles consist which is the team update the release plan and integrated product delivers to the users.

3.1.2.2.3 Wrap Up

The wrap up will perform deployment, post-deployment based on the user environment

3.1.2.3 Dynamic Software Development Method(DSDM)

Dynamic Software Development is rapid application development strategy which requires the users must be connected to the project and given rights to make decisions (GeeksforGeeks, 2019). There are three techniques used in DSDM:

- Time Boxing: This technique will allocate certain time period which known as times boxes and assign to specific taks and activities inside the project. It will help the project to stay on schedule.
- MoSCow Rules: This technique will categorize requirements into four different groups which is Must have, Should have, Could have, Won't have. It help to prioritize the requirements and determine the essential features of the project and the handle the expectation of the stakeholder.
- Prototyping: This technique will create simple version of the project and gather feedback, and follow up the requirements, and identify risk and solve risk. But visualize the project in simple way will help stakeholders to determine the improvements throughout the project development process.

Besides, there are seven stages to focus during DSDM which is

- Pre-Project: During Pre-Project steps, the needs will be identified, setup the development team,funding and defined the project objectives and scope.
- Feasibility Study: In this stage, the examination of the technical, financial, organizational aspects will determine the feasibility of the project to make sure this project is worth to proceed.
- Business Study: Business Study phase will make sure the stakeholder understand the business needs, objectives to achieve, and constraints of

businesss. It will analyse the stakeholder requirements, confirm the flow of the businesses and estimate the outcome of the project.

- Functional Model Iteration: The development team will develop the project functional prototype step-by-step and collect the feedback of the stakeholders about the functional prototype which aim to use these feedbacks to improve the product's functionality and determine the requirement more precisely.
- Design and Build Iteration: The development team will design, develop and test the product iteratively and make sure the the stakeholders involvement to make sure the product and outcome fulfil the customer needs and expectations and can make improvements based on the feedback from the stakeholders.
- Implementation: The final product will deployed and operate which involve training to the user, deploy the product to production and ensure the outcome fulfil the performance expected and quality standard during this stage.
- Post-Project: This stage will focus on reviewing the outcomes and identify the lessons that learned throughout the process and always updating the product to always meet stakeholders needs and expectations which development will aim to maintenance of the product and support the user needs.

3.1.2.4 Feature Driven Development(FDD)

This method will focus on Designing and Building features, the features will be break into more details work which for each steps within the development process will need to be obtained separately. (Javatpoint, n.d.)

3.1.2.5 Lean Software Development

Lean Software Development obeying the just in time production,which aim to increase the development speed of the product and reducing the cost at the mean time (GeeksforGeeks, 2023). There are seven phases in Lean Software Development (Javatpoint, n.d.)

- Eliminating Waste: This phases will reduce and eliminate the process, feature or activities that doesn't benefit the customer but consume a lot of resources.
- Amplifying learning: Emphasize the importance the continuous improvement in the development process. Besides, to make sure the improvement fulfil the user needs, the team should always seek for feedback from user and absorb the experience and lesson from those experiments.
- Defer Commitment: Delay the decisions as late as possible until as many as information gathered for better decision making. To avoid any decision made with only unclear information.
- Early Delivery: Delivery important functionality to customer as early as development team can, with the earlier delivery, teams can make improvements based on the feedback, and be more responsive to the requirements.
- Empowering the team : Empower the development team to make sure they are motivated, innovative and it able to help the team to have better quality of outcome, besides the team should have their right to manipulate their work and make decision in certain situations.
- Building Integrity: To make sure the functional requirements and non-functional requirements was met and obeying the regulatory and ethics.
- Optimize the whole: Aim to optimize the entire value stream but not only focus on small parts of the process or project.. This stage will focus on the impact towards the entire system.

3.1.2.6 eXtreme Programming(XP)

The eXtreme Programming is used the requirements is not clearly confirmed by the stakeholders and regularly changing demands and requirements and also when stakeholders doesn't sure the performance expected by them. (Javatpoint, n.d.)

3.1.3 Advantages and Disadvantages

Table 5 Advantages and Disadvantages of Agile Model

Advantages	**Disadvantages**
Reduce time taken on project development	Lack of formal documentation will lead to confusion and team members will interpret those important decision made during each of different phases.
Emphasize connection between team members which development team will have deep collaboration and understanding.	Not suitable for project that is complex.
Adapt to any changes requested by stakeholder.	Customer with unclear information and unclear expectations may lead the development team into wrong direction.
Prioritize customer needs and make sure the outcome meets their expectation and requirements.	Development team member must be expertise since the requirements of the project may change regularly and rapidly therefore the development team should able adapt to changes.
Development team work together to deliver the product which is better than programmer working alone.	With no proper documentation, maintenance and support could trouble the development team member.

3.1.4 Summary

In this project, the selected method and technique is prototyping methodology. Which this technique will create simple version of the project and gather feedback, and follow up the requirements, and identify risk and solve risk. But visualize the project in simple way will help stakeholders to determine the improvements throughout the project development process. There are few necessary steps needed to taken for produce a better outcome:

3.1.4.1 Requirements Gathering and Analysis

This is the starting point of the prototyping model. In this steps, there are several steps taken to gather the requirement and analyze those requirements. These requirements help stakeholders to have better understanding on the project and clearer information about the project. The method that used to

collect the requirement and information is Literature Review, Questionnaire, and Observation.

3.1.4.1.1 Questionnaires

A questionnaire aim to collect 20-30 respondents and the objective of this questionnaire is to gather insights and perceptives of public regarding the development of this project. Besides, throught the conduct of this questionnaire, the the expectaction which can provide a better guide to the development of this project. This questionnaire collected is the public aware of existing of mobile application assistive technology for visually impaired individuals and what do the public think that how mobile application assistive technology to help the visually impaired individuals and how this technology could assist in their life and improve their quality of life. The important concerns in this questionnaire is to know what expection of functions or features that could really benefits the visually impaired individuals and how it benefits them.

3.1.4.1.2 Literature Review

The literature conduct to review the existing technology to assist visually impaired individuals. There are different technologies that existed to assist visually impaired and I believe these technologies do assist visually impaired in some ways. Besides, the aspects that I had reviewed is the SDLC, Programming Language, Embedded System which also important in this project.

With the literature review conducted, the SDLC, Programming Language and embedded system can be determined and make the development team and stakeholders have better insight of the project and is there any improvement or aspect need to focus to make the AsSee to really assistive to end users.

It is very important to have a appropriate programming language which it determine the performance of the project, speed of respond, library that can be use to make this project more complete, therefore with the programming language review, the most suitable and appropriate programming language determined. Besides, the SDLC is also an important

role in developing this project, it helps to determine which methodology applied to this project and ensure the outcomes satisfied the stakeholder and fulfil end users requirements. Furthermore, selecting an appropriate embedded system also important which it relates to its processing speed, how it configures to respond to certain situations, and the speed and accuracy of processing data and information gathered by components connected to it.

3.1.4.1.3 Observations

With the observation, it helps to identify what problems or challenges that encounters by the visually impaired individuals and we get to have clearer insight of their daily life and daily routines, and identify what assistance needed in certain situations. By inspecting into their life, it helps development team to have better understanding what they are facing and how does the situation affect them and come out with the functions or features that can assist them in these situations.Observation helps to identify the need of the end users precisely and understand their life and the problems encountered better.

3.1.4.2 Quick Design

This is the phase that design the requirements quickly so that an overview can be conducted to understand the requirements. This phase helps stakeholders to have better understanding and access to basic design to the system. Therefore with quick design, there are more clearer system and user requirements that can help to improve the following steps design and improvement on the whole system development process.

3.1.4.3 Build Prototype

Prototype helps stakeholders to have better understanding and access to how the system will looks like.

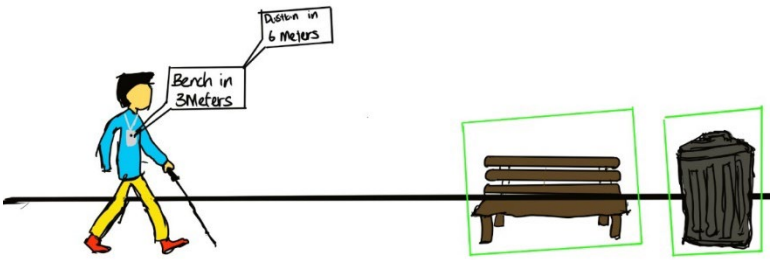


Figure 3.3 Prototype for AsSee - 1

This showed the AsSee system able to identify obstacles on the road and notify user for them to take corresponding actions.

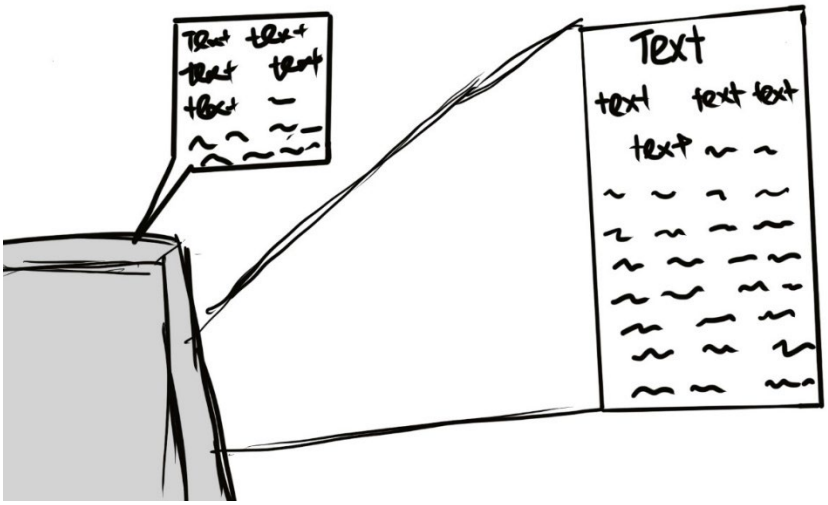


Figure 3.4 Prototype for AsSee -2

This showed the AsSee System able to identify text and translate it to speech for them to understand written text with no braille.

3.1.4.4 Initial User Evaluation

In initial user evaluation stage, the stakeholders will involve in the testing of this product, with the weakness and streghtness identified in this stages, there are improvements could be made to make sure the product was satisfied before it released. It will continue to improve the prototype, and review it until it is good enough to deploy to production. Besides, during this process, the system will improve iteratively. With those information collection in this phase, the

development team can make sure the user requirement are achieved and the outcome will satisfy the end user and stakeholders.

3.1.4.5 Refining Prototype

In this stage, with all the information and feedback collected, the development team will start to implement changes and improvement toward the AsSee system which to make sure the system make the user requirements, after the refining process, the system will be evaluate again and again after the prototype are good enough to deploy or be use without any other feedback given. This stage normally will repeat few times, to ensure the quality of the prototype and it is fulfil the requirement of the end user and stakeholders.

3.1.4.6 Product Implementation and Maintenance

This is the last stage which is the approved prototype was deployed to production, and the development team will start working on design,coding, and testing during production. The product after testing and deployment, will be periodically maintained ot ensure the the downtime and large-scale of failure minimized to increase the usability and reliability of the product.

The maintenance of the AsSee system is extremely important due to it is helping those who really need help, therefore the error rate and the maintenance are necessary to make sure the error rate minimizes and the accuracy and performance of system maximized (GeeksforGeeks, 2024a).

3.2 Project Plan

3.2.1 Work Breakdown Structure (WBS)

0.0 AsSee: Assitive Mobile Application for Visual Impaired Individual

1.0 Project Planning and Requirements Gathering

- 1.1 Draft Introduction
- 1.2 Draft Importance of Study
- 1.3 Identify Problem Statements
- 1.4 Identify Project Objectives
- 1.5 Propose Project Solutions
- 1.6 Propose Project Approach
- 1.7 Define Project Scope

1.7.1 Identify Target User

1.7.2 Define Main Features and Functions

1.8 Conduct Literature Review

1.8.1 Review On Existing Asssitive Technology

1.8.1.1 Review of Computer Vision Based Assistive Technology

1.8.1.1.1 Review Travel Aid of Visually Impaired

1.8.1.1.2 Review The vOICe

1.8.1.1.3 Review Finger Reader

1.8.1.2 Review Braille Technology

1.8.1.2.1 Braille Note Takers

1.8.1.2.2 Braille Printer

1.8.1.2.3 Refreshable Braille Display

1.8.1.3 Review Smart Cane

1.8.1.3.1 Review Smart Cane Device

1.8.1.3.2 Review WeWalk Smart Cane

1.8.1.3.3 Review CAN Go

1.8.1.4 Summarize Existing Assistive Technology

1.8.2 Review On SDLC

1.8.2.1 Review V-Shaped Model

1.8.2.2 Review Iterative Model

1.8.2.3 Review Agile Model

1.8.2.4 Review Waterfall Model

1.8.2.5 Review Spiral Model

1.8.2.6 Summary of SDLC

1.8.3 Review On Development Framework

1.8.3.1 Review Programming Language

1.8.3.1.1 Review Java

1.8.3.1.2 Review Kotlin

1.8.3.1.3 Review Flutter

1.8.3.1.4 Review React Native

1.8.3.1.5 Review Python

1.8.3.1.6 Summary of Programming Language

1.8.3.2 Review Existing Assistive Mobile Applications

1.8.3.2.1 Review Lookout by Google

1.8.3.2.2 Review TapTapSee

1.8.3.2.3 Review Supersense

1.8.3.2.4 Review Intelligent Eye

1.8.3.2.5 Review Slightly Helper

1.8.3.2.6 Review Object Recognition App

1.8.3.2.7 Sumamry of Existing Assitive Mobile Application

1.9 Summarize of Literature Review

1.10 Methodology and Work Plan

1.10.1 Determine methodology

1.10.1.1 Define and design phases of methodology

1.10.2 Define Work Breakdown Structure (WBS)

1.10.2.1 Identify work in project

1.10.3 Define Gantt Chart

1.10.3.1 Define project plan in details

1.10.4 Define Project Tools

1.10.5 Conduct analysis of questionnaire

1.10.5.1 Design Questionnaire

1.10.5.2 Deploy Questionnaire

1.10.5.3 Analyse Questionnaire Responds

1.10.6 Requirement Specifications

1.10.6.1 Identify Functional Requirements

1.10.6.2 Identify Non-Functional Requirements

1.10.7 Use Case Modelling

1.10.7.1 Use Case Diagram

1.10.7.2 Use Case Description

2.0 Data Acquisition

2.1 Identify and collect relevant dataset

2.1.1 Identify and collect relevant dataset for obstacles recognition training

2.1.2 Identify and collect relevant dataset for text recognition training.

2.1.3 Identify and collect relevant dataset for command recognition training.

2.2 Process Data

2.2.1 Process the data

2.2.1.1 Modify the dataset to make dataset applicable for training

2.2.1.1.1 Resizing Data

2.2.1.1.2 Crop Data

2.2.1.1.3 Normalization data

2.2.1.2 Augment dataset

2.2.1.2.1 Rotate Data

2.2.1.2.2 Zoom Data

2.2.1.2.3 Flip Data

2.2.2 Finalize Dataset

3.0 System Development

3.1 Setup System

3.1.1 Setup Development Environment

3.1.2 Download and import necessary libraries

3.2 Model Development

3.2.1 Implement TensorFlow Lite library for object, obstacles detection

3.2.2 Implement ML Kit for text-recognition

3.2.3 Implement Android TextToSpeech API for text to speech.

3.2.4 Develop Model

3.2.5 Training Model

3.2.5.1 Train Object Detection Model

3.2.5.2 Train Text Recognition Model

3.2.5.3 Train TTS Model

3.2.5.4 Train Speech Recognition Model

3.2.6 Improve Model

3.2.7 Develop method to ultilize these model in corresponding situations

3.3 Integration and Testing

3.3.1 Integrate Model trained into Mobile Application

3.3.2 Testing Model by simulating working environment

3.3.3 Improve System

3.3.3.1 Fix Bugs

3.3.3.2 Improve performance,usability and reliability

4.0 Documentation and Deployment

4.1 Document system design, implementation and testing results

4.2 Generate User Manual and Training Module for end user

4.3 Deploy System

4.4 Maintain system

5.0 Lesson Learned

5.1 Identify lesson learned in this project

5.2 Documentation of lesson learned

5.3 Identify advantages and disadvantages of this project

3.3 Gantt Chart

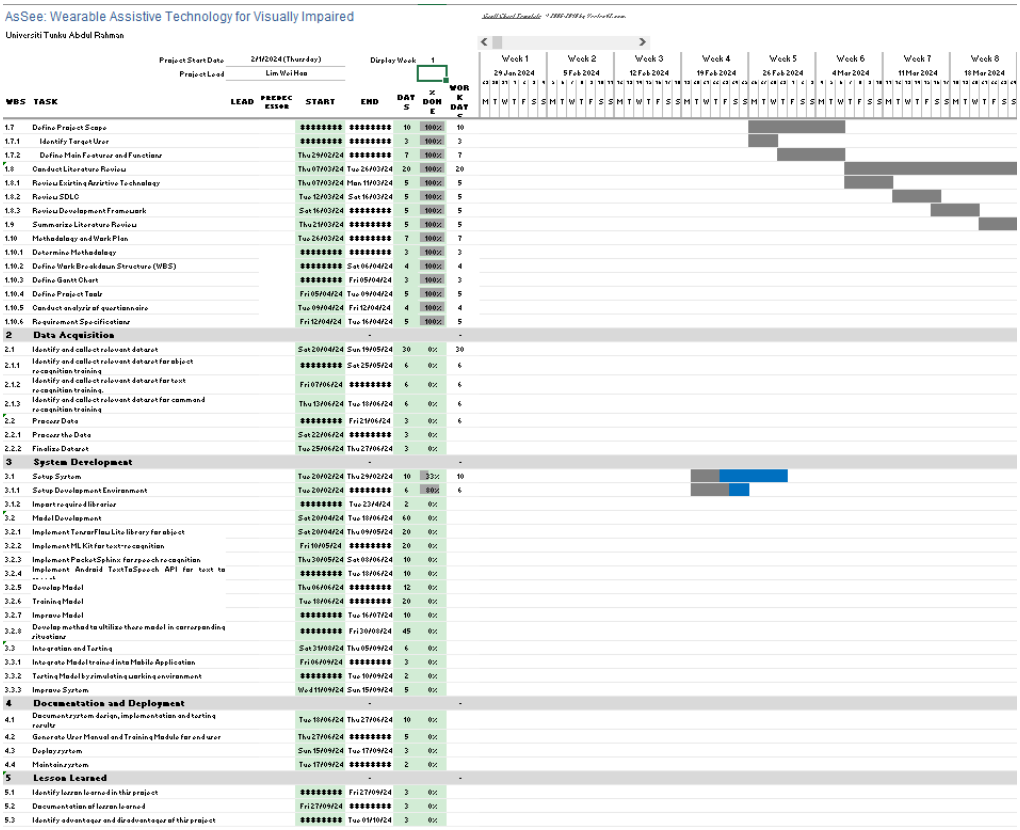


Figure 3.5 Gantt Chart 1

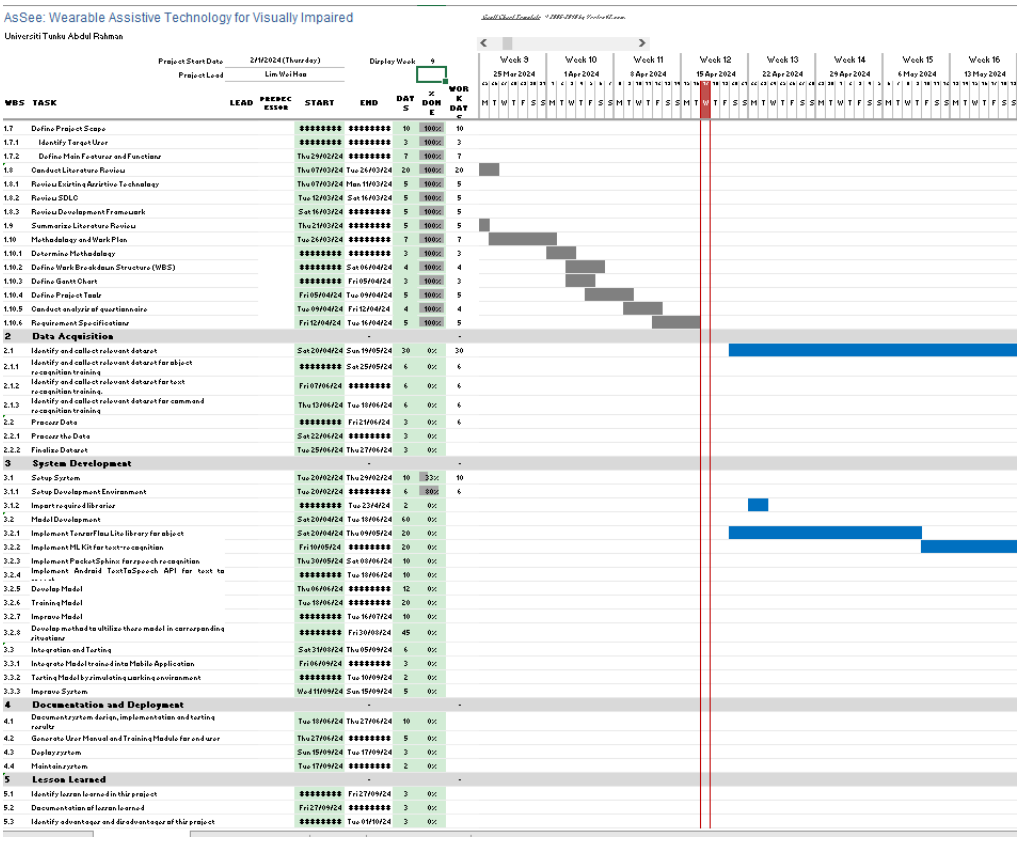


Figure 3.6 Gantt Chart 2

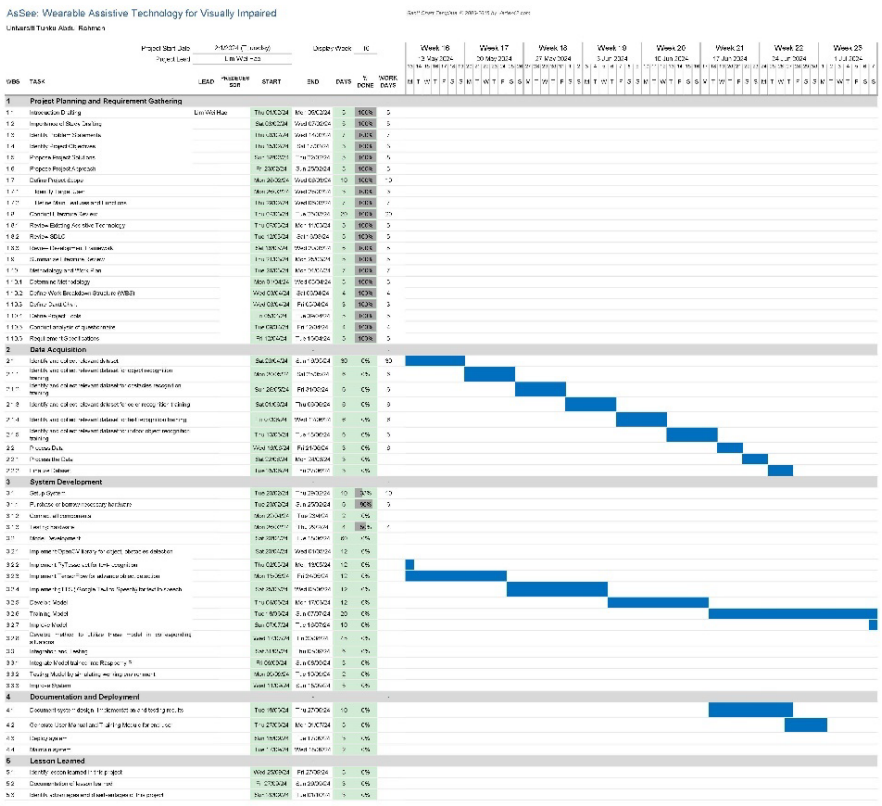


Figure 3.7 Gantt Chart 3

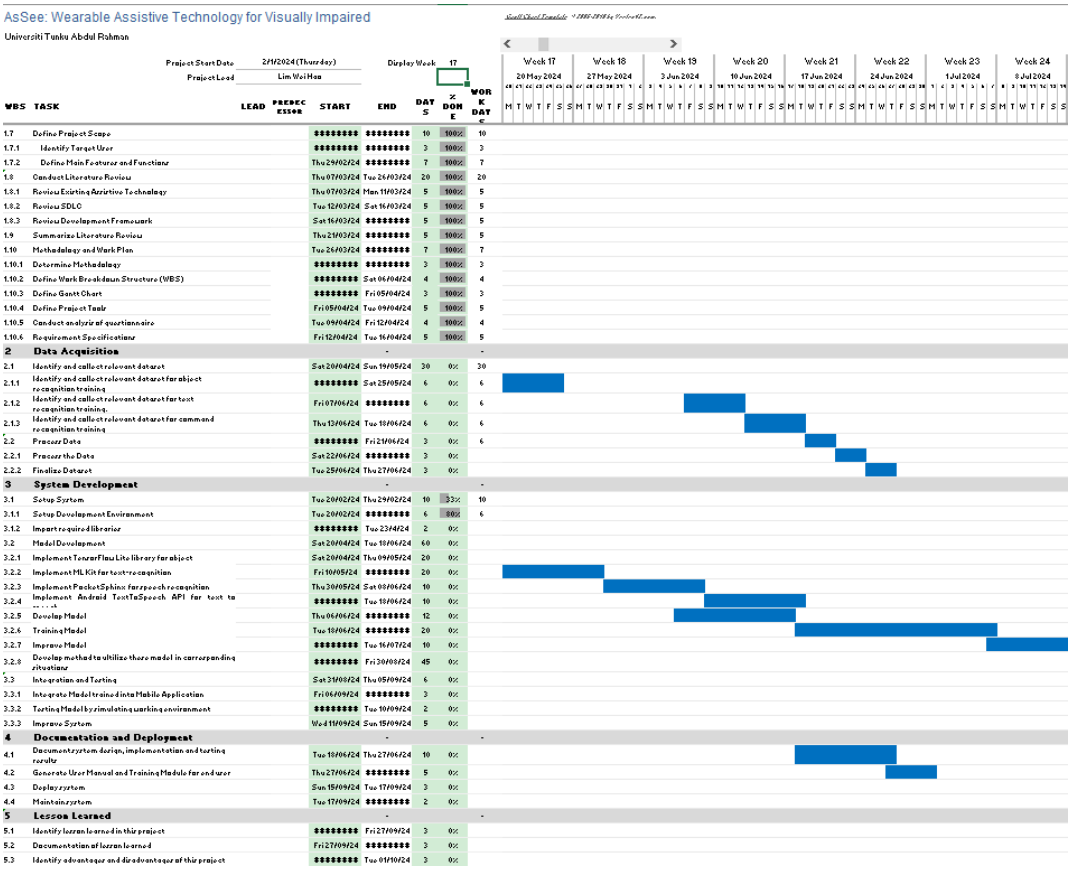


Figure 3.8 Gantt Chart 4

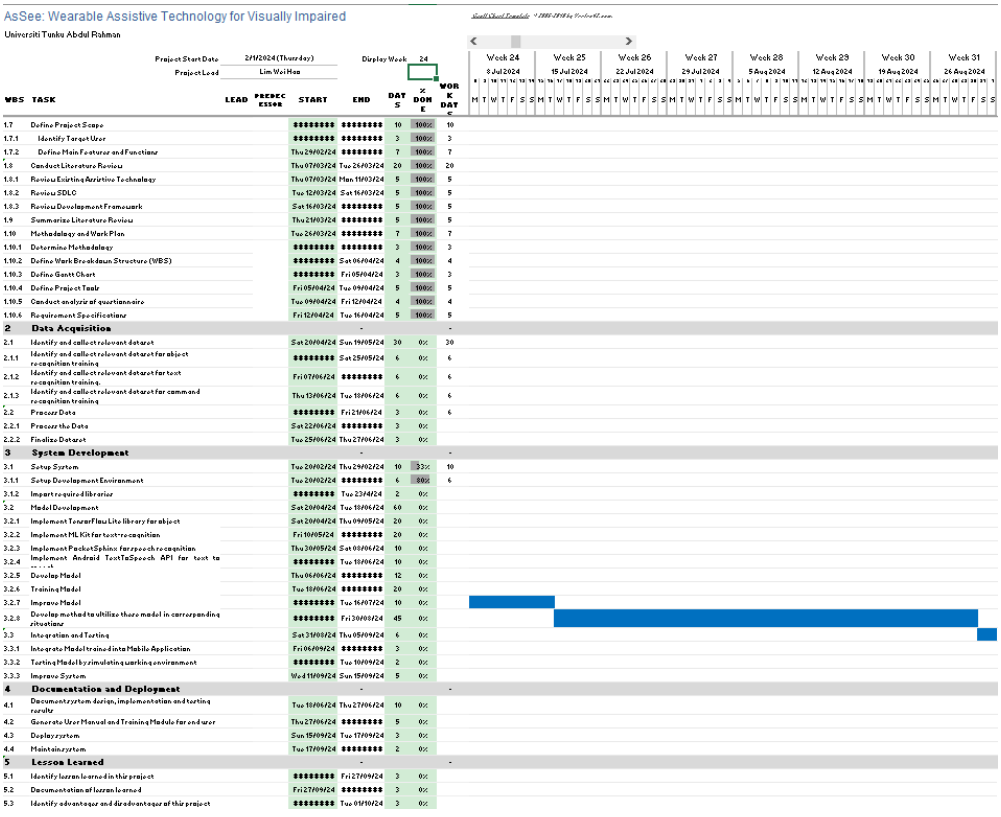


Figure 3.9 Gantt Chart 5

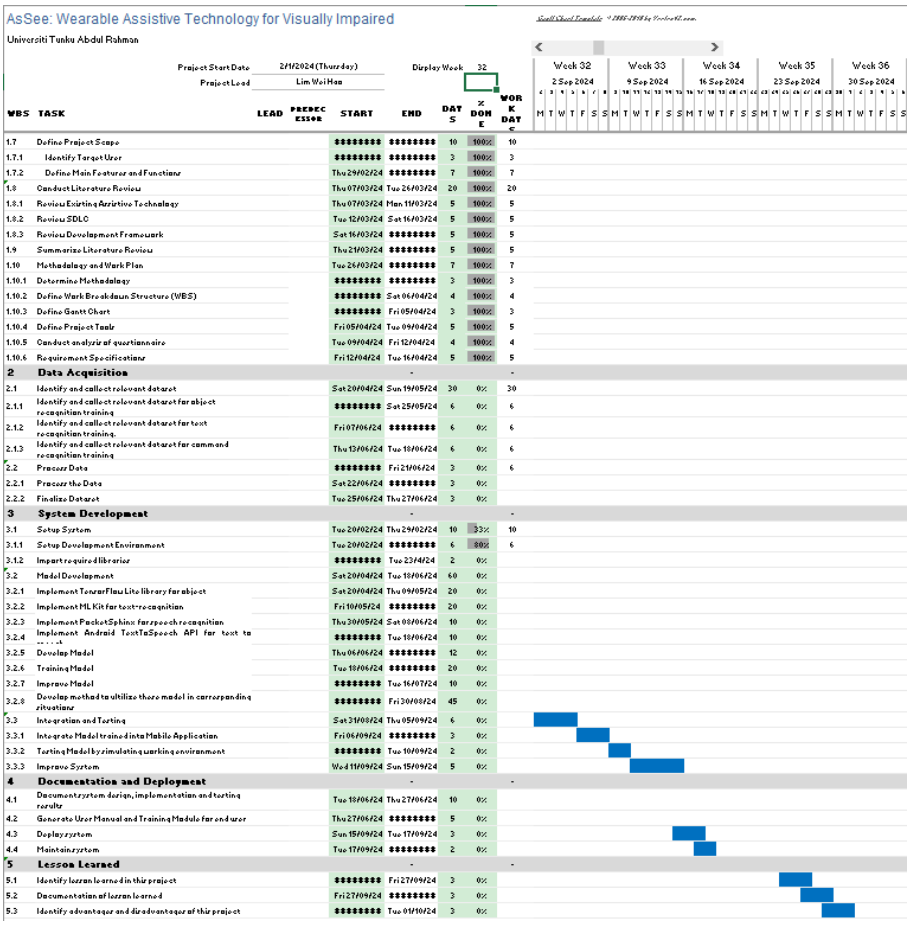


Figure 3.10 Gantt Chart 6

3.4 Project Development Tools

Project development tools plays crucial roles in system development process. The tools selected to develop AsSee system consists is Kotlin,TensorFlow Lite,Android TextToSpeech API,ML Kit, Visual Studio Code and PocketSphinx.

3.4.1 Kotlin

Figure 3.11 Kotlin - Development Tools

The programming language that selected fo this project is Kotlin due to its support from Android Ecosystem and it was officially supported by Google. Besides, it is an native language for Android application, therefore the performance and tools to make application more perfect are ready. Besides, due to its null safety where can increase the development speed. Furthermore, due to its code is easy to read and fewer code needed to develop an application, which suitable for the development within limited time. In addition, there are a lot of supported Libraries and frameworks which make the features and functions could be support by mature libraries and frameworks which required lesser code from scartch which also an advantages. Last but not least, the ability to reduce the bugs and error inside the code where it helps to reduce the time for maintenance and debugging which also helps developers to save time.

In conclusion,Kotlin are suitable and perfect suite for Android applications development where due to its interoperability with Java and the support from libraries and frameworks make it selected for the language used for this project (Ravi Makhjia, 2023).

3.4.2 Visual Studio Code

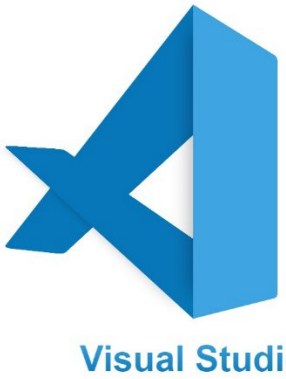


Figure 3.12 Visual Studio Code - Development Tools

Visual Studio Code is one of the famous platform for developers to code their software and it compatible with multiple languages and with different extension enable developers to develop software and make the UI of visual studio code is more prettier and code looks neat. Therefore, Visual Studio Code selected for programming purposes of AsSee system.

3.4.3 Android TextToSpeech API

This is an Android Application Programming Interface (API) that let your device speak, which supports multiple languages. This API will convert the text that accepted and convert it into speech and make your devices speak it out. With all the preset configuration, the API able to identify and speak out the language that user want based on the settings.In conclusion, this API enables your application and phone to speak which it is suitable for development of mobile application for disabilities (Anon, n.d.).

3.4.4 TensorFlow Lite

Figure 3.13 TensorFlow Lite

TensorFlow Lite is an end-to-end platform for machine learning, which it enables developers to deploy their trained model for mobile and edge devices and helps developer convert TensorFlow model to make it fit with mobile application. Besides, the compability of TensorFlow Lite on different operating systems also one of the factors selected for object detection library. Furthermore, the TensorFlow Lite which is the deep learning framework that implement edge computing which enable real-time detection on edge devices. Lastly, TensorFlow Lite work with TensorFlow where it convert the TensorFlow trained model to make it work with edge devices with less memory and cost consumption (Gaudenz Boesch, 2024).

3.4.5 ML Kit

Figure 3.14 ML Kit

ML Kit is a SDK that integrate the Google's machine learning expertise into mobile applications. ML Kit allows users to process real-time camera captured images or videos regardless of devices without connection to internet (Google, 2024). ML Kit was used to handle use cases like Text recognition, face recognition, barcode scanning, image labelling and Landmark recognition, where in this project it will used for text recognition and work with Android TextToSpeech API to perform text to speech functionality.

3.4.6 Vosk Speech Recognition

Figure 3.15 Vosk Speech Recognition

The decision to use Vosk is driven by its open-source nature, accuracy, and the ease with which it can be employed in various industries. (PySquad, 2024). The vosk is an offline opensource toolkit that support various different devices to integrate with. Besides, it supports up to 20+ languages and dialects where it also provides user to train own dictionary for own use case. In addition, it also supports different languages environment to implement it with different languages and different platforms.

3.5 Summary

In conclusion, with the methodology that chosen and integrate to this system, the AsSee system should be more complete and beneficial to the end user and satisfy end users expectation and requirements. Besides, with the well-planned

WBS and Gantt Chart, it ensure that the development process can be on-time and always reviewable to ensure the system in on schedule and to do changes to ensure the development process is secure and trackable.

CHAPTER 4

PROJECT INITIAL SPECIFICATION

4.1 Introduction

This section discussed the requirements gathering results and the requirements for the application. Besides, also discussed the tools employed into the development of this project and show the expectations of user toward this project and with the prototype showed to demsotrate the working environment of the application.

4.2 Fact Finding

There are 20 response that I had received through the survey on AsSee which is the Assistive Mobile Application Technology for Visually Imparied Individuals. The fact finding section summarized all the collected responds. The responder is all Malaysian.

4.2.1 Question 1

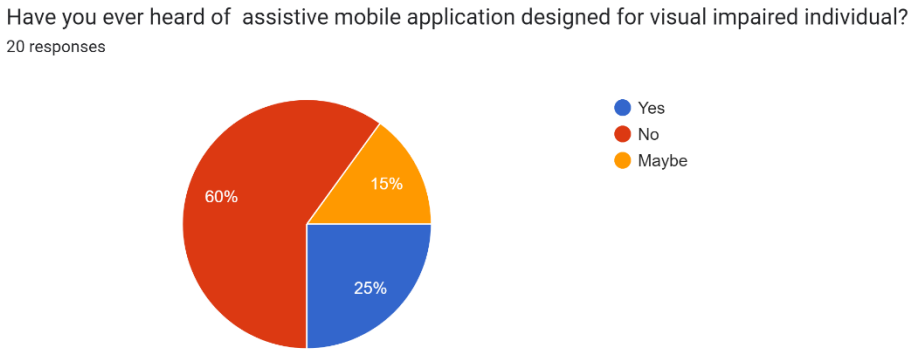


Figure 4.1 Questionnaire - Q1

The first question is “Have you ever heard of assistive mobile application designed for visual impaired individuals?”. There are 60% of respondents answered No, 15% Maybe and 25% Yes. Through this question, we can tell that in Malaysia, the assistive mobile application for impaired individuals is not fully popualrized in Malaysia.

4.2.2 Question 2

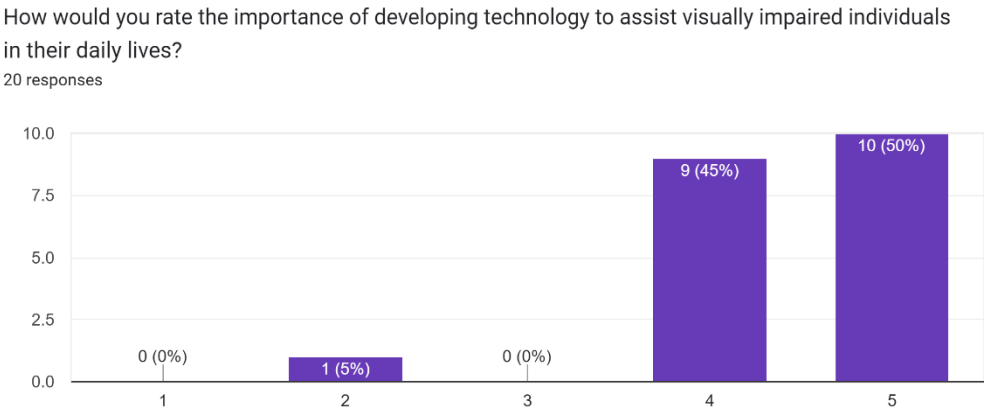


Figure 4.2 Questionnaire - Q2

The second questions show that 50% of responder think that it is extremely important to develop technology to assist virtually impaired individuals in their daily lifes, 45 % think it is important and 5% think it is moderate that develop technology will assit them in daily lives. In summary, it show that it is necessary to have and develop the technology to help them in daily lives, which can increase daily tasks productivity and their experience in daily routines.

4.2.3 Question 3

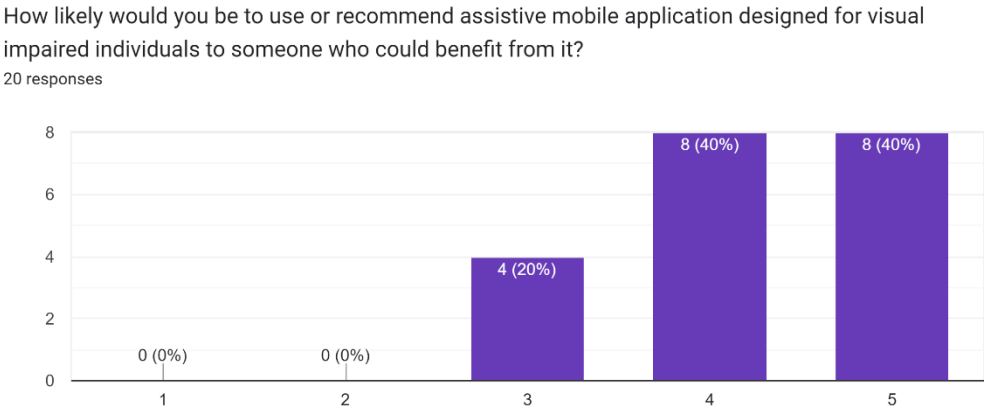


Figure 4.3 Questionnaire - Q3

The third question showed that 40% which is 8 responder will definitely recommend assistive mobile application designed for visually impaired individuals to someone they think who could benefit from it, besides, there are

40% would recommend and 4 respondents which is 20% may recommend to who may benefit from them.

4.2.4 Question 4

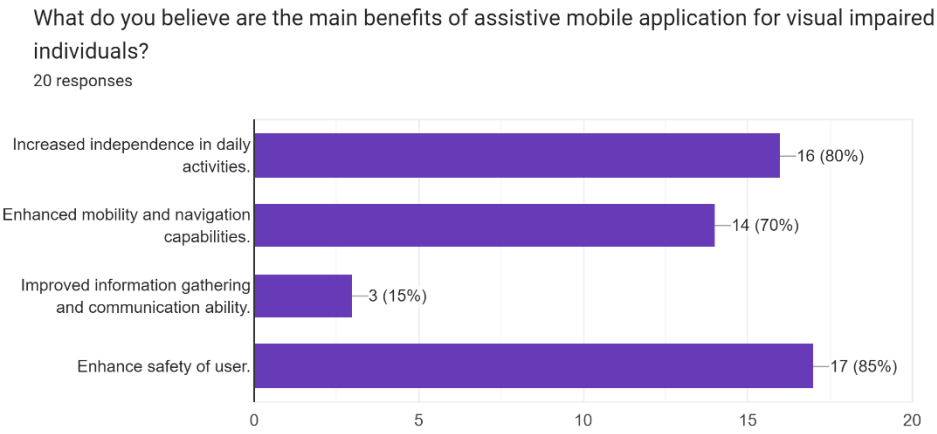


Figure 4.4 Questionnaire - Q4

There are 16 respondents which is 80% think that the introduction of assistive mobile application for visually impaired individuals will help them in increase their independence in daily activities, 70% of respondents think that it can enhance their mobility and navigation capabilities, 3 respondents think that assistive mobile application can improve their information gathering and communication ability and 17 which is 85% of the respondents think that it can enhance the safety of user.

4.2.5 Question 5

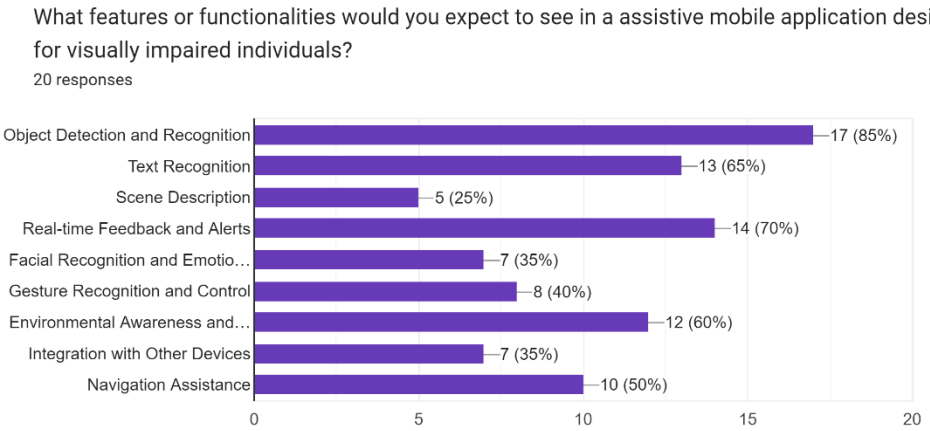


Figure 4.5 Questionnaire - Q5

The expected features or functionality in assistive mobile application designed for visually impaired individuals is 85% think it should help user in object detection and recognition, 70% think it should have real-time feedback and alerts, 65% think it should help user to recognize the text, 60% think it should aware and adapt to environment, 50% think it should provide navigation assistance, 40% think it should recognize and control gesture, 35% think it should recognize face and emotion and integrate with other device, and only 25% think it should recognize the scene.

4.2.6 Question 6

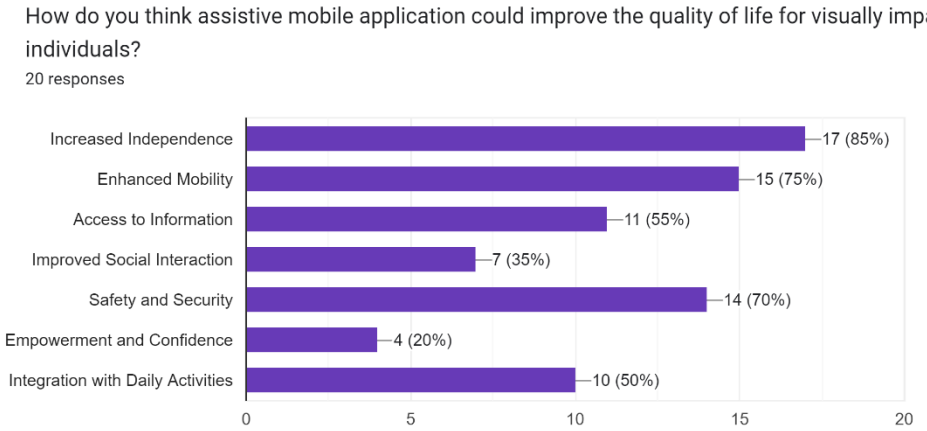


Figure 4.6 Questionnaire - Q6

The assistive mobile application can improve the quality of life for visually impaired individuals by a lot of different aspects. There are 85% think that it increased their independence so their quality of life increased too, 75% think that it enhanced their mobility, 70% think that it increased their safety and security, 55% think that increase their access to information for them to make decisions, 50% think it integrated with daily activities so that quality of life could be improved, 35% think it improved their social interaction and only 20% think it enhance their empowerment and confidence.

4.2.7 Question 7

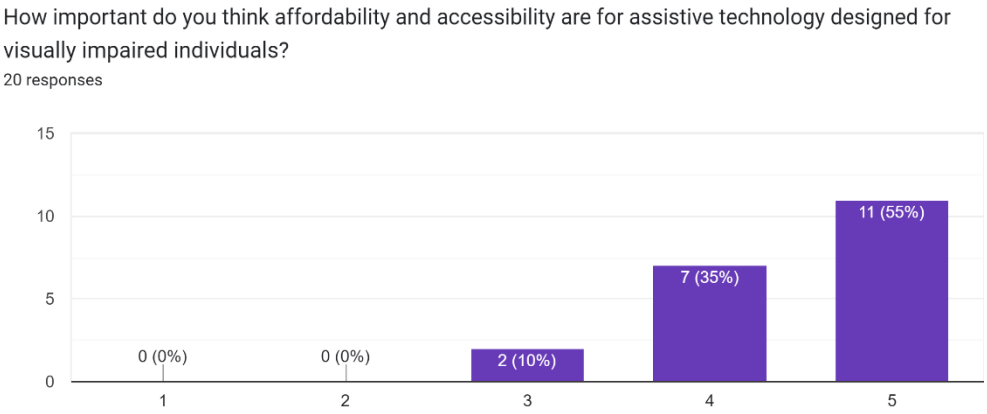


Figure 4.7 Questionnaire - Q7

Among all the respondents, there are 55% think that it is extremely important that the technology is accessible and affordable.Besides, there are 35% think that the accessibility and affordability of the technology is important and 2 respondents which is 10% think that it is moderate to have accessible and affordable technology.

4.2.8 Question 8

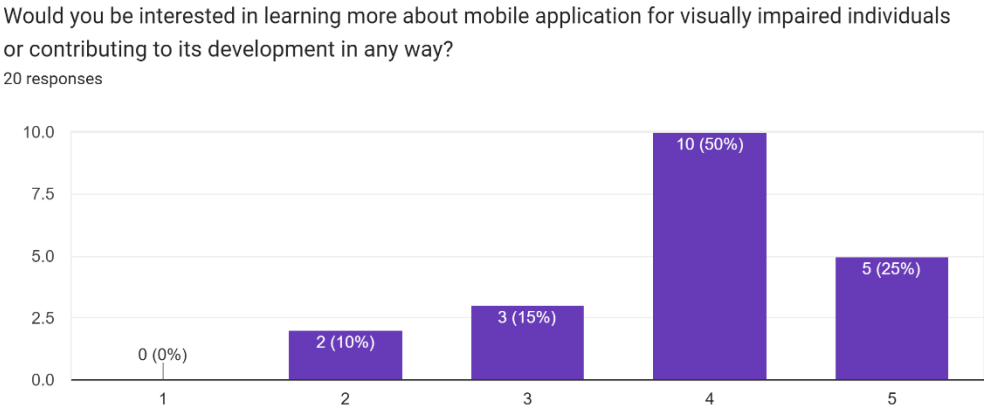


Figure 4.8 Questionnaire - Q8

There are 5 respondents which is 25% think that they are very likely to contribute and participate in assistive mobile application for visually impaired individuals, 10 respondents which is 50% think that there are willing to participate and contribute to assistive mobile application development, 3 respondents which is 15% think that they may participate and contribute to the

technology development, and 2 respondents which is 10% not interested in participating and contributing to the technology development.

4.3 Functional Requirements

The AsSee functional requirements:

Table 6 Functional Requirements

No.	Functional Requirement
i.	AsSee system should be able to recognize object.
ii.	AsSee system should be able to recognize object or text real-time.
iii.	AsSee system should be able to recognize multiple objects that are captured by the camera.
iv.	AsSee system should be able to let the user toggle detection mode.
v.	AsSee system should be able to convert recognized objects into speech.
vi.	AsSee system should understand simple commands from the user.
vii.	AsSee system should be able to recognize text.
viii.	AsSee system should be able to convert captured text into speech.

4.4 Non-functional Requirements

The AsSee system non-functional requirements:

Table 7 Non-functional Requirements

No.	Requirement
i.	The AsSee system should be reliable, providing consistent assistance to visually impaired individuals with a low error rate.
ii.	The AsSee system should have low latency and high performance, providing real-time responses to users with feedback returned.
iii.	The AsSee system should have high usability, requiring little to no training time for users to master.
iv.	The AsSee system's returned speech format information should be understandable, accurate, and clear.
v.	The AsSee system's gathered information should be accurate.
vi.	The AsSee system should be robust and resilient, functioning well in various situations and weather conditions.

vii.	The AsSee system should be affordable for individuals who need this technology to improve their quality of life.
viii.	The AsSee system should have high availability.
ix.	The AsSee system should provide functionality without requiring an internet connection.
x.	The AsSee system should be user-friendly for visually impaired individuals.

4.5 Prototype

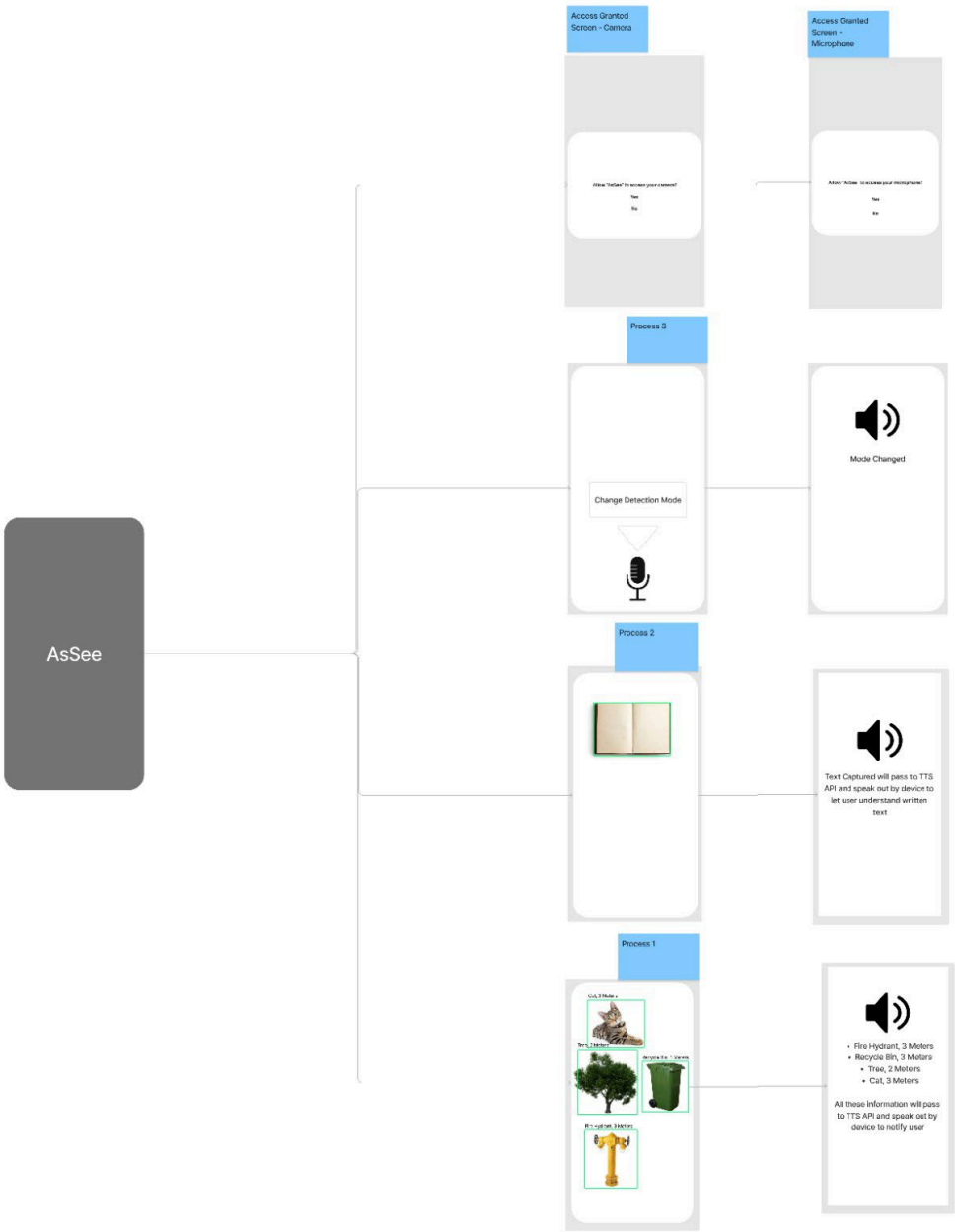


Figure 4.9 Prototype for AsSee – 1

This is the prototype of the AsSee system which showed 4 process in the system. First and foremost, when the user open the application for the first time, the user will be requested to grant access to AsSee application with the camera and microphone access for ultilzize the function of the application.

Besides, the process 1 which showed the object recognition function where the user point the camera towards the object surrounded and the system will recognize and label the object, after the recognition the details of the object will speak out by the devices to notify user with the object category.

Furthermore, the process 2 where indicates the text recognition function where when user point camera towards the document wanted to read and the mobile application will automatically capture the text if document spotted and recognized. After the text recognized, the content will speak out by the device to help user understand the written text.

Last but not least, the speech recognition function where to accept user simple command to interact with the mobile application like user wish to toggle detection mode between object detection mode and text recognition mode.

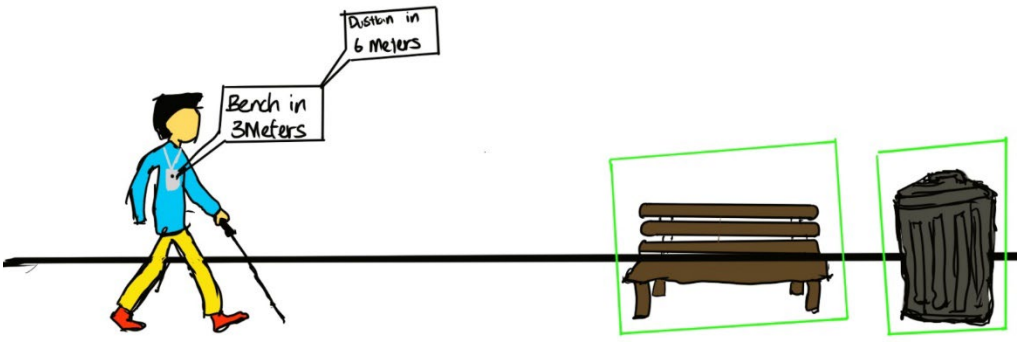


Figure 4.10 Demostration of AsSee Using Scenario - Object Detection

This is the demonstration of the Object Detection which is the Process 1, when the camera recognizes the objects and it will speak out the details of the object like category to notify user.

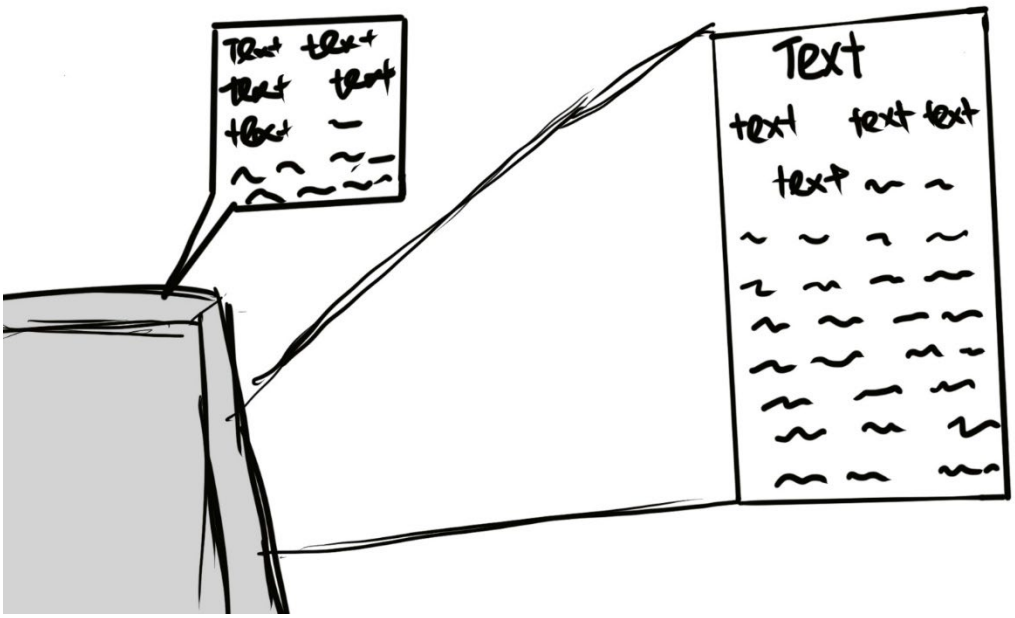


Figure 4.11 Demostration of AsSee Using Scenario - Text Recognition

This is the demonstration of the Text Recognition where the camera spotted and recognize the document and convert it into speech that enables user to understand the written document.

4.6 Dataset

This section briefly describes the dataset will use to train the model.

4.6.1 COCO

Common Object in Context where will be use to train the object detection model where it is the dataset that contains more than 330k images and the categories was clearly specific and up to 80 object cateogories which is every suitable for real-time environment object detection function.(COCO, n.d.)

4.6.2 IAM Handwriting Database

IAM Handwritting Database which selected to train the text recognition model where it contains 115320 isolated and labelled words with contribution of 657 writers of their handwriting which is very suitable to train the model to make it understand and read document.(kaggle, n.d.)

4.6.3 Mozilla Common Voice

The dataset that use to train the speech recognition model is Common Voice since it come with more than 30k hours of recorded data to train the model(Common Voice, n.d.)

4.7 Use Case Modelling

4.7.1 Use Case Diagram

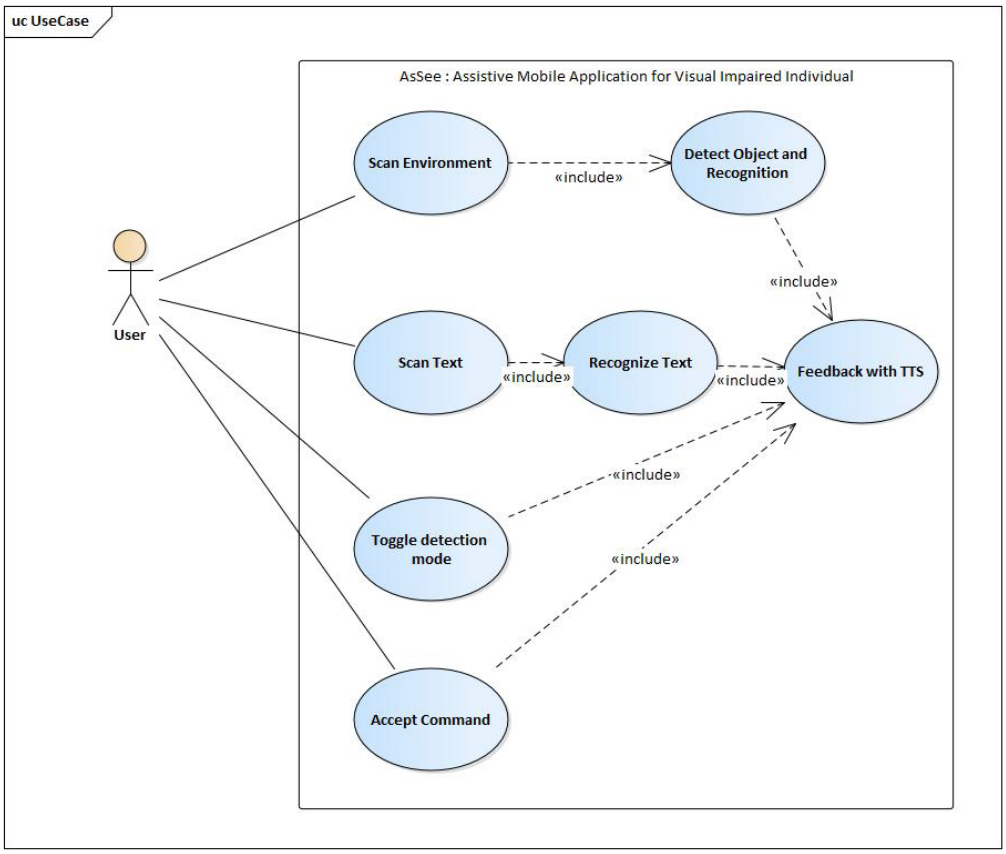


Figure 4.12 Use Case Diagram

4.7.2 Use Case Description

Use Case Name: Scan Environment	ID: UC01	Importance Level: High
Primary Actor: User	Use Case Type: Detailed, Essential	
Stakeholders and Interests: User – User using application wish to gain information surrounding.		
Brief Description: This use case describes how a user use AsSee to gather information of the environment.		
Trigger: When a user wish to gather and understand information of the environment.		
Relationships: Association: User Include: Detect Object and Recognition Extend: - Generalization: -		
Normal Flow of Events: 1. User open the application 2. Application start scanning environment with camera		
Sub-flows:		
Alternative/Exceptional Flows: A-1: User Open the application for the first time 1. The mobile application request granted with camera access 2. User refuse to granted camera access. 3. Use Cases terminates.		
A-2 User Open the application for the first time 1. The mobile application request granted with camera access 2. User granted access. 3. Use Case Continue		
A-3 Camera malfunctioning or blocked 1. Camera couldn't capture any object 2. Use Case Terminates.		

Figure 4.13 Use Case Description - UC01

Use Case Name: Detect Object and Recognition	ID: UC02	Importance Level: High
Primary Actor: User	Use Case Type: Detailed, Essential	
Stakeholders and Interests: User – User using application wish to detect and recognize object and wish mobile application feedback the details to user.		
Brief Description: This use case describes how a user use AsSee to detect object and recognition.		
Trigger: When a user pointing camera toward object.		
Relationships: Association: User Include: Feedback with TTS Extend: - Generalization: -		
Normal Flow of Events: 1. Perform UC01 – Scan Environment 2. User pointing camera toward one direction or object 3. Camera capture every object available 4. System recognize object captured 5. System generate and label information of the object, distance, name and pass to TTS API 6. Perform UC03		
Sub-flows:		
Alternative/Exceptional Flows: A-1: Unclear environment or incorrect user action 1. Camera couldn't detect a full object 2. Use Case Terminate A-2: System couldn't recognize the object 1. Camera couldn't recognize object information 2. Use Case Terminate		

Figure 4.14 Use Case Description - UC02

Use Case Name: Feedback with TTS	ID: UC03	Importance Level: High
Primary Actor: User	Use Case Type: Detailed, Essential	
Stakeholders and Interests: User – User using application wish to understanding the environment and text with the speech from the application		
Brief Description: This use case describes how a user use AsSee to feedback to user with the speech.		
Trigger: When a user scanned text or environment.		
Relationships: Association: User Include: - Extend: - Generalization: -		
Normal Flow of Events: 1. Perform UC01/UC04/UC07 2. System received captured object information/captured text/command received 3. System speak out the information/captured text content/respond to command.		
Sub-flows:		
Alternative/Exceptional Flows: A-1: Speaker malfunctioning 1. System couldn't speak out information gathered 2. Use Case Terminates A-2: Connected audio device malfunctioning 1. System couldn't speak out information gathered through device 2. Use Case Terminates A-3: Unstable connection with audio device 1. System spoke out information but user receive it blurly. 2. Rerun UC01 or UC04		

Figure 4.15 Use Case Description - UC03

Use Case Name: Scan Text	ID: UC04	Importance Level: High
Primary Actor: User	Use Case Type: Detailed, Essential	
Stakeholders and Interests: User – User using application wish to gain information and understand written text.		
Brief Description: This use case describes how a user use AsSee to understand written text.		
Trigger: When a user wish to understand written text.		
Relationships: Association: User Include: Recognize Text Extend: - Generalization: -		
Normal Flow of Events: 1. User open the application 2. Application start scanning text with camera 3. Perform UC05		
Sub-flows:		
Alternative/Exceptional Flows: A-1: User Open the application for the first time 1. The mobile application request granted with camera access 2. User refuse to granted camera access. 3. Use Cases terminates.		
A-2 User Open the application for the first time 1. The mobile application request granted with camera access 2. User granted access. 3. Use Case Continue		
A-3 Camera Malfunctioning 1. The camera couldn't scan text 2. Use Case Terminates		

Figure 4.16 Use Case Description - UC04

Use Case Name: Recognize Text	ID: UC05	Importance Level: High
Primary Actor: User	Use Case Type: Detailed, Essential	
Stakeholders and Interests: User – User using application wish application to recognize written text.		
Brief Description: This use case describes how a user use AsSee to recognize written text.		
Trigger: When a user wish AsSee recognize written text.		
Relationships: Association: User Include: Feedback with TTS Extend: - Generalization: -		
Normal Flow of Events: 1. Perform UC04- Scan Text 2. Camera capture text detected 3. Send captured text to TTS API 4. Perform UC03		
Sub-flows:		
Alternative/Exceptional Flows: A-1 Unknown Language or Blur text 1. The camera couldn't capture and recognize text 2. Use Case Terminates		

Figure 4.17 Use Case Description - UC05

Use Case Name: Toggle detection mode	ID: UC06	Importance Level: High
Primary Actor: User	Use Case Type: Detailed, Essential	
Stakeholders and Interests: User – User using application wish to swap detection mode from object detection to text recognition or text recognition to object detection		
Brief Description: This use case describes how a user toggle detection mode of AsSee		
Trigger: When a user interact with system to toggle mode		
Relationships: Association: User Include: - Extend: - Generalization: -		
Normal Flow of Events: 1. User press on button to toggle detection mode 1. Perform UC03 to notify user mode changed.		
Sub-flows:		
Alternative/Exceptional Flows: A-1: Speaker malfunctioning 1. System couldn't inform use changes made 2. Use Case Terminates A-2: Connected audio device malfunctioning 1. System couldn't inform use changes made 2. Use Case Terminates A-3: Unstable connection with audio device 1. System spoke out information but user receive it blurly. 1. Rerun UC06		

Figure 4.18 Use Case Description - UC06

Use Case Name: Accept Command	ID: UC07	Importance Level: High
Primary Actor: User	Use Case Type: Detailed, Essential	
Stakeholders and Interests: User – User using application wish to use simple command to interact with AsSee system		
Brief Description: This use case describes how a user interact with AsSee system with simple command		
Trigger: When a user wish AsSee respond to simple command from user.		
Relationships: Association: User Include: Feedback from TTS Extend: - Generalization: -		
Normal Flow of Events: 1. User open the application 2. Application start listening to user 3. Perform UC05		
Sub-flows:		
Alternative/Exceptional Flows: A-1: User Open the application for the first time 1. The mobile application request granted with microphone access 2. User refuse to granted microphone access. 3. Use Cases terminates.		
A-2 User Open the application for the first time 1. The mobile application request granted with microphone access 2. User granted access. 3. Use Case Continue		

Figure 4.19 Use Case Description - UC07

CHAPTER 5

SYSTEM DESIGN

5.1 Introduction

In this chapter, the system design will be explained, to have better understanding of the system and dataflow, overall system design and the function working flow.

5.2 System Architecture Design

In this project, there are multiple AIs implemented, Vosk for Speech Recognition, EfficientDet work with Tensorflow Lite for Object Detection, and ML KIT Text Recognition selected for text recognition. There are different function assigned to different AI implemented and with different class for them to perform own unique function. This diagram showed the workflow of the AsSee application.

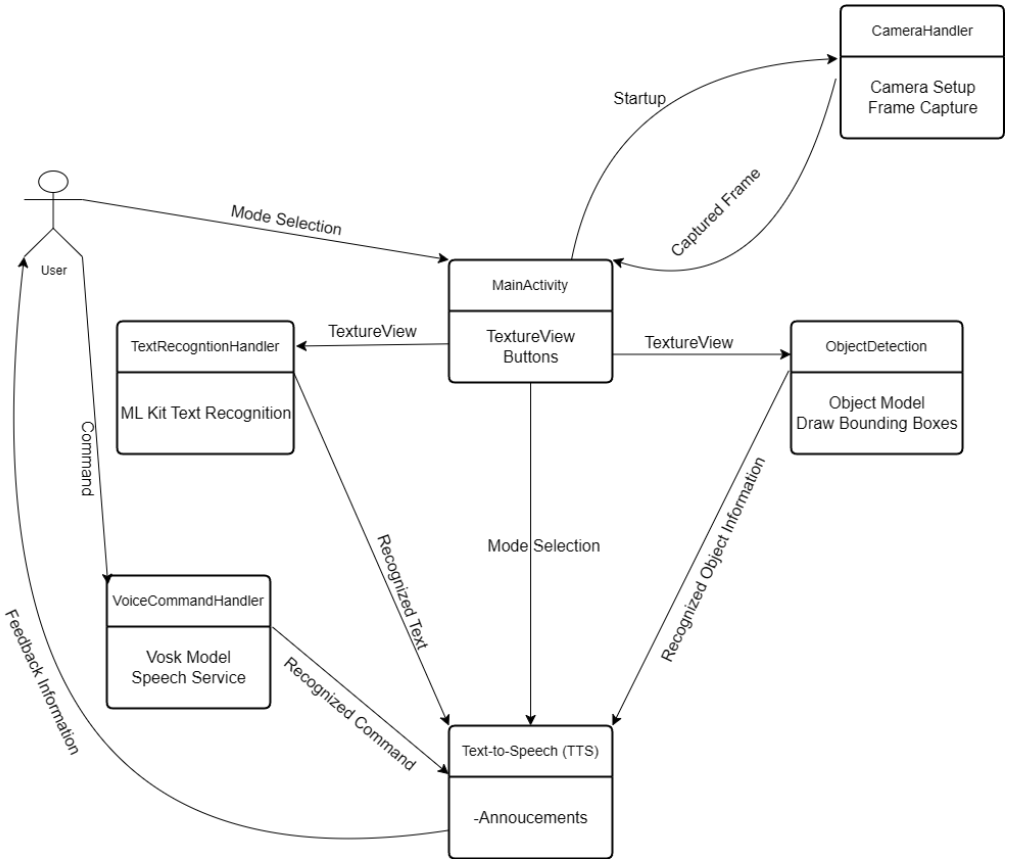


Figure 5.1 AsSee System Architecture Design

5.2.1 MainActivity

The MainActivity serves as the central hub of the application, where it interacts with different components. Besides, it helps to initialize the UI elements and helps user in button to change different mode, which is object detection mode and text recognition mode. Furthermore, it helps user to wake the voice assistant when they need voice assistant to help them to do corresponding action.

5.2.2 CameraHandler

This class manages the camera, and passes the captured frame to the other components like object detection and text recognition. Besides, this class handle and manage to open the camera preview session as the TextureView which ultilize by the ObjectDetection and TextRecognitionHandler

5.2.3 TextRecognitionHandler

The TextRecognitionHandler will also captures the current frame of the TextureView if the user long press the screen. The TextRecognitionHandler will pass the frame to the ML Kit Text Recognition tools to recognize the text, and then pass the recognition result to text to speech function to return information to the user.

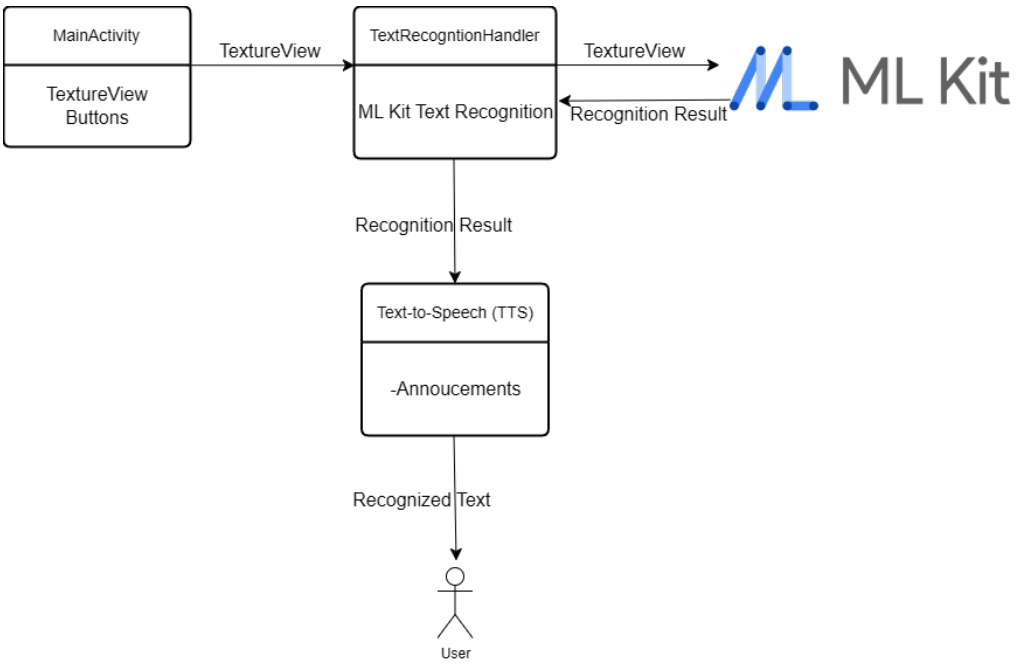


Figure 5.2 Text Recognition Module Architecture Design

5.2.4 ObjectDetection

ObjectDetection is the class that provides function and provide assistance in the object detection and return information to the user. This function will process the frame that passed by the CameraHandler function and then process it with the TensorFlow Lite model that support by EfficientDet, after the process of detection, this function will draw bounding box around the object with the label of the object recognized.

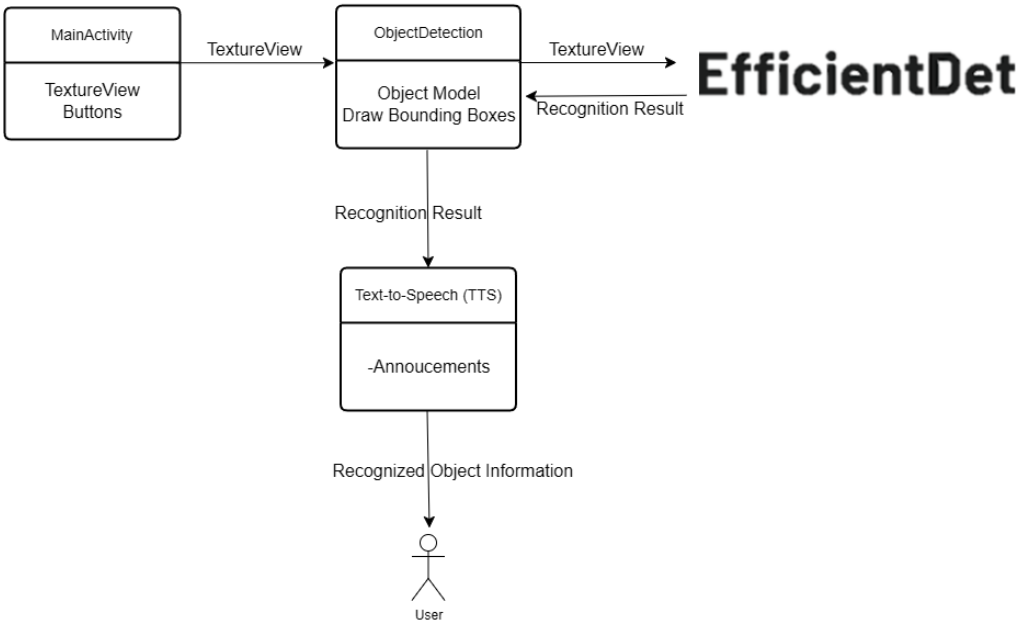


Figure 5.3 Object Detection Module Architecture Design

5.2.5 VoiceCommandHandler

This function will wake if the user press the voice assistant button, then the user will speak out command like ‘object mode’ or ‘text mode’ to switch between different mode to assist them in specific situations and the speech recognition doesn’t required the internet connection, because the model is install locally in the device.

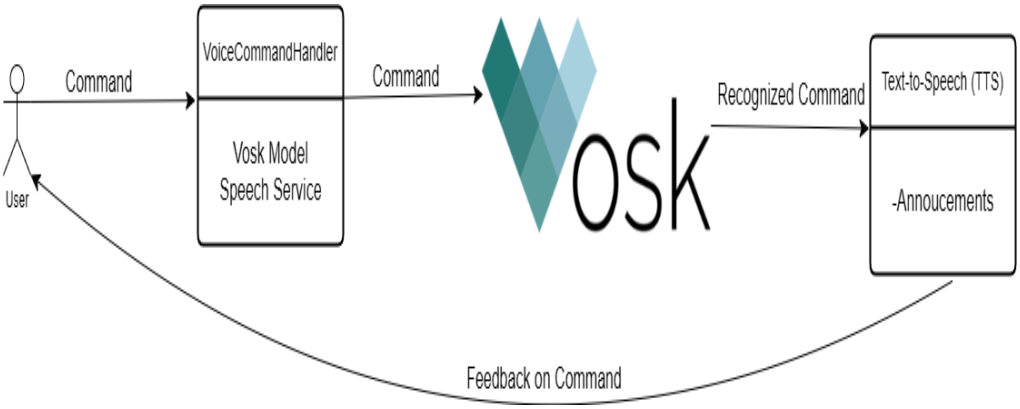


Figure 5.4 Speech Recognition Module Architecture Design

5.3 System Design

5.3.1 Object Detection

Figure 5.5 Object Detection Application Screenshot

This is the working environment demonstration of the Object Detection, from the image above, it showed the Object Detection is working fine, with the class that covered by the COCO dataset which up to 80 classes. The object detection will detect the object and draw bounding boxes around the object detected and return information with speak out the class name of the object to the user.

5.3.2 Text Recognition

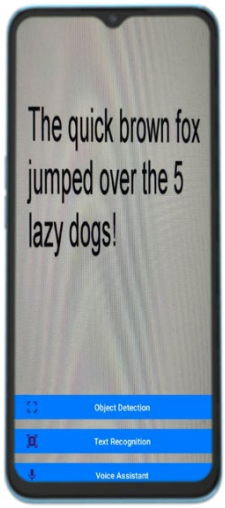


Figure 5.6 Text Recognition Application Screenshot

This is the working environment demonstration of the Text Recognition, from the image above, it showed the Text Recognition is working fine, and it will speak out the recognized text to the user for them to access information on document or written format information.

5.3.3 Speech Recognition

Figure 5.7 Speech Recognition Simulation

This is the working environment demonstration of the Speech Recognition, from the image above, it showed the Speech Recognition, and then the application will accept command from the user and then do action according to the command.

5.4 Activity Diagram

5.4.1 Change To Text Recognition Mode

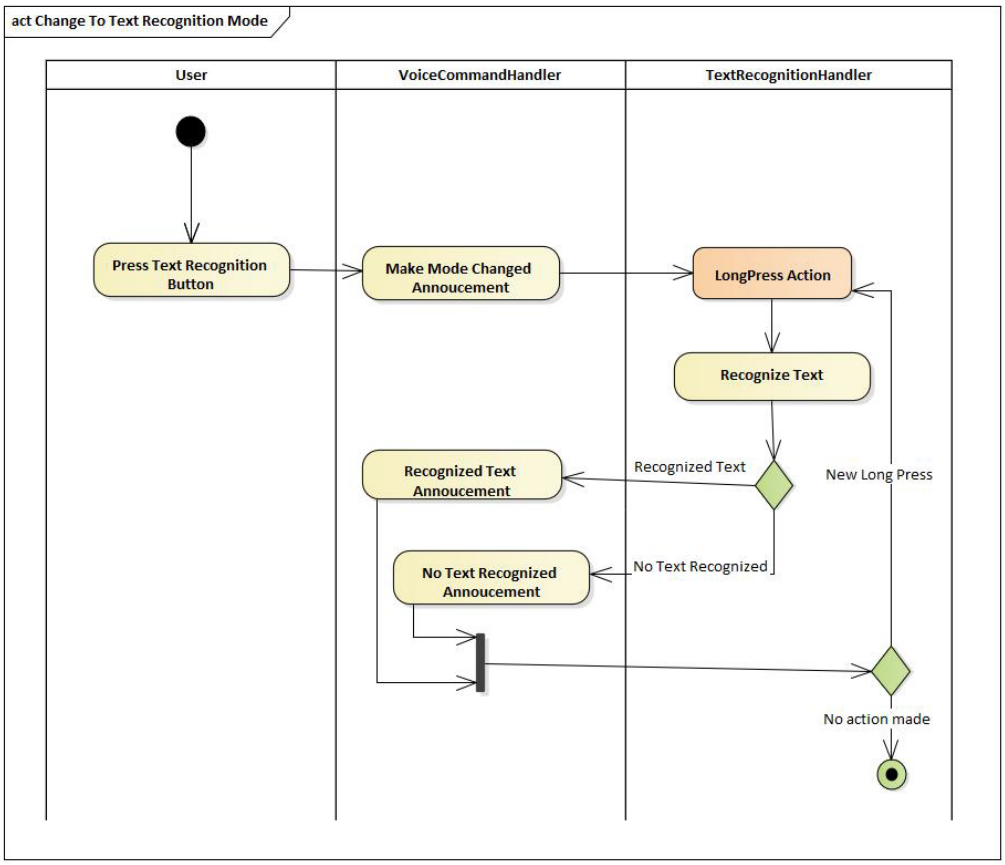


Figure 5.8 Activity Diagram for Change to Text Recognition Mode

This is the activity diagram of changing mode to the Text Recognition Mode where it showed that the user to toggle recognition mode with button. The user need to press the Text Recognition button on the screen and then the announcement of mode changed will be made, after the application detect the long press action from the user, it will start recognizing the current frame to recognize the text available, if there is text recognized then the system will speak the recognized text else the “No Text Recognized” announcement will be made.

5.4.2 Change To Object Detection Mode

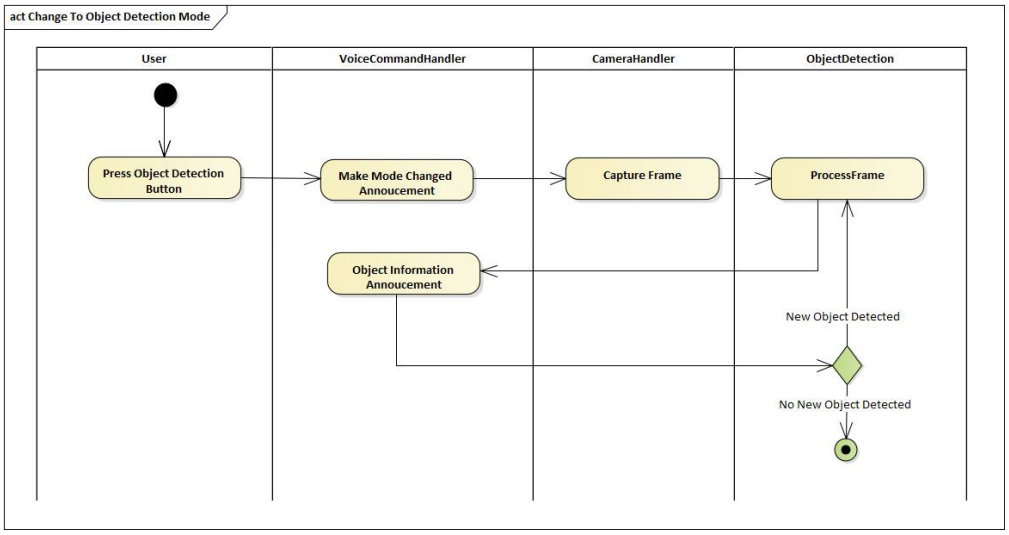


Figure 5.9 Activity Diagram for Change to Object Detection Mode

This is the activity diagram of changing mode to the Object Detection Mode where it showed that the user to toggle recognition mode with button. The user need to press the Object Detection button on the screen and then the announcement of mode changed will be made, the application will start detecting the frame that captures by the camera and speak out the object detected information to the user.

5.4.3 Speech Recognition Accept Command

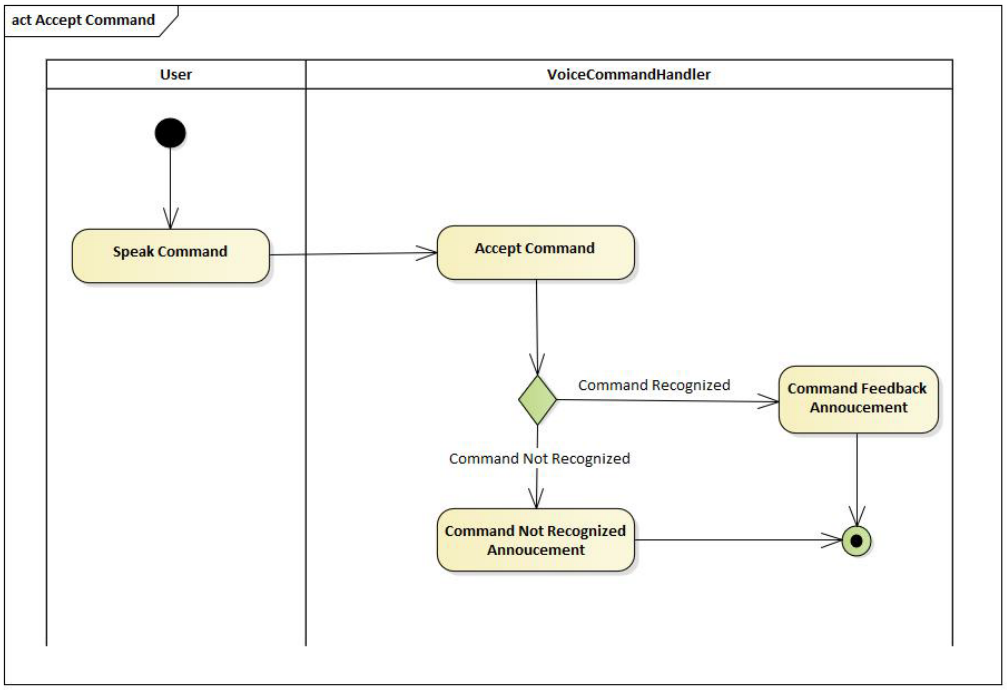


Figure 5.10 Activity Diagram for Speech Recognition Accept Command

This is the activity diagram of the application accepting command from the user, when the command recognized, it will do correspond action, for example, change to object detection mode if “Object” detected from the user speech.

CHAPTER 6

SYSTEM IMPLEMENTATION

6.1 Introduction

In this chapter, the system implementation will be discussed, to have more insight to the system implementation in different aspects.

6.2 MainActivity.kt

```
private lateinit var cameraHandler: CameraHandler
private lateinit var objectDetection: ObjectDetection
private lateinit var textRecognition: TextRecognitionHandler
private lateinit var voiceCommandHandler: VoiceCommandHandler

private lateinit var textureView: TextureView
private lateinit var button1: Button
private lateinit var button2: Button
private lateinit var button3: Button
```

Figure 6.1 Screenshot of MainActivity.kt

In the MainActivity.kt, will initial all the Handler and Function class in the MainActivity. TextureView is the real time stream from the camera and the buttons for “Object Detection”, “Text Recognition” and “Voice Assistant”.

```
enum class RecognitionMode {
    OBJECT_DETECTION, TEXT_RECOGNITION
}
```

Figure 6.2 Screenshot of MainActivity.kt 2

And the enum to determine the Recognition mode and depends on the mode to do corresponding action.

6.3 ObjectDetection.kt

```
for (i in 0 ≤ until < scores.floatArray.size) {
    val score = scores.getFloatValue(i)
    if (score > 0.75) {
        val objectClass = labels[classes.getIntValue(i)]
        if (!detectedObjects.contains(objectClass)) {
            detectedObjects.add(objectClass)

            paint.color = colors[classes.getIntValue(i) % colors.size]
            paint.style = Paint.Style.STROKE
            val location = locations.floatArray.sliceArray( indices: i * 4 ≤ until < (i + 1) * 4)
            canvas.drawRect(
                RectF( left: location[1] * w, top: location[0] * h, right: location[3] * w, bottom: location[2] * h), paint
            )
            paint.style = Paint.Style.FILL
            canvas.drawText(objectClass, x: location[1] * w, y: location[0] * h - 10, paint)
        }
    }
}
```

Figure 6.3 Screenshot of ObjectDetection.kt

This is the code in ObjectDetection where to filter the detections with confidence scores above 0.75 where aim to provide high believable only to the user, and then for those filtered results, will start drawing bounding boxes based on the locations of the object.

```
if (objectsToSpeak.isNotEmpty() && objectsToSpeak != lastSpokenObjects) {
    lastSpokenObjects = objectsToSpeak
    tts.speak(objectsToSpeak, TextToSpeech.QUEUE_FLUSH, params: null, utteranceId: null)
}
```

Figure 6.4 Screenshot of ObjectDetection.kt 2

And after filtered the result and draw bounding boxes, will send to tts to speak the detected object information.

6.4 TextRecognitionHandler

```
val touchListener = View.OnTouchListener { _, event ->
    when (event.action) {
        MotionEvent.ACTION_DOWN -> {
            longPressHandler = android.os.Handler()
            longPressHandler?.postDelayed(longPressRunnable, longPressDuration)
            true
        }
        MotionEvent.ACTION_UP, MotionEvent.ACTION_CANCEL -> {
            longPressHandler?.removeCallbacks(longPressRunnable)
            longPressHandler = null
            true
        }
        else -> false
    }
}

private val longPressRunnable = Runnable {
    recognizeText()
}
```

Figure 6.5 Screenshot of TextRecognitionHandler.kt

This is the touchListener where to handle the Long press action from the user, and if the long press was detected, will execute the recognizeText function.

```
private fun recognizeText() {
    val bitmap = textureView.bitmap
    if (bitmap != null) {
        val image = InputImage.fromBitmap(bitmap, rotationDegrees: 0)
        recognizer.process(image)
            .addOnSuccessListener { visionText ->
                recognizedText = visionText.text
                if (recognizedText?.isNotEmpty() == true) {
                    tts.speak(recognizedText, TextToSpeech.QUEUE_FLUSH, params: null, utteranceId: null)
                } else {
                    tts.speak( text: "No text recognized", TextToSpeech.QUEUE_FLUSH, params: null, utteranceId: null)
                }
            }
            .addOnFailureListener { e ->
                tts.speak( text: "Error recognizing text", TextToSpeech.QUEUE_FLUSH, params: null, utteranceId: null)
            }
    } else {
        Log.e( tag: "TextRecognitionHandler", msg: "Bitmap is null")
    }
}
```

Figure 6.6 Screenshot of TextRecognitionHandler.kt 2

And then the recognizeText will take the bitmap from the textureview to process and then if there is text detected, will speak out the recognized text else will speak “No Text Recognized”

6.5 VoiceCommandHandler.kt

```
private fun initVoskModel() {
    val modelDir = File(context.filesDir, child: "vosk-model-small-en-us-0.15")

    if (!modelDir.exists()) {
        modelDir.mkdirs()
    }

    try {
        voskModel = Model(modelDir.absolutePath)
    } catch (e: IOException) {
        Log.e( tag: "Vosk", msg: "Error initializing Vosk model: ${e.message}", e)
    }
}
```

Figure 6.7 Screenshot of VoiceCommandHandler.kt

This initVoskModel will init the vosk model which is the speech recognition model.

```
private fun handleCommand(hypothesis: String) {
    when {
        hypothesis.contains( other: "text", ignoreCase = true) -> {
            MainActivity.switchToTextRecognitionMode();
        }
        hypothesis.contains( other: "object", ignoreCase = true) ->
            MainActivity.switchToObjectDetectionMode();
        }
        else -> {
            speak( text: "Command not recognized")
        }
    }
}
```

Figure 6.8 Screenshot of VoiceCommandHandler.kt 2

And this will check the speech recognized, is it contain “text” or “object” to perform mode switching function.

6.6 Justification for Object Detection

Object detection is a technique that uses neural networks to localize and classify objects in images. This computer vision task has a wide range of applications, from medical imaging to self-driving cars. (What Is Object Detection? | IBM, n.d.). In simple words, object detection is computer vision task that helps to assist machine to locates objects in the images. Is it an instance that train computer to see like what human do with data to train it until it able to classification and recognize objects with categories.

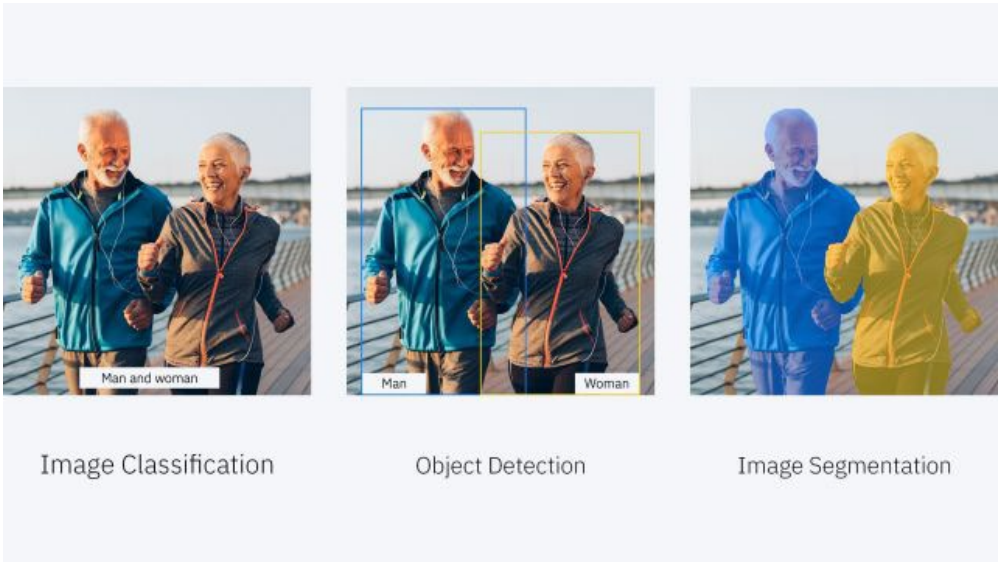


Figure 6.9 Object Detection Justification

Picture above showed the difference between Image Classification, Object Detection and Image Segmentation, where object detection handles some subtask of the other function and provide other information like categories and estimation of location in images.

6.6.1 Justification for EfficientDet Object Detection Model using TensorFlow Lite

There are multiple frameworks that's available provide neural networks unit that enable user and developer to train and assemble the neural networks models. But there are two major machine learning frameworks that mostly used by developers and users.

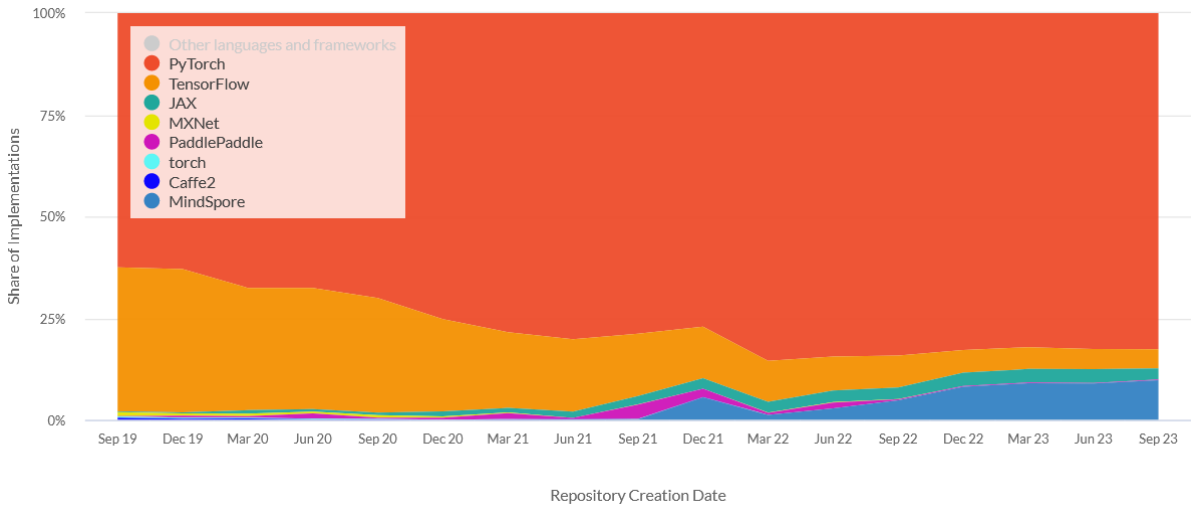


Figure 6.10 Comparison Graph for Object Detection Model (Boesch, G.,2023)

From the Graph above, can tell the coverage which is the share of implementation of PyTorch and TensorFlow covered most of the usage, and PyTorch was becoming the major framework for developer to assemble and train models.

6.6.1.1 TensorFlow Lite (Light version of TensorFlow)

TensorFlow Lite is an end-to-end platform for machine learning, which it enables developers to deploy their trained model for mobile and edge devices and helps developer convert TensorFlow model to make it fit with mobile application. Besides, the compatibility of TensorFlow Lite on different operating systems also one of the factors selected for object detection library. Furthermore, the TensorFlow Lite which is the deep learning framework that implement edge computing which enable real-time detection on edge devices. Lastly, TensorFlow Lite work with TensorFlow where it converts the

TensorFlow trained model to make it work with edge devices with less memory and cost consumption (Gaudenz Boesch, 2024).

6.6.1.2 Pytorch

PyTorch is the framework for user to assemble and train machine learning models that introduced in 2016. The PyTorch was the first framework that provide deep learning framework with focus on both speech and usability. Therefore, the usage of the PyTorch increase rapidly after years. Besides, PyTorch implement Pythonic Programming style which make it ease to use and learn.

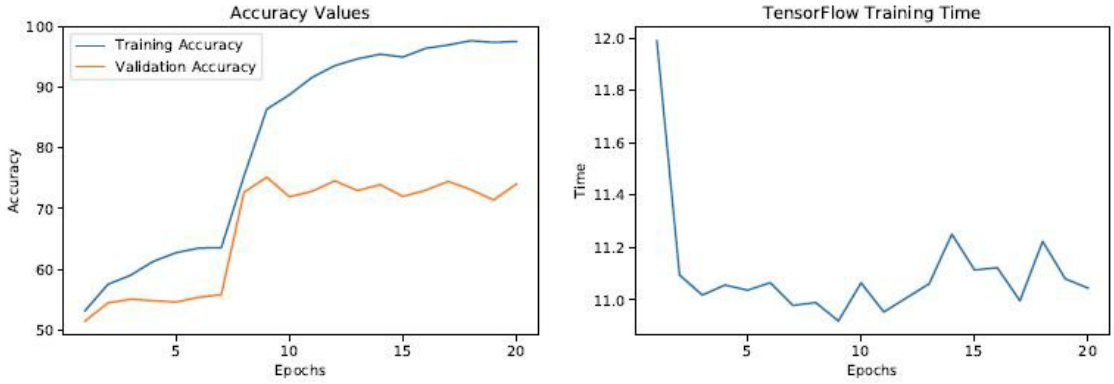


Figure 1: TensorFlow Accuracy and Training Time

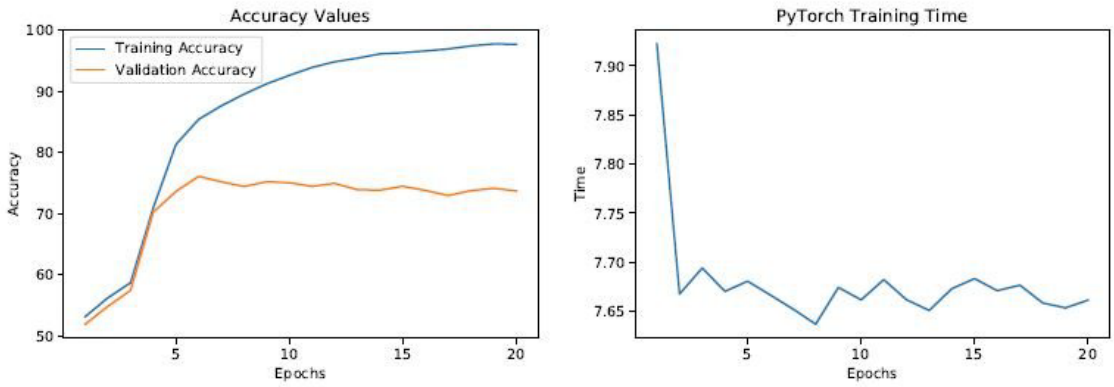


Figure 2: PyTorch Accuracy and Training Time

Figure 6.11 Comparison Graph for Pytorch vs Tensorflow (Boesch, G.,2023)

The graph above showed the **Training Time** and **Accuracy** under training of different dataset. For the accuracy, each framework was able to produce an AI Model that is accurate up to 78% after 20 epochs. But for the training time, the

PyTorch required lower training time which only cost 7.67 seconds at average and 11.19 seconds for TensorFlow training time. Therefore, the duration to train a complete TensorFlow model is slightly longer than PyTorch, but the required RAM to train the model where only 1.7GB RAM for TensorFlow and 3.5GB Ram for PyTorch.

Besides, the maturity and support of different APIs is much better than PyTorch.

6.6.2 Justification for TensorFlow Lite

The framework that chooses for assemble and export the model to make it compatible with the Android. The reason why TensorFlow Lite selected is because, it reduces the barrier of the computer vision and machine learning into different devices, therefore it makes those computer-vision or machine learning models applicable to mobile, so TensorFlow Lite selected for the framework to assemble the model. Furthermore, the TensorFlow Lite enable to convert the TensorFlow model to TensorFlow Lite model which applicable to the android application, and there is a lot of well-known models also with ready conversion function to convert their model to. tflite model. Besides, the TensorFlow Lite was able to provide minimal latency models that provide real-time performance with inference time decreased. In addition, the well-designed of the TensorFlow system provide a simple way for user to implement and integrate .tflite model into the Android or iOS application.

Figure 6.12 Screenshot of Android Studio

The Android Studio provide simple way to import the TensorFlow Lite Model into the application.

6.6.3 Model Comparison

There is multiple different model available for object detection on TensorFlow Lite.

Mode Type	Latency of model	(Accuracy)mAP	FPS	Real-Time
R-CNN	High	~60	<1	NO
Fast R- CNN	Medium	~70	<1	NO
Faster R-CNN	Medium	~70	7	NO
YOLO	Low	~60	46	YES
SSD	Low	~74	59	YES

Figure 6.13 Comparison Table for Different Models (Bhagwat, S.,2021)

There are multiple models available for object detection, and there is different performance for different model, from the table above that comparing the Latency where to show the response time needed for the model, where YOLO (You Only Look Once) and SSD (Single-Shot Detector) where show the real-time performance which is best fit for the AsSee application, besides, the accuracy which is the mAP metrics of the model showed that SSD has the best performance where to compare with others where up to 74% in Android devices. Besides, for the FPS comparison, the SSD also have higher performance among all the models. But there are one more model that is not compared therefore, there is another comparison research to compare between YOLO and EfficientDet.

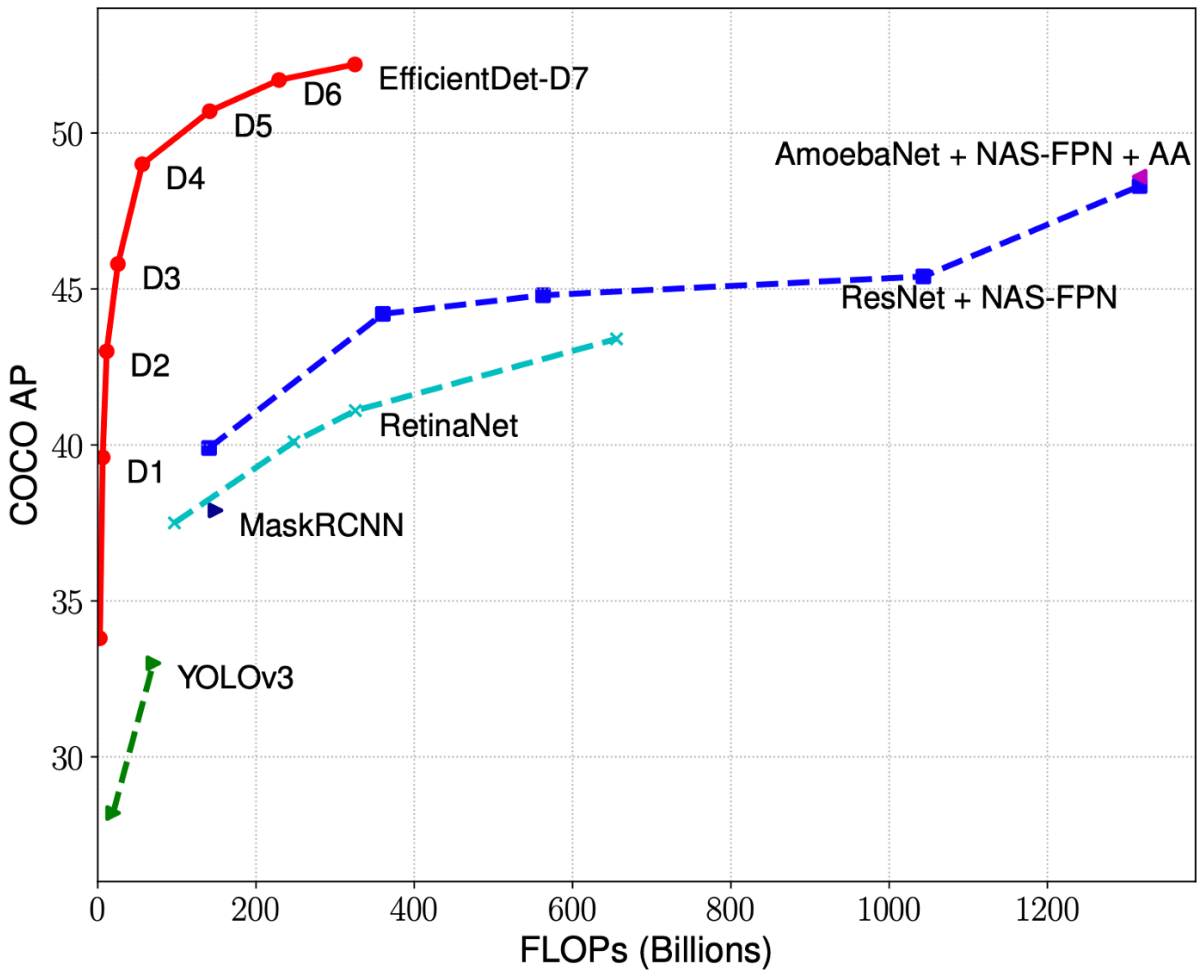


Figure 6.14 EfficientDet FLOPs and COCO AP Graph (Tan, Pang, & Le, 2020)

The graph showed that the EfficientDet from D1 to D7 is much more accurate than the YOLOV3 but from D2 onwards required more computational power for detecting, but only D1 to D4 are suitable for mobile devices where D4 and onwards are more suitable for devices with high computational power like PC, Machine etc.

Figure 6.15 Comparison Training Time of YOLOV3 and EfficientDet (Jacob, S.,2020)

The graph above shows the training time required by each of the model, where YOLO required more time overall to train the model and make it able to use in real environments. Therefore, to implement a complete YOLO model, it requires a lot of time to train a mature model compared to EffiecientDet.

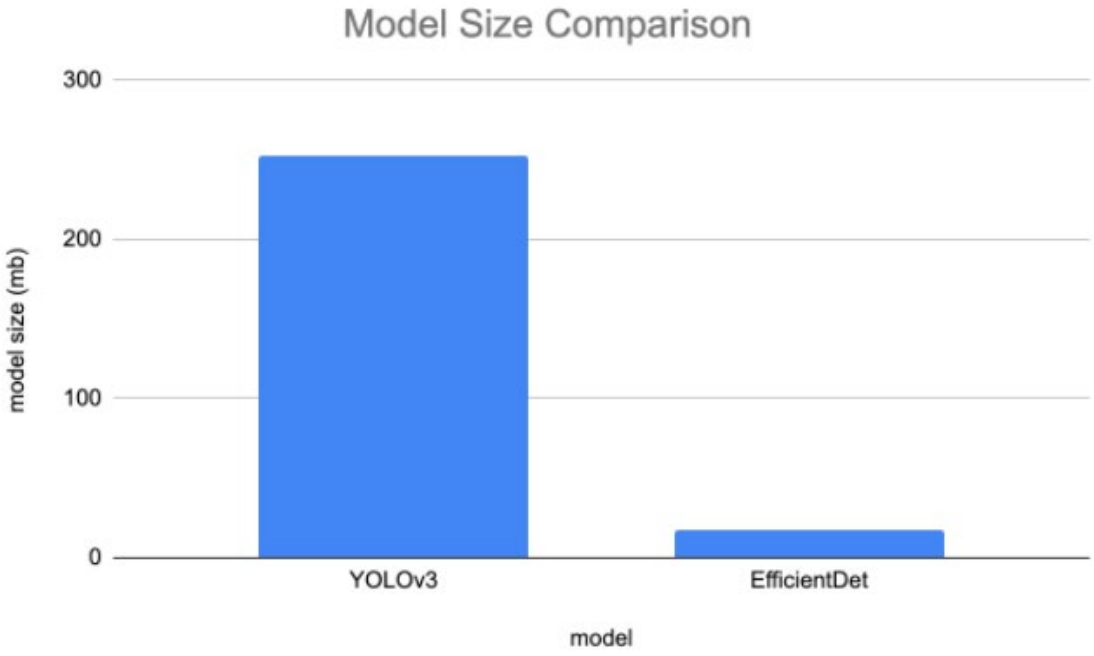


Figure 6.16 Comparison Model Size of YOLOV3 and EfficientDet (Jacob, S.,2020)

The graph above shows the model size after training, the YOLOv3 is significantly bigger size than EfficientDet, therefore, it needs more computational power and space to implement the model into applications and make it working in applications.

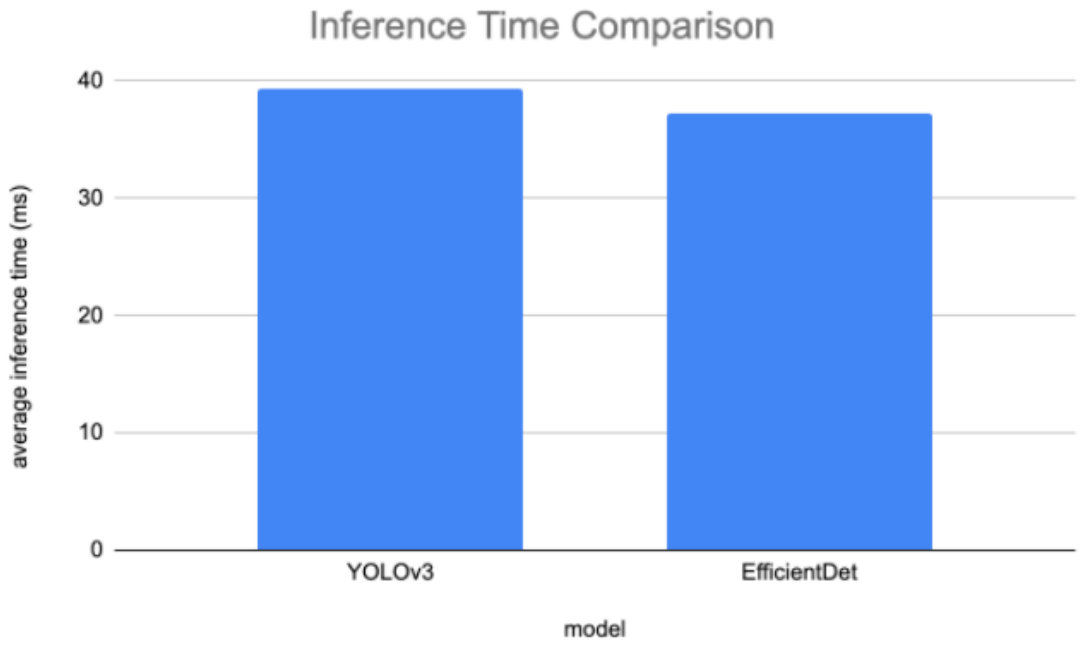


Figure 6.17 Comparison Inference Time of YOLOV3 and EfficientDet (Jacob, S.,2020)

Even with higher training time and bigger size model, the YOLOv3 still required longer inference time compared to EfficientDet, even it doesn't affect a lot in real-time performance, but it indicates that EfficientDet provide almost same accurate results with shorter time compared to YOLOv3 refer to the graph below. It showed that the accuracy of both model is doing well but the time needed to detect the object required by EfficientDet is lower.

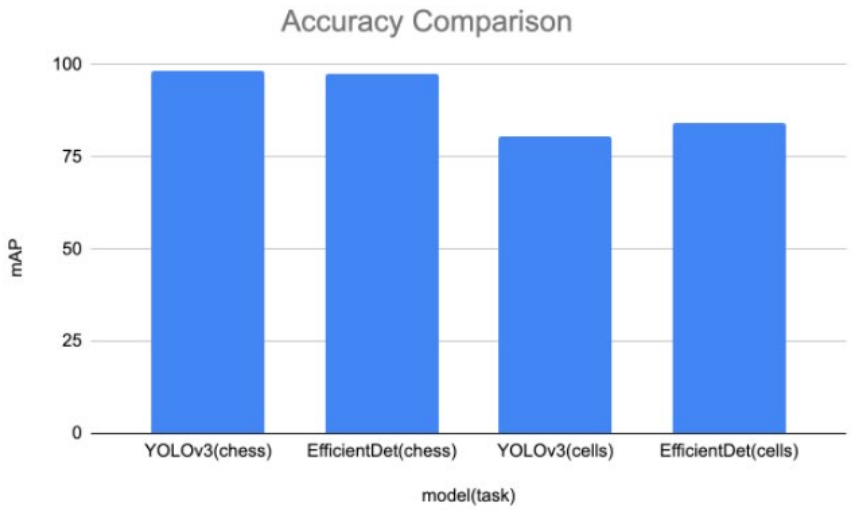


Figure 6.18 Comparison Accuracy of YOLOV3 and EfficientDet (Jacob, S.,2020)

Besides, with all the implementation into the AsSee application, there are several points that make sure

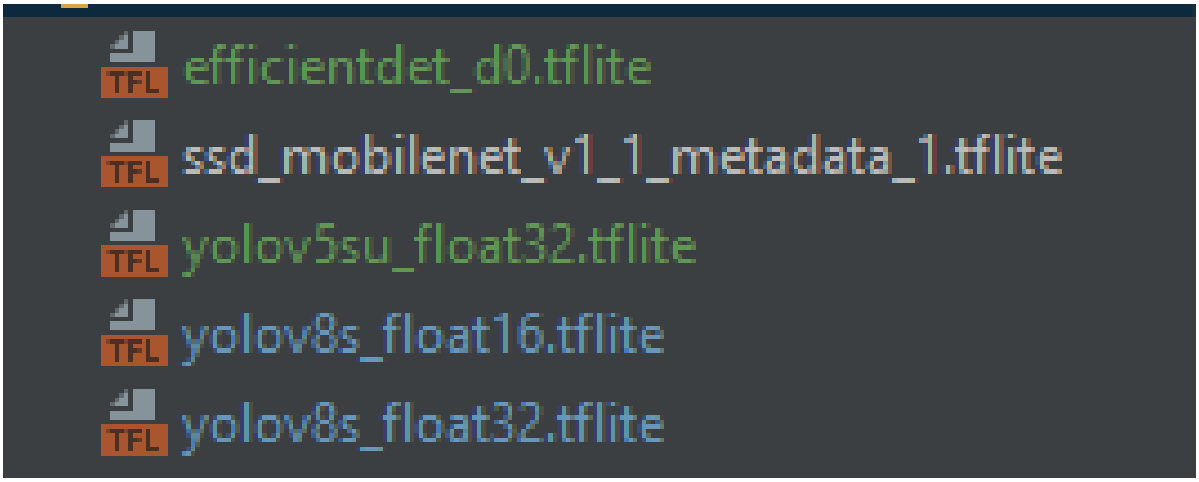


Figure 6.19 Implementation of Different Models in Android Studio

Throughout the implementation, YOLO model it is not very easy to implement into the android application, and it require more computational power where may not afford by every device, and sometimes it will burst out the memory

and computational power of the android devices, therefore YOLO is not selected for the application, due to the setup complexity and the performance. Furthermore, the SSD mobile net and EfficientDet is easier to implement, and it is much easier to use. Besides, the performance is much better than YOLO, YOLO performance is much better in devices with more computational power but due to research and own implementation, it is slightly low perform compared to EfficientDet and SSD mobile net in mobile devices. In addition, the SSD mobile net performance is not consistent as EfficientDet and required more stable situations to perform well, even with slightly difference, but the EfficientDet was selected and implemented into this application. Furthermore, the YOLO required more complex functions and methods to make it work with the application, therefore the processing time might have slightly longer than other model and with more resources needed, the YOLO implementation will slow down the application.

6.7 Justification for Text Recognition

Text Recognition also known as Optical Character Recognition (OCR) where is the process that extract text and information from image and converts it into a machine-readable text format. The OCR can ensure the process of reading information in business processes like forms, receipts, invoices etc it is more fast and more accurate and extract important information from those documents and convert it into paperless mode where paperwork waste a lot of time and consume a lot of resources. Automating this process ensure the efficiency and productivity.

The OCR convert image into machine-readable text with these following steps, image acquisition, preprocessing, text recognition, layout recognition and postprocessing. There are a lot of industries implemented the OCR Text Recognition like healthcare industry to extract medical records, finance industry extract cheque and invoice information and retailers use it to inventory management. Therefore, it helps to save time and resources with lower error rate compared to human and when it comes to huge amounts of data, the performance of the OCR is much more consistent and perfect than human eyes and brain and even shorter the duration of data entry into system.

The correct and high performance of OCR implementation in AsSee application is crucial, due to the information extract from the document is important to the user of the application because it helps them to understand documents and read text that is not in braille using AsSee applications. Therefore, there are research and comparison conducted and referred to select the best fit OCR technology to implement into AsSee applications to ensure the performance of the application.

For the text recognition model, there are research and text conducted to select the best to implement into the application to ensure the accuracy of the result and improve the quality of assistant that provided to user. The research conducted on different use case and included different situations where include the following use case where covered the daily life common text or logo that needs to be recognize.

Figure 6.20 Different Usage of Text Recognition

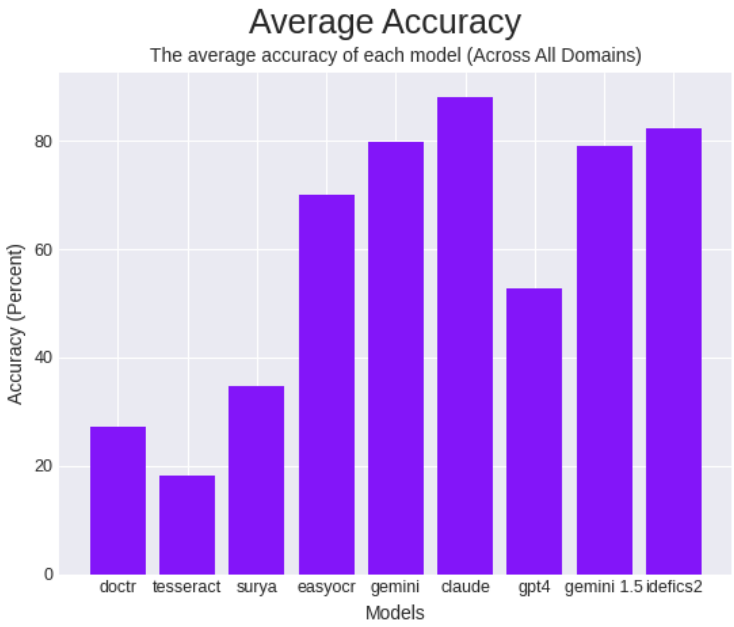


Figure 6.21 Comparison of Average Accuracy of Text Recognition Models (L,Ueno.,2024)

The graph above shows that there are several AI models that had been conduct test to show the accuracy of the model to recognize the text and extract information from the document. Which accuracy is always the first to consider when selecting an AI model to be implemented into application. From the graph we can know that the highest accuracy which is claude model where with more than 80% of accuracy which is having the best performance in accuracy where the tesseract model only has less than 20% of the accuracy.

The accuracy is not the only to consider in AsSee application but speed also an important aspect to concern to fulfil the real-time performance, if the speed was slow and time taken to process the image, it is crucial to the application function and assistant that provide to the user.

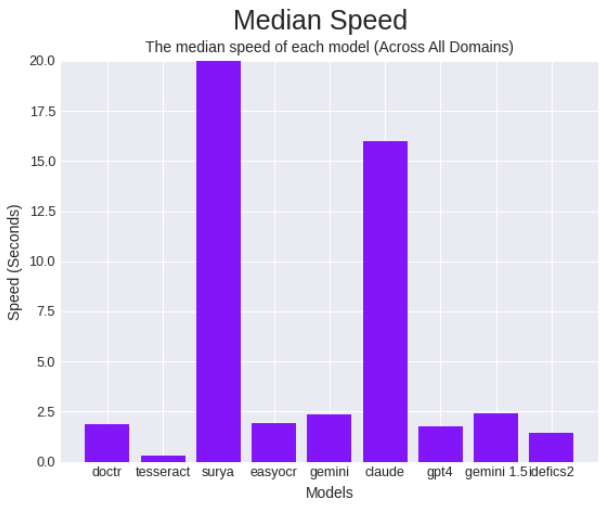


Figure 6.22 Comparison of Median Speed of Text Recognition Models (L,Ueno.,2024)

The graph above shows the median speed that taken of the model in processing image and return the result. Furthermore, the longer speed needed by the model to process image means the real-time requirement is not fulfilled, therefore the model needs longer time was not suitable for the AsSee application. In addition, the graph below shows the performance and relationship of the speed and accuracy of the model. Based on the graph, where the model Surya have high accuracy but required longer time to process the image, but the claude need shorter time to process but with even higher or same accuracy. The performance is uneven for every AI model.

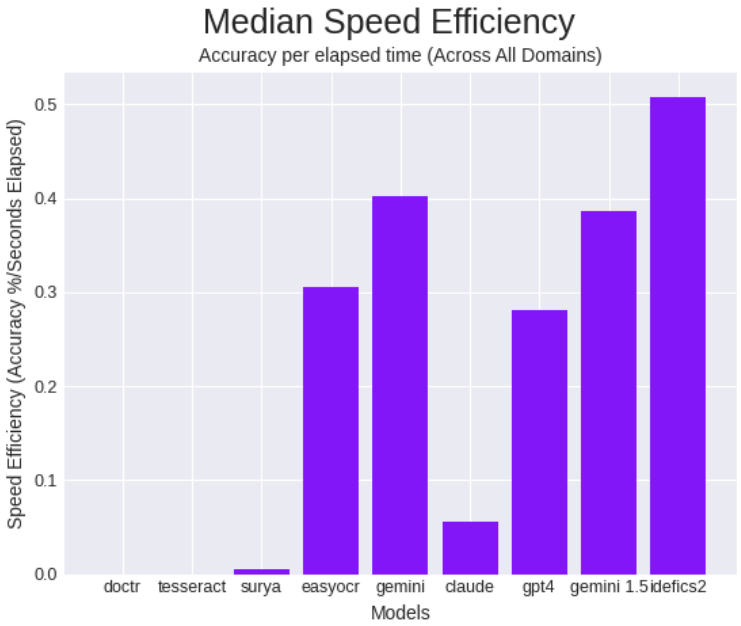


Figure 6.23 Comparison of Median Speed Efficiency of Text Recognition Models (L,Ueno.,2024)

From the graph above, it shows the most effective AI model is idefics2 and follow with gemini, gemini 1.5, easyocr, gpt4 etc. The efficiency is the most important aspect to be considered because it represents the speed and accuracy of the AI model and how effective of the AI model.

For AsSee, there are still one aspect needed to be considered which is the cost, the AsSee application is non-profit application and aim to provide services to the user with no charges. Therefore, the cost is also the importance aspect to affect the selection of the AI model.

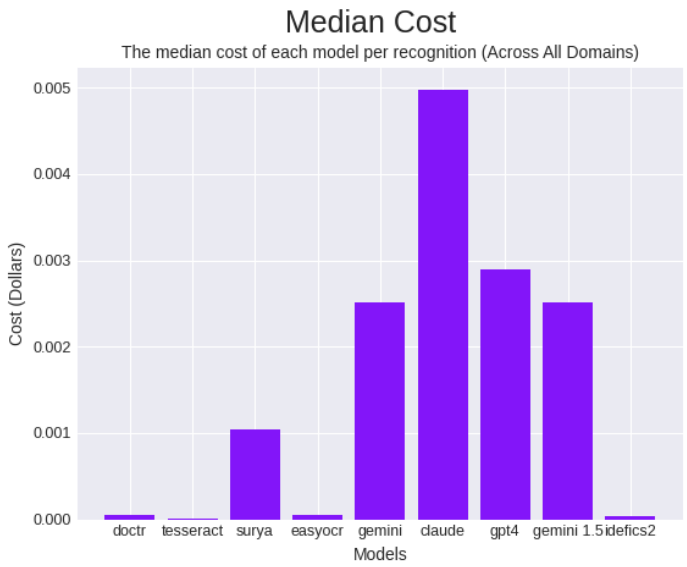


Figure 6.24 Comparison of Median Cost of Text Recognition Models (L,Ueno.,2024)

The cost of the AI models was in dollars and the claude which have the best performance in accuracy have the most expensive price. And there are some few open-source and free option but with good performance in efficiency like idefics2 model.

Therefore, the idefics2 model seem like the best choice to be implement into the application but the idefics2 model is not compatible and not convertible to model that compatible with Android system therefore, even with best performance and efficiency, idefics2 model was not selected to implement into AsSee application.

6.7.1 Comparison between Tesseract vs EasyOCR vs ML Kit Text Recognition API

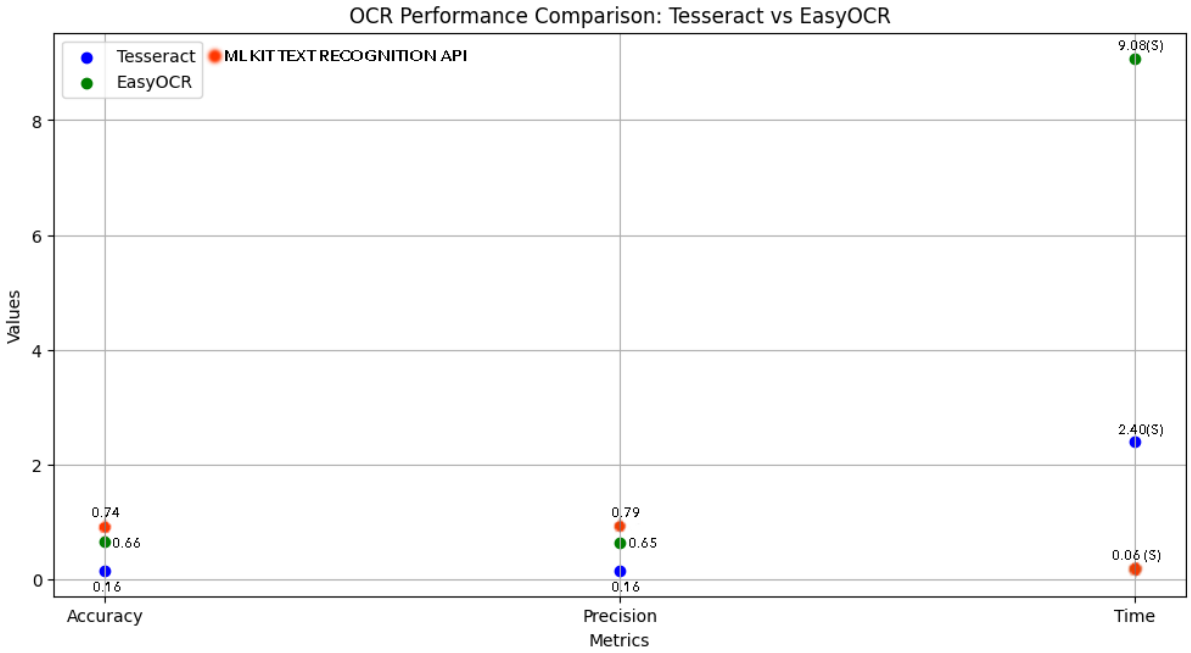


Figure 6.25 Comparison of Tesseract vs EasyOCR vs ML Kit Text Recognition API

The dataset that used for testing is ICDAR 2013 test and the environment was set in Kaggle to test the AI model that compatible with Android and Android-friendly AI model which is the tesseract and EasyOCR to compare with the ML Kit Text Recognition. Where the results show the research that referred to is creditable because the result was nearly identical.

Conclusion, the methodology that implemented into the application is ML Kit Text Recognition API due to it is free, with higher accuracy and faster speed compared to tesseract and EasyOCR the performance of the ML Kit Text Recognition API was more efficient and more accurate with highest speed and believe the speed was fastest among all the model in the research and self-setup environment even with slightly low accuracy but able to handle command document in daily life and the result was precise enough and high credibility.

6.7.2 Justification for ML Kit Text Recognition

ML Kit Text Recognition API is the mobile SDK text recognition that developed and introduced by Google. ML Kit Text Recognition can recognize

text in Latin-based character set. Google development teams doesn't have any academic work published to evaluate and introduce the ML Kit Text Recognition API.

The ML Kit Text Recognition API will take image as input and output with the recognized text.

Feature	Unbundled	Bundled
Library name	com.google.android.gms:play-services-mlkit-text-recognition com.google.android.gms:play-services-mlkit-text-recognition-chinese com.google.android.gms:play-services-mlkit-text-recognition-devanagari com.google.android.gms:play-services-mlkit-text-recognition-japanese com.google.android.gms:play-services-mlkit-text-recognition-korean	com.google.mlkit:text-recognition com.google.mlkit:text-recognition-chinese com.google.mlkit:text-recognition-devanagari com.google.mlkit:text-recognition-japanese com.google.mlkit:text-recognition-korean
Implementation	Model is dynamically downloaded via Google Play Services.	Model is statically linked to your app at build time.
App size	About 260 KB size increase per script architecture.	About 4 MB size increase per script per architecture.
Initialization time	Might have to wait for model to download before first use.	Model is available immediately.
Performance	Real-time on most devices for Latin script library, slower for others.	Real-time on most devices for Latin script library, slower for others.

Figure 6.26 Screenshot of ML Kit Text Recognition API Library

The google play services runtime implementing TensorFlow Lite, this enables android device to run machine learning models from TensorFlow Lite without statically bundling TensorFlow Lite libraries into the application. Therefore, with these advantages, it doesn't require manually setup and install the models into the application which significantly increase the size of the application but with high performance to ensure the accuracy and efficiency of the application.

The ML Kit Text Recognition API combine both Convolutional Neural Networks (CNNs) and Recurrent Neural Network (RNNs) which is the CRNNs where to recognize text based on an image. There are different stages for ML Kit Text Recognition API to recognize text from image, there are transformation, extraction, sequence modelling and prediction.

The reason why selecting ML Kit Text Recognition API is because it is free to implement into the application, and without any complex setup required to make it work in the application. Furthermore, it does not require

any additional models install into the application which may increase the size and make the size of the application bigger. In addition, the real-time performance which is the crucial metrics to evaluate the text recognition to support the needs of the user and with the higher accuracy compared to the model which is android friendly.

6.8 Justification for Speech Recognition

Speech recognition, also known as automatic speech recognition (ASR), computer speech recognition or speech-to-text, is a capability that enables a program to process human speech into a written format. (What Is Speech Recognition? | IBM, n.d.)

There are multiple key features of the speech recognition, the accuracy and speed which determine the speech recognition can perform real-time and respond to user inputs. Besides, the NLU which represent Natural Language Understanding enables system to handle complex commands and queries which make the system more user-friendly. Furthermore, the supportive for the multiple languages and dialect where allow different users to communicate with the system with speech recognition, for example, there are different dialect and different pronunciation for specific languages, with the support of different speech recognition make sure everyone can interact and communicate with the system. Lastly, the Background noise Handling where also a critical feature of the speech recognition, where enable user to communicate with the system under a noisy background and able to capture the command precisely and correctly and filter the noise.

There is different usage of the speech recognition which well-known virtual assistant is the most popular usage of speech recognition, and to ensure the application also accessible by people with disabilities, speech recognition play an important role in this aspect. Furthermore, the automotive system where integrated speech recognition to accept command from user to do corresponding action and there are still a lot of implementations of speech recognition to make the software and device more accessible and responsive to command.

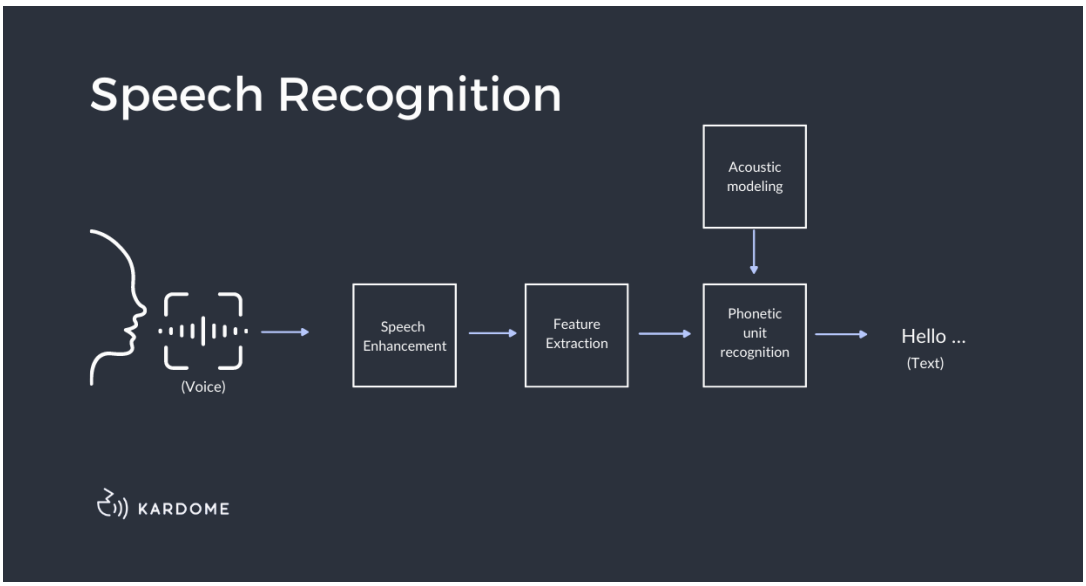


Figure 6.27 Speech Recognition Workflow

6.8.1 Justification for Vosk Speech Recognition Model

There are a lot of different well-known ASR AI Model like Kaldi, DeepSpeech, Vosk, Wav2Letter++, OpenSeq2Seq, Whisper, Wav2Vec and each of them got their own pros and cons, after comparison in different aspects, Vosk, Wav2Vec2 and Whisper Model was selected for the comparison in Word Error Rate (WER) and Time Taken Comparison for select appropriate model to be implemented into AsSee.

Table 8 Comparison of Speech Recognition Models

	Kaldi	DeepSpeech	Vosk	Wav2Letter++
Pros	- Customizable - State of Art Performance - Lots of Features - Active Community	- Easy To use - Pre-built Models - Community Support - Cross Platform	- Low Resource - Real-Time Processing -Multi-Language -Easy To Integrate	- Fast - End-to-End - Extensible - GPU Acceleration
Cons	- Steep Learning Curve - Resource Hungry	- Accuracy Variability - Limited Language - Recording Limitations	- Accuracy -No advanced features	- Hard to Setup - Resource Hungry -Limited Community

Table 9 Comparison of Speech Recognition Models 2

	OpenSeq2Seq	Whisper	Wav2Vec
Pros	- Versatile - Fast - Pre-trained Models - Documentation	- Accuracy - Language Support - Easy To Use - Advanced Features	- Self Supervised - Accurate - Scalable - Tunable
Cons	- Complex - Resource Hungry - Not Speech Specific	- Resource Hungry - Not Customizable	- Hard to setup - Resource Hungry - Not Real-Time

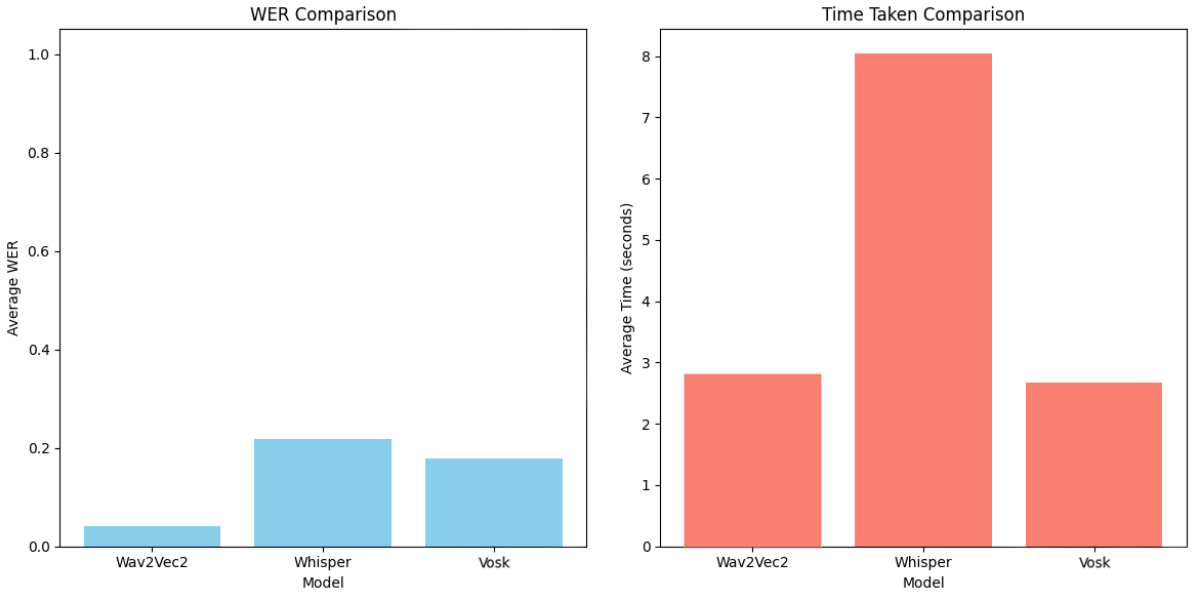


Figure 6.28 Comparison of Wav2Vec2 vs Whisper Model vs Vosk

Through out the comparison of different model, Wav2Vec, Whisper Model and Vosk was selected for the comparison, because there are ready model for Android which means them are Android Friendly, the whisper provided whisper-tiny.tflite, and Vosk with vosk-small-en-us-0.15 that is user friendly, and Wav2Vec with ready TensorFlow model that able to be convert to tflite and implement into Android, therefore, the resources needed to run the model is lesser which enable Android devices to run it smoothly completely. Therefore, these three models were selected for the comparison. The dataset that selected is LibriSpeech and a small module under the LibriSpeech.

Based on the graph above, the WER Comparison is the Wav2Vec have the best performance where only with 0.04 which means the accuracy is around 96% with the highest performance of the different models, followed by Vosk with WER around 0.18 which represent accuracy up to 82% and Whisper with 0.22 WER where accuracy is around 78%. Furthermore, the graph above shows the average time taken of the models to process, the performance of Wav2Vec2 was like Vosk, but Vosk was slightly better in average time taken which is 2.8 seconds and Wav2Vec is 2.9 seconds in average, and the one with worst performance is Whisper model is 8 seconds. After the comparison in different aspect, the Vosk was selected to be implemented into the AsSee application.

The reason why the Vosk was selected to be implemented into AsSee application is because the resources needed is much lesser than other which represents the performance of Vosk will be much better than others in Android Native and support more devices to run the models, furthermore, the different languages support by Vosk model which enable future development that support different language command from different user dialect. In addition, the lightweight and fast performance of Vosk which may slightly lowly perform than Wav2Vec2 but the offline capability of the Vosk which it can be performed without internet connection, but Wav2Vec2 required internet connection to perform recognition therefore, to ensure that the application still working under different environments like without internet connection, Vosk was better option to be implemented into the AsSee application. Besides, the customization of different model for Vosk where it enables to customize the dictionary and train own model for different usage. Besides, the real-time performance in Android Native is slightly better than other models. In different aspects and consideration, Vosk was selected for the implementation into the AsSee application.

CHAPTER 7

SYSTEM TESTING

7.1 Introduction

In this chapter, the system testing will be discussed, and testing was conducted to ensure that the project meets the requirements and project objectives with proper function.

7.2 Unit Testing

Table 10 Unit Testing for AsSee 1

Test Case ID	TC-001	Module Name	Camera Handler	
Test Title	Test Case for Camera Handler			
Pre-Condition	Device Camera Working			
Test Case Description	Execution Steps	Expected Result	Actual Result	Status
Open Application	1. Open Application	The application will show frame captured by the camera	The application show frame captured by the camera	Pass

Table 11 Unit Testing for AsSee 2

Test Case ID	TC-002	Module Name	Object Detection	
Test Title	Test Case for Object Detection			
Pre-Condition	1. Device Camera Working 2. Detect Object is covered by Model 3. Device Speaker Working			
Test Case Description	Execution Steps	Expected Result	Actual Result	Status
1.)Object Detection Function	1. Open Application 2. Switch to Object Detection Mode 3. Point Camera toward Object would like to be recognize 4. Application speak the information of detected object	The application will draw bounding boxes of object detected and speak information of detected object	The application draw bounding boxes of object detected and speak information of detected object	Pass

Table 12 Unit Testing for AsSee 3

Test Case ID	TC-003	Module Name	Text Recognition	
Test Title	Test Case for Text Recognition			
Pre-Condition	1. Device Camera Working 2. Device Speaker Working			
Test Case Description	Execution Steps	Expected Result	Actual Result	Status
1.)Text Recognition Function	1. Open Application 2. Switch to Text Recognition Mode 3. Point Camera toward text would like to be recognize 4. Long press the screen to trigger the text recognition function 5. Application speak the recognized text	The application will speak recognized text	The application speak recognized text	Pass
2.)Text Recognition Function	1. Open Application 2. Switch to Text Recognition Mode 3. Point Camera toward anywhere with no text 4. Long press the screen to trigger the text recognition function 5. Application speak No Text Recognized	The application will speak No Text Recognized	The application speak No Text Recognized	Pass

Table 13 Unit Testing for AsSee 4

Test Case ID	TC-004	Module Name	Voice Command Handler	
Test Title	Test Case for Voice Command Handler			
Pre-Condition	1. Device Microphone Working 2. Device Speaker Working			
Test Case Description	Execution Steps	Expected Result	Actual Result	Status
1.)Change mode to object detection	1. Open Application 2. Wake up the Voice Assistant with button 3. Speak command "Object" 4. Application made announcement "Mode changed to Object Detection"	The application will change to object detection mode with announcement	The application changed to object detection mode with announcement made	Pass
2.)Change mode to text recognition	1. Open Application 2. Wake up the Voice Assistant with button 3. Speak command "Text" 4. Application made announcement "Mode changed to Text Recognition"	The application will change to object detection mode with announcement	The application changed to object detection mode with announcement made	Pass
3.)Wrong Command Made	1. Open Application 2. Wake up the Voice Assistant with button 3. Speak anything other than "Object" and "Text" or no speaking 4. Application made announcement "Command Not Recognized"	The application will make announcement "Command No Recognized"	The application made announcement "Command No Recognized"	Pass

7.3 User Acceptance Test

Table 14 User Acceptance Test for AsSee

Test Case ID	Description	Expected Outcome	Pass/Fail	User Comment	Rating (1-5)
TC001	Application launches successfully	The application should open without errors			
TC002	Camera functionality	The application should open without errors			
TC003	Object detection	The application should detect objects accurately			
TC004	Text recognition	The application should recognize and read text correctly			
TC005	Text-to-speech functionality	The application should read the recognized text aloud			
TC006	User interface clarity	The UI should be clear and easy to navigate			
TC007	Overall performance	The application should run smoothly without significant lag			
TC008	Voice Command to Switch Mode	The application was able to switch between different mode with voice command			

Instructions for Users

1. Pass / Fail : Indicate if the test case passed or failed.
2. User Comments : Provide any additional comments or observations.
3. Rating (1-5) : Rate the overall experience of the specific feature or test case, where 1 is very poor and 5 is excellent.

7.4 System Usability Scale

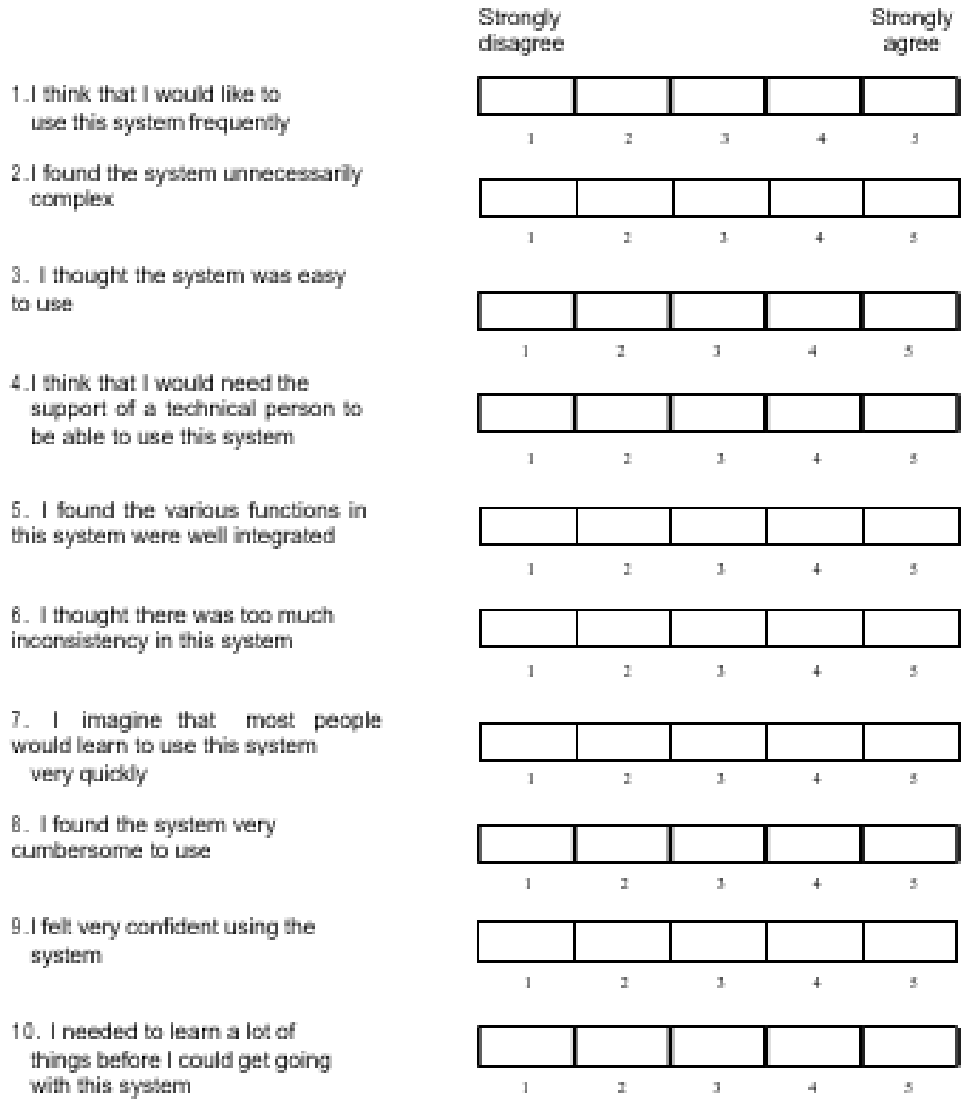


Figure 7.1 System Usability Scale for AsSee

This is the SUS, SUS was released by the John Brooke in 1986 where consist of 10 questions with 5 scale of answer. It have own special unique scoring system. There are few rules and special calculations for the SUS.

i. For odd-numbered questions: substract one from the user response. For example, question scale 1 is 5, then 5 – 1 = 4 for question 1.

ii. For even-numbered questions: subtract user response from 5. For example, question 2 scale is 2, then 5 – 2 = 3 for question 2.

iii. This ensure it convert every question scale from 0 to 4, where 4 represents the most positivite response.

iv. After convert every question scale to the new scale, then multiply by 2.5 to get the final score where the range from 0 to 100 instead of only 0 to 40.

Table 15 System Usability Scale Grading System

SUS Score	Grade	Adjective Rating
>80.3	A	Excellent
68-80.2	B	Good
67	C	Okay
51-66	D	Poor
<51	F	Awful

By implement the SUS, from the score from the SUS, it can evaluate and rate the project. If the score is 80.3 or higher, means the project is loved by the user and it will be recommend to other people, beside, if the score is under 51, then a quick fix and improvement needs to be implemented as soon as possible to make the project usable and perform well and acceptable by user.

7.5 Test Guide

Precautions:

i.) You need to wait for awhile for application to accept next command

ii.) You need to wait for the application complete current action and operation to accept next command

iii.) You need to grant access to camera and microphone to the application to make sure it works well.

iv.) Text recognition supports only English

Testing Procedure:

1.) Say '**Text**' To change to Text Recognition Mode
   a. Switched to Text Recognition Mode alert by application
   b. Point your camera toward the text\paragraph\photo\document that contains text you want to recognize and long press on screen
      i. Can Use Paragraph in Page 2
   c. Application will recognize the text and read out.

2.) Say '**Object'** Change to Object Detection Mode

a. Switched to Object Detection Mode alert by application

b. Pointing your camera toward the common object in daily life it will start detecting and return necessary information to you.

7.5.1 Summary

7.5.1.1 Summary of UAT

To summarize the test from 3 person invited for the testing of the AsSee.

For the User Acceptance Test (UAT), there are 3 respond.Where the average marks is (40 + 40 + 39) / 3 = 39.6667 marks for the User Acceptance Test where indicates the application fulfil the user requirements and user satisfaction is near to perfect. There for the average mark for UAT shows the AsSee application is acceptable and meets user expectations and the functionality fulfilled the customer.

7.5.1.2 Summary of SUS

To summarize the software usability scale (SUS), there are 3 different respond received.

I. Feedback 1

Odd-Numbered Questions (1, 3, 5, 7, 9)	For these questions, subtract 1 from the response:	Question 1: 4 - 1 = 3 Question 3: 5 - 1 = 4 Question 5: 4 - 1 = 3 Question 7: 5 - 1 = 4 Question 9: 5 - 1 = 4	3 + 4 + 3 + 4 + 4 = 18
Even-Numbered Questions (2, 4, 6, 8, 10):	For these questions,subtract the response from 5:	Question 2: 5 - 1 = 4 Question 4: 5 - 1 = 4 Question 6: 5 - 2 = 3 Question 8: 5 - 1 = 4 Question 10: 5 - 1 = 4	4 + 4 + 3 + 4 + 4 = 19
Total			18 + 19 = 37
Total SUS (Total x2.5)			37 × 2.5 = 92.5 Final SUS Score = **92.5**

Figure 7.2 Feedback 1 for SUS

II. Feedback 2

Odd-Numbered Questions (1, 3, 5, 7, 9)	For these questions, subtract 1 from the response:	Question 1: 5 - 1 = 4 Question 3: 5 - 1 = 4 Question 5: 5 - 1 = 4 Question 7: 5 - 1 = 4 Question 9: 5 - 1 = 4	4 + 4 + 4 + 4 + 4 = 20
Even-Numbered Questions (2, 4, 6, 8, 10):	For these questions,subtract the response from 5:	Question 2: 5 - 1 = 4 Question 4: 5 - 1 = 4 Question 6: 5 - 1 = 4 Question 8: 5 - 1 = 4 Question 10: 5 - 1 = 4	4 + 4 + 4 + 4 + 4 = 20
Total			20 + 20 = 40
Total SUS (Total x2.5)			40 × 2.5 = 100 Final SUS Score = **100**

Figure 7.3 Feedback 2 for SUS

III. Feedback 3

Odd-Numbered Questions (1, 3, 5, 7, 9)	For these questions, subtract 1 from the response:	Question 1: 4 - 1 = 3 Question 3: 5 - 1 = 4 Question 5: 5 - 1 = 4 Question 7: 5 - 1 = 4 Question 9: 5 - 1 = 4	3 + 4 + 4 + 4 + 4 = 19
Even-Numbered Questions (2, 4, 6, 8, 10):	For these questions,subtract the response from 5:	Question 2: 5 - 1 = 4 Question 4: 5 - 2 = 3 Question 6: 5 - 1 = 4 Question 8: 5 - 1 = 4 Question 10: 5 - 1 = 4	4 + 3 + 4 + 4 + 4 = 19
Total			19 + 19 = 38
Total SUS (Total x2.5)			38 × 2.5 = 95 Final SUS Score = **95**

Figure 7.4 Feedback 3 for SUS

Where the average SUS score is (95+100+92.5)/3 = 95.83, therefore, it indicates it is grade A and excellent application and it also means that the user would like to recommend the application to other person and it is usable.

CHAPTER 8

CONCLUSION

8.1 Introduction

In this chapter, conclusion and future project expectations discussed.

8.2 Challenges

This section discusses the challenges and problems encountered during the journey of the project.

i. Self pickup the Kotlin Programming Language.
ii. Self pickup the knowledge regarding different AI Models in different fields.
iii. Insufficient resources and time to train the model.

To encounter the challenges faced, to pickup the Kotlin programming there are a lot of tutorials and code refer to faster the process of learning, and due to the similarities of the Kotlin and Java, it makes it more easier to pickup. To have better understanding of the different AI Models, there are a lot of research have to be done to understand better and select the most appropriate AI Models to implement into the application. To encounter the problems of the insufficient and time to train the model, the pre-trained model was selected to implement into the application.

8.3 Future Enhancement

i. Develop support for multilingual speech recognition to accommodate various dialects.
ii. Integrate with smartwatches for continuous health monitoring and status updates.
iii. Implement location tracking features to enhance navigation and ensure user safety.
iv. Incorporate scene description functionality to provide users with a better understanding of their surroundings.
v. Implement with AI Model that can identify more object and classes.

As stated above, the limitations will be mitigated in future development for enhancement to make the application more perfect.

8.4 Conclusion

During the journey of this project, the development of this project and researching regarding the projects benefits me a lot and lessons learned could ultilized for upcoming projects and journey of life.

The project considered success and with all the research and development that comeout with a deliverables that satisfy and achieved the objectives. There are different models that implement after several research and comparison, with all the results referred to, the AI models selected for the project is TensorFlow Lite EfficientDet Object Detection AI model, Vosk speech recognition AI model and ML Kit Text Recognition AI. With established the connection between different AI model and switch different AI model in different situations assist them with different functions. With the application, there are assistance provided to the user in specific situations which is recognizing the object and with position information, besides recognizing the text when needed and make command to the application to switch between different mode with button or voice commands.

As mentioned in the introduction, there are several difficulties facing by visual impaired individuals in daily routines, therefore with the AsSee application, there are some situations in daily life like provide assistance for them to access the information from the written format information, besides, it provides the object detection for them to have better navigation and mobility. In addition, believe the application slightly arise the awareness of existing technologies to assist individuals with disabilities. Therefore, with the safety level increased and assistsanec provided to the user, the objective was achieved.

Last but not least, there are a lot of lesson learnt in this project. With all the guidance and comment from supervisors and moderator, it enhances the development process and there are a lot of guidance that could be applied to incoming project and life.

REFERENCES

Abrar Al-Heeti, 2023, *Google's Lookout App Adds Detailed, AI-Powered Image Descriptions - CNET* [Online]. Available at: https://www.cnet.com/tech/mobile/googles-lookout-app-adds-detailed-ai-powered-image-descriptions/ [Accessed: 16 April 2024].

AFB, *Electronic Notetakers (Braille) | American Foundation for the Blind* [Online]. Available at: https://www.afb.org/assistive-technology-products/braille-notetakers [Accessed: 17 April 2024].

Alexandra, 2024, *What Is SDLC? Understand the Software Development Life Cycle - Stackify* [Online]. Available at: https://stackify.com/what-is-sdlc/ [Accessed: 17 April 2024].

Anon, *Flutter App Development - Why You Should Choose Flutter* [Online]. Available at: https://www.appify.digital/post/flutter-app-development [Accessed: 16 April 2024a].

Anon, *Using Text-to-Speech | Android Developers* [Online]. Available at: https://www.dre.vanderbilt.edu/~schmidt/android/android-4.0/out/target/common/docs/doc-comment-check/resources/articles/tts.html [Accessed: 17 April 2024b].

Anon, *What is Flutter? - Flutter App Explained - AWS* [Online]. Available at: https://aws.amazon.com/what-is/flutter/ [Accessed: 16 April 2024c].

Anon, *What is Object Detection? | IBM* [Online]. Available at: https://www.ibm.com/topics/object-detection [Accessed: 21 August 2024d].

Anon, *What Is Speech Recognition? | IBM* [Online]. Available at: https://www.ibm.com/topics/speech-recognition [Accessed: 23 August 2024e].

AWS, *What is SDLC? - Software Development Lifecycle Explained - AWS* [Online]. Available at: https://aws.amazon.com/what-is/sdlc/ [Accessed: 17 April 2024].

Battle For Blindness, *Overcoming challenges in daily life as a blind person - Battle for Blindness* [Online]. Available at: https://battleforblindness.org/overcoming-challenges-in-daily-life-as-a-blind-person/ [Accessed: 16 April 2024].

Bhagavatiprasad Vaghela, 2023, *Benefits of Reactive Native | IEEE Computer Society* [Online]. Available at: https://www.computer.org/publications/tech-news/trends/benefits-of-react-native [Accessed: 16 April 2024].

Boldu, R. et al., 2018. FingerReader2.0. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 2(3), pp.1–19. Available at: https://dl.acm.org/doi/10.1145/3264904 [Accessed: 17 April 2024].

CAN, *Features – CAN* [Online]. Available at: https://can.co/pages/features [Accessed: 17 April 2024].

Carmen Willings, 2024, *Braille Technology for Students who are Blind or Visually Impaired - Teaching Students with Visual Impairments* [Online]. Available at: https://www.teachingvisuallyimpaired.com/braille-technology.html [Accessed: 17 April 2024].

Chris Raroque, *9+ Reasons Why Python Is A Popular Programming Language* [Online]. Available at: https://aloa.co/blog/why-python [Accessed: 17 April 2024].

COCO, *COCO - Common Objects in Context* [Online]. Available at: https://cocodataset.org/#home [Accessed: 18 April 2024].

Common Voice, *Common Voice* [Online]. Available at: https://commonvoice.mozilla.org/en [Accessed: 18 April 2024].

Envision, 2024, *Challenges blind people face when living life* [Online]. Available at: https://www.letsenvision.com/blog/challenges-blind-people-face-when-living-life [Accessed: 16 April 2024].

Fable, *Braille Note Taker | What is a Braille Note Taker? | Fable* [Online]. Available at: https://makeitfable.com/glossary-term/braille-note-taker/ [Accessed: 17 April 2024].

Gaudenz Boesch, 2024, *TensorFlow Lite - Computer Vision on Edge Devices [2024 Guide] - viso.ai* [Online]. Available at: https://viso.ai/edge-ai/tensorflow-lite/ [Accessed: 17 April 2024].

GeeksforGeeks, 2019, *Dynamic Systems Development Method (DSDM) - GeeksforGeeks* [Online]. Available at: https://www.geeksforgeeks.org/dynamic-systems-development-method-dsdm/ [Accessed: 17 April 2024].

GeeksforGeeks, 2023, *Lean Software Development (LSD) - GeeksforGeeks* [Online]. Available at: https://www.geeksforgeeks.org/lean-software-development-lsd/ [Accessed: 17 April 2024].

GeeksforGeeks, 2024a, *Prototyping Model - Software Engineering - GeeksforGeeks* [Online]. Available at: https://www.geeksforgeeks.org/software-engineering-prototyping-model/ [Accessed: 17 April 2024].

GeeksforGeeks, 2024b, *Software Development Life Cycle (SDLC) - GeeksforGeeks* [Online]. Available at: https://www.geeksforgeeks.org/software-development-life-cycle-sdlc/ [Accessed: 17 April 2024].

Google, *Use Lookout to explore your environment - Android Accessibility Help* [Online]. Available at: https://support.google.com/accessibility/android/answer/9031274?hl=en [Accessed: 16 April 2024].

Hossain, E., Qaiduzzaman, K.M. and Rahman, M., 2020. Sightless Helper: An Interactive Mobile Application for Blind Assistance and Safe Navigation. Available at: https://www.researchgate.net/publication/343284975 [Accessed: 16 April 2024].

IBM, *What is Computer Vision? | IBM* [Online]. Available at: https://www.ibm.com/topics/computer-vision [Accessed: 17 April 2024a].

IBM, *What Is Java? | IBM* [Online]. Available at: https://www.ibm.com/topics/java [Accessed: 16 April 2024b].

Indian Institute of Technology Delhi, *SmartCane* [Online]. Available at: https://assistech.iitd.ac.in/smartcane.php [Accessed: 17 April 2024].

InFocus, 2022, *Braille Technology For People with Visual Impairment | InFocus* [Online]. Available at: https://infocus-charity.org.uk/news-and-stories/blog/braille-technology-what-it-means-to-people-with-visual-impairment/ [Accessed: 17 April 2024].

Javatpoint, *Agile Model (Software Engineering) - javatpoint* [Online]. Available at: https://www.javatpoint.com/software-engineering-agile-model [Accessed: 17 April 2024].

kaggle, *iam_handwriting_word_database* [Online]. Available at: https://www.kaggle.com/datasets/nibinv23/iam-handwriting-word-database [Accessed: 18 April 2024].

Khneisser, E. et al., 2018. Intelligent eye: A mobile application for assisting blind people. Available at: https://www.researchgate.net/publication/325562479 [Accessed: 16 April 2024].

Laura Medcalf, 2021, *Supersense – AI app for individuals with visual impairments - Assistive Technology at Easter Seals Crossroads* [Online]. Available at: https://www.eastersealstech.com/2021/08/24/supersense-ai-app-for-individuals-with-visual-impairments/ [Accessed: 16 April 2024].

Ben Lutkevich, *What is Kotlin and Why Use it? | Definition from TechTarget* [Online]. Available at: https://www.techtarget.com/whatis/definition/Kotlin [Accessed: 16 April 2024].

MIT Fluid Interfaces Media Lab, *Finger Reader - Future Health Systems* [Online]. Available at: https://cargocollective.com/futurehealth/Finger-Reader [Accessed: 17 April 2024].

PySquad, 2024, *Vosk with Python: Future of Audio Processing with OpenSource Tools | by PySquad | Medium* [Online]. Available at: https://medium.com/@pysquad/vosk-with-python-future-of-audio-processing-with-opensource-tools-4868ef8c0f12 [Accessed: 10 September 2024].

Ravi Makhjia, 2023, *12 Reasons for Using Kotlin for Android App Development - Stackify* [Online]. Available at: https://stackify.com/12-reasons-for-using-kotlin-for-android-app-development/ [Accessed: 17 April 2024].

Shittu Olumide, 2023, *What is Java Used For in 2023? The Java Programming Language and Java Platform Strengths* [Online]. Available at: https://www.freecodecamp.org/news/what-is-java-used-for/ [Accessed: 16 April 2024].

Sivan, S. and Darsan, G., Computer Vision based Assistive Technology For Blind and Visually Impaired People. Available at: http://dx.doi.org/10.1145/2967878.2967923 [Accessed: 17 April 2024].

Supersense, *Supersense - AI for Blind / Scan text, money and objects* [Online]. Available at: https://www.supersense.app/ [Accessed: 16 April 2024].

TapTapSee, *TapTapSee - Blind and Visually Impaired Assistive Technology - powered by CloudSight.ai Image Recognition API* [Online]. Available at: https://taptapseeapp.com/instruction.html [Accessed: 16 April 2024].

The vOICe, 2024, *The vOICe - New frontiers In Artificial Vision* [Online]. Available at: https://www.seeingwithsound.com/ [Accessed: 17 April 2024].

Tian, Y., Yang, X., Yi, C. and Arditi, A., 2013. Toward a Computer Vision-based Wayfinding Aid for Blind Persons to Access Unfamiliar Indoor Environments. *Machine vision and applications*, 24(3), p.521. Available at: /pmc/articles/PMC3636776/ [Accessed: 17 April 2024].

TSBVI, *Braille Display Devices | Texas School for the Blind and Visually Impaired* [Online]. Available at: https://www.tsbvi.edu/statewide-resources/services/braille/display [Accessed: 17 April 2024].

WeWALK, *WeWALK Smart Cane – Smart Cane for the Visually Impaired* [Online]. Available at: https://wewalk.io/en/ [Accessed: 17 April 2024].

WHO, 2023, *Blindness and vision impairment* [Online]. Available at: https://www.who.int/news-room/fact-sheets/detail/blindness-and-visual-impairment [Accessed: 16 April 2024].

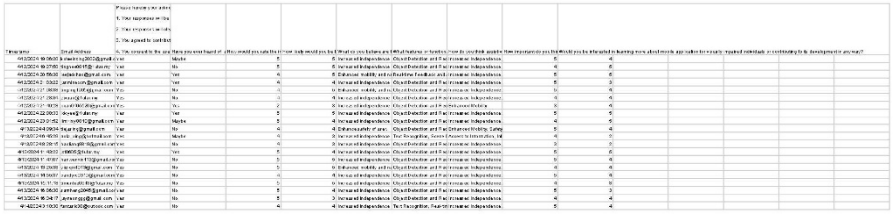
WON WEI CHENG, 2021. *OBJECT RECOGNITION MOBILE APP FOR VISUALLY IMPAIRED USER WON WEI CHENG A project report submitted in partial fulfilment of the requirements for the award of Bachelor of Science (Honours) Software Engineering*,

APPENDICES

Appendix A: Graphs

-

Appendix B: Tables



Appendix B: B Questionnaire Response Table

Test Case ID	Description	Expected Outcome	Pass/Fail	User Comment	Rating (1-5)
TC001	Application launches successfully	The application should open without errors	Pass	No Comment. Good performance.	5
TC002	Camera functionality	The application should open without errors	Pass	No Comment. Good performance.	5
TC003	Object detection	The application should detect objects accurately	Pass	No Comment. Good performance.	5
TC004	Text recognition	The application should recognize and read text correctly	Pass	No Comment. Good performance.	5
TC005	Text-to-speech functionality	The application should read the recognized text aloud	Pass	No Comment. Good performance.	5
TC006	User interface clarity	The UI should be clear and easy to navigate	Pass	No Comment. Good performance.	5
TC007	Overall performance	The application should run smoothly without significant lag	Pass	No Comment. Good performance.	5
TC008	Voice Command to Switch Mode	The application was able to switch between different mode with voice command	Pass	No Comment. Good performance.	5

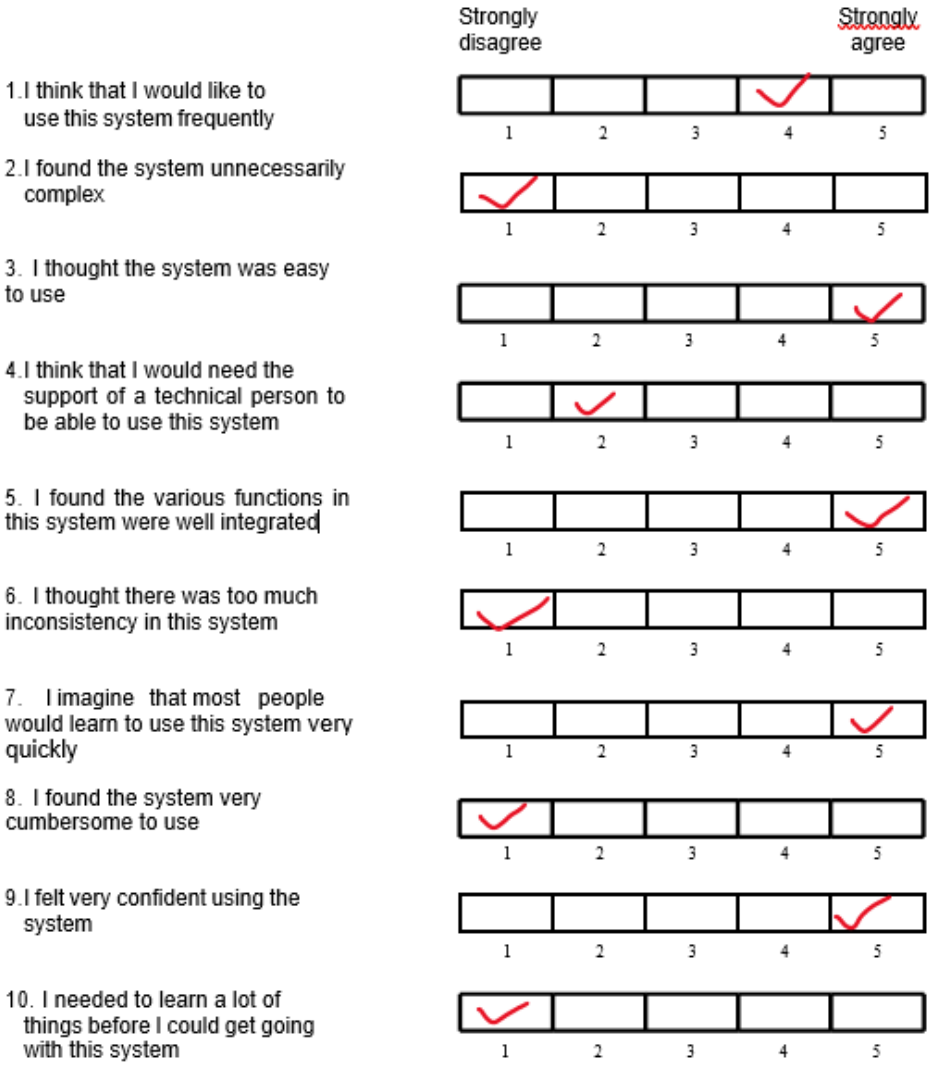
Appendix B: C User Acceptance Test Respond 1

Test Case ID	Description	Expected Outcome	Pass/Fail	User Comment	Rating (1-5)
TC001	Application launches successfully	The application should open without errors	Pass	App open successfully.	5
TC002	Camera functionality	The application should open without errors	Pass	Can see camera captured frame	5
TC003	Object detection	The application should detect objects accurately	Pass	Able to detect objects	5
TC004	Text recognition	The application should recognize and read text correctly	Pass	Can recognize text in english	5
TC005	Text-to-speech functionality	The application should read the recognized text aloud	Pass	Clear feedback from application	5
TC006	User interface clarity	The UI should be clear and easy to navigate	Pass	UI is easy to understand and use	5
TC007	Overall performance	The application should run smoothly without significant lag	Pass	Good performance	5
TC008	Voice Command to Switch Mode	The application was able to switch between different mode with voice command	Pass	Can recognize command to switch mode	4

Appendix B: D User Acceptance Test Respond 2

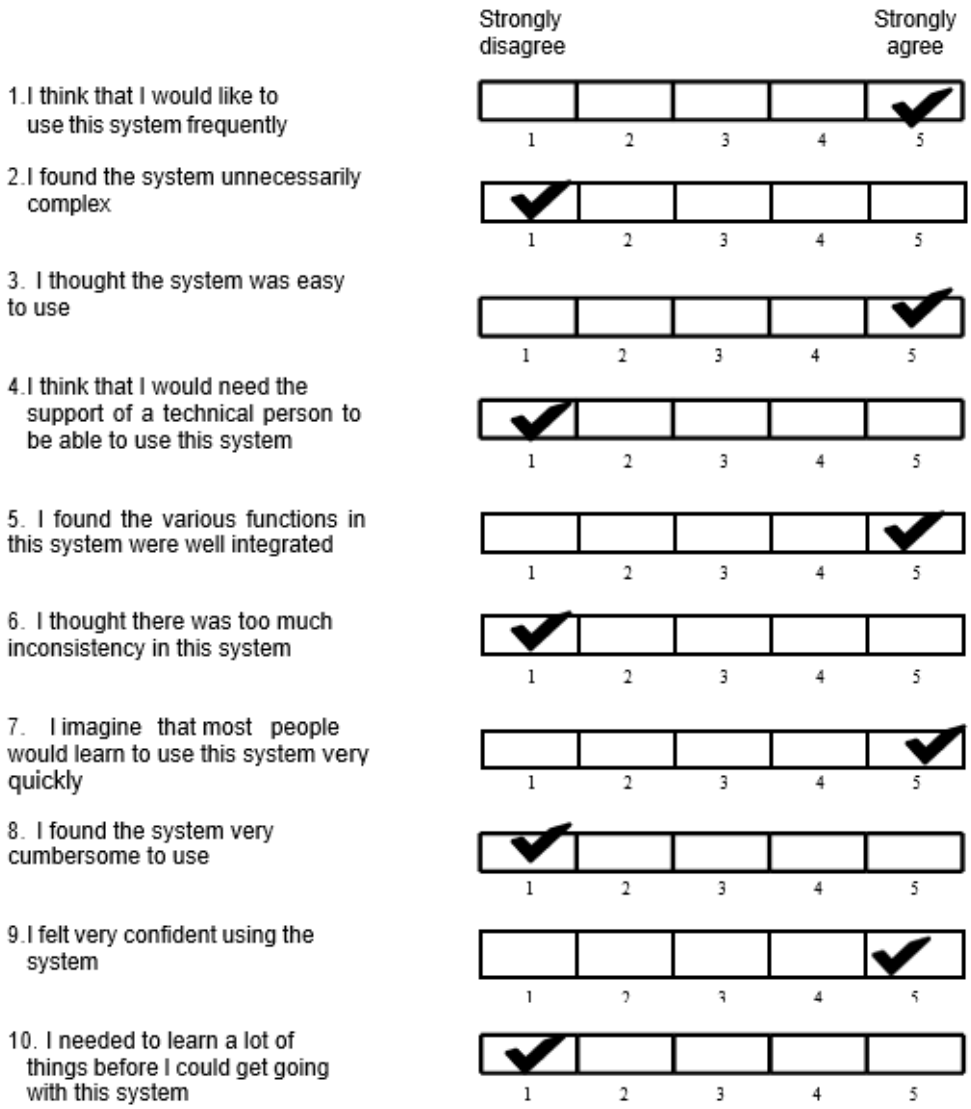
Test Case ID	Description	Expected Outcome	Pass/Fail	User Comment	Rating (1-5)
TC001	Application launches successfully	The application should open without errors	Pass	All Good.	5
TC002	Camera functionality	The application should open without errors	Pass	All Good.	5
TC003	Object detection	The application should detect objects accurately	Pass	All Good.	5
TC004	Text recognition	The application should recognize and read text correctly	Pass	All Good.	5
TC005	Text-to-speech functionality	The application should read the recognized text aloud	Pass	All Good.	5
TC006	User interface clarity	The UI should be clear and easy to navigate	Pass	All Good.	5
TC007	Overall performance	The application should run smoothly without significant lag	Pass	All Good.	5
TC008	Voice Command to Switch Mode	The application was able to switch between different mode with voice command	Pass	All Good.	5

Appendix B: E User Acceptance Test Respond 3



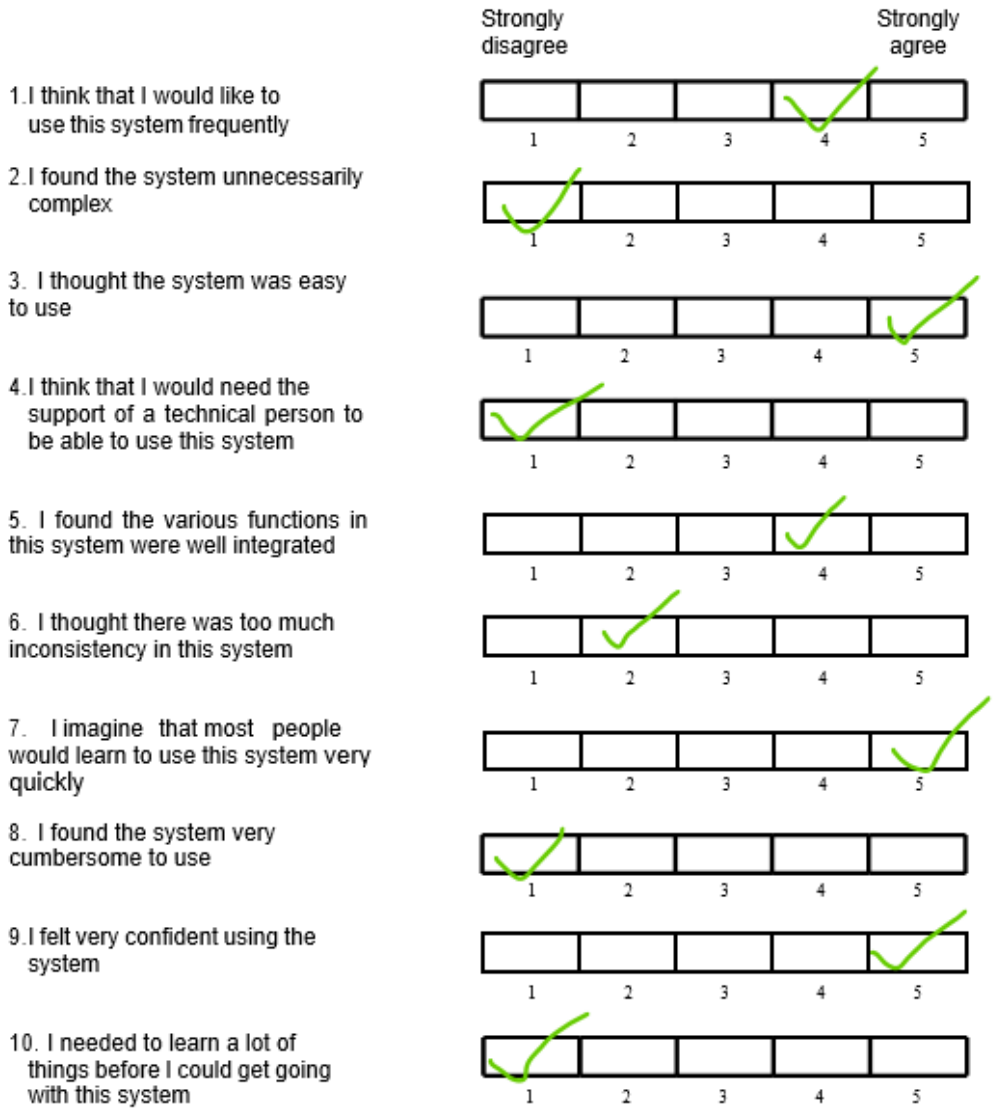
Appendix B: F Software Usability Scale Respond 1

System Usability Scale



Appendix B: G Software Usability Scale Respond 2

System Usability Scale



Appendix B: H Software Usability Scale Respond 3