Virtual Eye for the Blind

Bhagyashri Bhamare
Dept. of Information Technology
NITK Surathkal
Mangalore, India
bhamare.bhagyashri1999@gmail.com

Chinmayi C. R.
Dept. of Information Technology
NITK Surathkal
Mangalore, India
chinmayicr27@gmail.com

K. Keerthana
Dept. of Information Technology
NITK Surathkal
Mangalore, India
keerthanakanapuram@gmail.com

***Abstract*—Technological advancements can be used to assist the blind. We have built an android application that helps blind people detect objects and hurdles in the path. This is done by giving a brief about the surrounding environment. A speech system is incorporated to the application to dictate the obstacles in the surrounding. An user interface(UI) is built for the application keeping in mind all the Human Computer Interaction Design Principles. This user friendly UI has been tested on blindfolded people for empirical analysis. This eliminates the need for eye to a great extent. There are some extra functionalities like text recognition and colour recognition. Since its an application, it can be installed in the phone and used easily. This app can be run on any low-end smartphone devices.**

***Index Terms*—Human Computer Interaction, user interface, empirical analysis, blind, android application**

I. Introduction

Some groups of the society face difficulty in carrying out activities that most of us take for granted. It is estimated that 39 million persons globally are blind and 246 million have low vision, and that about 90% of them live in low-income settings [1]. There are devices designed for accessibility features but hardly any specially designed for non-sighted ones. Visually impaired people face detecting objects around them as an extreme challenge. Hence, it is important to use technology to give the visually impaired an alternative gift of vision. This can be achieved with a visual assistance using AI. Most products are designed only for this purpose [2] which increases overhead expenses to the blind. A common use of technological features is through mobile phones. Its difficult to find someone who doesn't possess a mobile phone. Virtual Eye is an android application designed and developed for visually impaired people. Every mobile device has a built in speech recognition system. The app is flexible, portable and user friendly. The user interface is built to enhance simplicity and thus increase ease of use. This application is different from all the other applications that are close to it that have been done before because we collect real time data. There is no capturing and collection of data for processing. This makes it extremely user friendly. A detailed empirical analysis is done on 30 participants to further prove the efficiency of the application.

- A complete speech controlled mechanism is used to command and perform operations in the application.
- The application detects obstacles around the person and dictates out it in the form of speech.
- This application includes a text reader that reads out the text on an object or a sheet of paper that the camera is focused on.

II. Literature Survey

A technique [3] for detecting obstacles, darkness, and tracking the visually impaired individual was proposed. This method [4] uses an ultrasonic sensor and water sensor to detect obstacles including pits and pebbles. Color, light, entity, and banknote detection is possible with this Android app [5]. Blinds were asked to take photos in this experiment [6], and the findings were wildly inaccurate. The use of a GPS module aids blind people in their navigation. [7] Many special standalone devices were designed for people with loss of vision such as talking watches, thermometers,scales, blood glucose reader and blood pressure measuring,among others. Colorino Talking Color Identifier [9] is a standalone device that is used to recognize colors and detect light. When the Infrared beam hits an obstacle, you'll be able to figure out what it is. Pen friend [10], another standalone device that has the shape of a pen, is packaged with magnetic coded labels so that the user can paste the label onto an object, and then uses his pen to recognize the associated code and record a name for the label with his voice. In addition to the above mentioned devices, there are other mobile applications on the internet that provide such services or features for the blind. An example is the "Seeing Assistant" [11].

Authors	Methodology	Merits	Limitations	Additional Details
Joe Louis Paul I, Sasirekha S etal	GPS tracking, sensors, Light Dependent Resistor	Portable, Offline method	Overhead expenses, no text detection	Use of Raspberry Pi
Mukesh Prasad Agrawal, Atma Ram Gupta	Water sensor, ultrasonic sensor, Microcontroller	Efficient, unique, user-friendly	No text or colour detection, no speech	RF_Module, GPS_GSM,
Milios Awad, Jad El Haddad, Edgar Khneisser, Tarek Mahmoud, Elias Yaacoub, Mohammad Malli	CraftAR On-DeviceImage-Recognition SDK, Convolutional Neural Networks (CNN) on Android	Colour recognition, light recognition, Banknote recognition	No speech system	PHP and MySQL for the database
Noboru Takagi, Yuichiro Mori	Voting of Local Features, Image Retrieval using Binary Decision Tree	Simple to use	Inaccurate results	Use of ORB vector
Akhilesh Krishnan, Deepakraj G, Nishanth N, Dr.K.M.Anandkumar	Iterative Deepening Depth First Search	Static and dynamic object recognition	Quality of dynamic object recognition	Echolocation, Bluetooth module

TABLE I
SUMMARY OF LITERATURE SURVEY

III. Problem Statement

We intend to use AI to describe people, text and objects to a blind person, allowing them to perceive their near surroundings. Our project will be able to tell visually-impaired people what is around them. We attempt to build an easy to use, reliable and hassle free UX for the visually impaired.

A. Objectives

- Bridge the barrier between technology and blind assistance
- Build an efficient speech control system to guide the visually impaired
- Enable text recognition to accurately read texts
- Provide an user-friendly, easy to use platform

IV. Methodology

A. Activity Diagrams

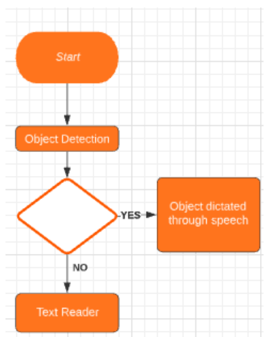

Fig. 1. Object Detection Flow Diagram

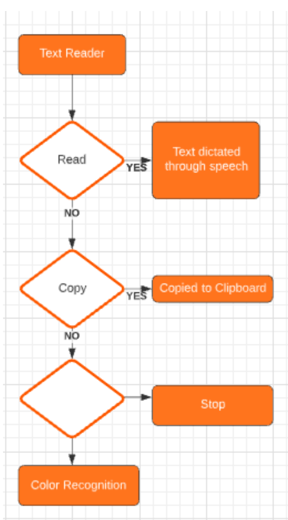

Fig. 2. Text Recognition Flow Diagram

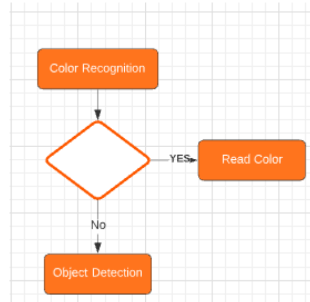

Fig. 3. Colour Recognition Flow Diagram

B. Object Detection

YOLO (You Only Look Once) [12] is a unified, real time object detection technique. A convolutional neural network (CNN) is used for object detection. The algorithm divides the image into regions and predicts bounding boxes and probabilities for each region using a single neural network applied to the entire image. The estimated probabilities are used to weight these bounding boxes.

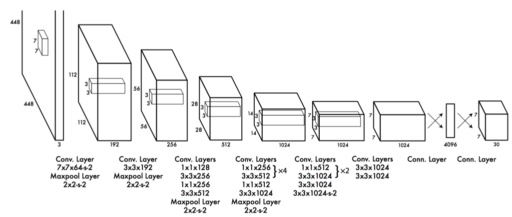

Fig. 4. YOLO Architecture

C. Text Recognition

OCR (Optical Character Recognition) [13] is used widely for text recognition. It automates image processing and converts into text. In our work, this text is then converted to speech. Matrix matching is one of the type of OCR algorithm. Matrix matching is used to compare an image to a stored glyph. This is done in a pixel-by-pixel basis.

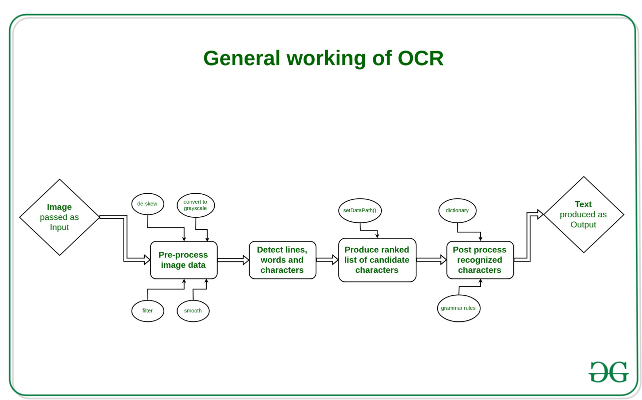

Fig. 5. OCR recognition

D. Algorithms/Pseudo Code

1) Text to Speech: Detected object is displayed as text. This text is converted to speech for the blind user.

```
# Import the required module
from gtts import gTTS

# Module to play converted audio
import os

# The text to convert to audio
text = 'Welcome to Virtual Eye'

# Language in which you want to convert
language = 'en'

tsobj = gTTS(text=mytext, lang=language, slow=False)

# Save the converted audio in a mp3 file
tsobj.save("speech.mp3")

# Playing the converted file
os.system("mpg321 speech.mp3")
```

Fig. 6. Text To Speech Code

V. Results and Analysis

A. Design Principles

- **Consistency**

A consistency is maintained as it is a complete speech controlled system.

- **Informative feedback**

Current status of the process is given by voice assistant.

- **Recognition rather than recall**

The system doesn't demand the user to recall anything. It works on recognition.

- **Match between system and the real world**

The objects are dictated in the language understandable to the blind.

- **User control and freedom**

The users are given freedom to use the system according to their control.

- **Flexibility and efficiency of use**

The application offers flexibility and overall efficiency of use. It can be installed in any device. Its a light weight application.

- **Aesthetic and minimalist design**

A simple and easy to use design has been used.

B. Empirical Analysis

The reviews of 30 users were taken to perform an empirical analysis on the application built for the purpose of this project. The findings from the reviews of these users are depicted in the following graphs. The trend in time taken for the app to take action shows that it took a marginally longer time the first time the application was used. However, the time taken for the user to navigate the screen drops steeply after the first time. The average rating of the app was found to be 7.76.

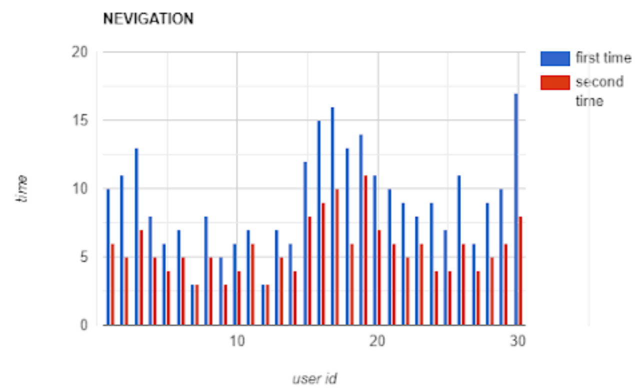

Fig. 7. Screen Analysis

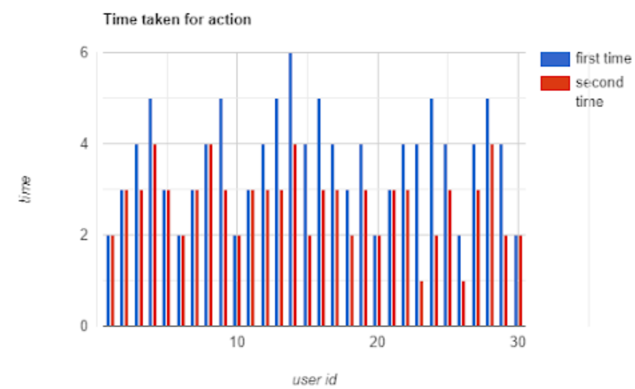

Fig. 8. Time taken for action

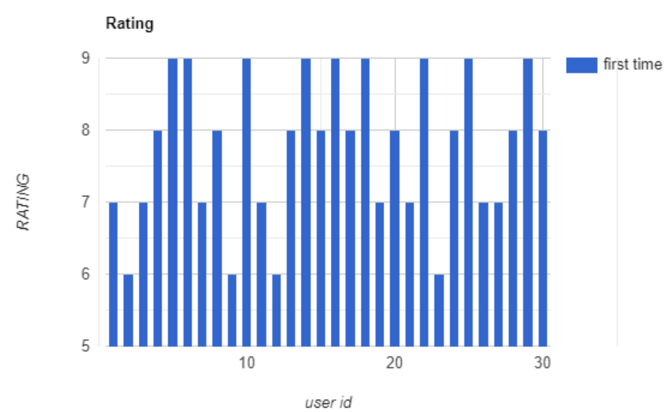

Fig. 9. Usability Rate

User	Time for navigation(seconds) First time	Time for navigation(seconds) Second time
1	10	6
2	11	5
3	13	7
4	8	5
5	6	4
6	7	5
7	3	3
8	8	5
9	5	3
10	6	4
11	7	6
12	3	3
13	7	5
14	6	4
15	12	8
16	15	9
17	16	10
18	13	6
19	14	11
21	10	6
22	9	5
23	8	6
24	9	4
25	7	4
26	11	6
27	6	4
28	9	5
29	10	6
30	17	8

TABLE II
SCREEN ANALYSIS

User	Time for navigation(seconds) First time	Time for navigation(seconds) Second time
1	2	2
2	3	3
3	4	3
4	5	4
5	3	3
6	2	2
7	3	3
8	4	4
9	5	3
10	2	2
11	3	3
12	4	3
13	5	3
14	6	4
15	4	2
16	5	3
17	4	3
18	3	2
19	4	3
20	2	2
21	3	3
22	4	3
23	4	1
24	5	2
25	4	3
26	2	1
27	4	3
28	5	4
29	4	2
30	2	2

TABLE III
TIME FOR ACTION

C. Model Based Design Evaluation

Goals, Operators, Methods, and Selection (GOMS) model. In principle GOMS can be used to answer two different questions concerning the usability of a user interface by disabled users. First, it can be investigated if the overall design of the user interface itself forces an unacceptable disadvantage for disabled users. Second, GOMS models can be used to evaluate if disabled users are in principle able to reach predefined performance goals.

- grant permissions
- choose language
- verification of a single tap on object detection
- speech to convert object description text

User id	Rating(/10)
1	7
2	6
3	7
4	8
5	9
6	9
7	7
8	8
9	6
10	9
11	7
12	6
13	8
14	9
15	8
16	9
17	8
18	9
19	7
20	8
21	7
22	9
23	6
24	8
25	9
26	7
27	7
28	8
29	9
30	8

TABLE IV
USER RATING

- single tap on text reader
- read text
- single tap on stop
- single tap on copy text
- copy text to clipboard
- Yes or No prompt to copy text
- colour recognition
- read the color to the blind user
- single tap on object detection
- repeat the steps

D. Task Modelling and Analysis

Hierarchical Task Analysis

Task 1: Open Camera

Task 1.1: Grant permissions when the app is set up
Task 1.2:Access system settings to open camera

Task 2: Tap on the screen

Task 2.1: Single tap on screen
Task 2.3: Double tap on screen to get voice assistance

Task 3: Repeat Task 2 for every surrounding assistance

Task 4: Exit the application

Fig. 10. Hierarchical Task Analysis

E. Screenshots

Follwing are the screenshots of our application

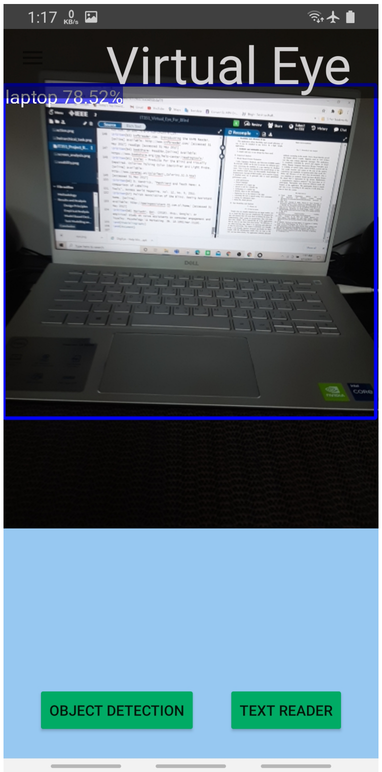

Fig. 11. Object Detection

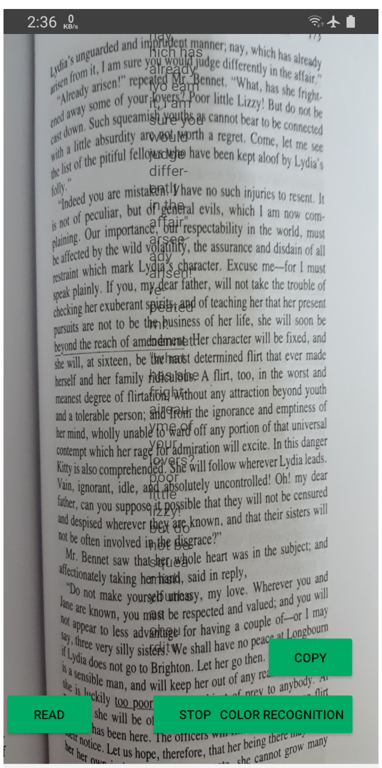

Fig. 12. Text Detection

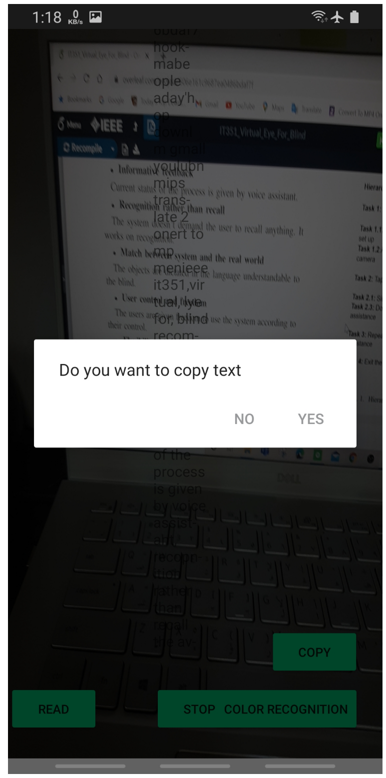

Fig. 13. Copy Text

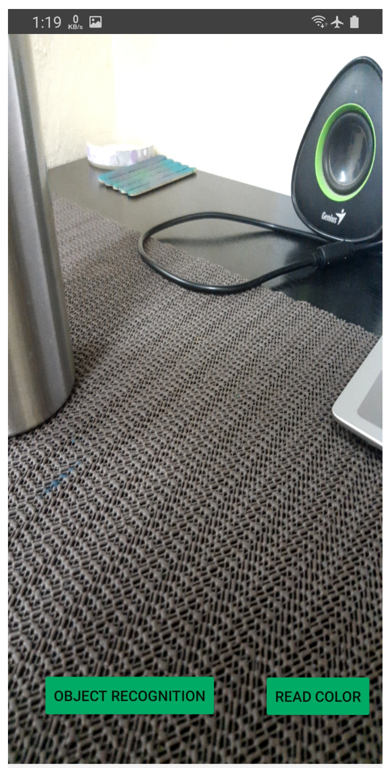

Fig. 14. Colour Detection

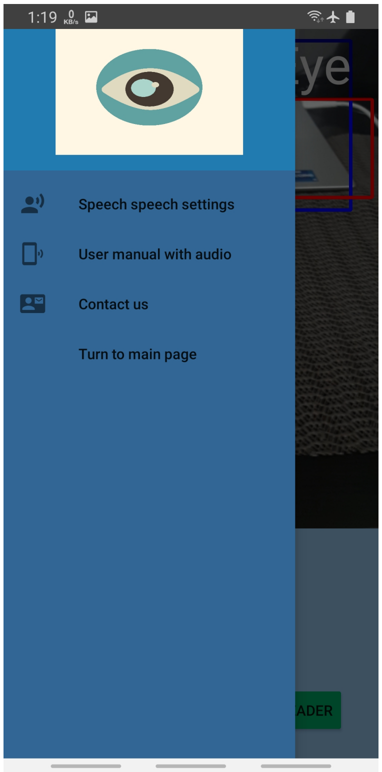

Fig. 15. Navigation

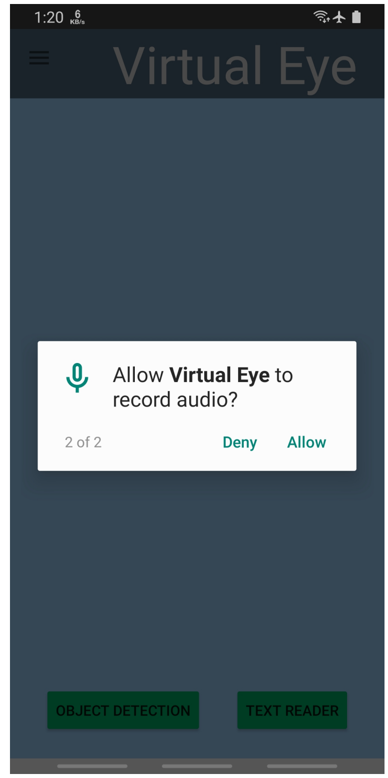

Fig. 17. Permission for Audio Access

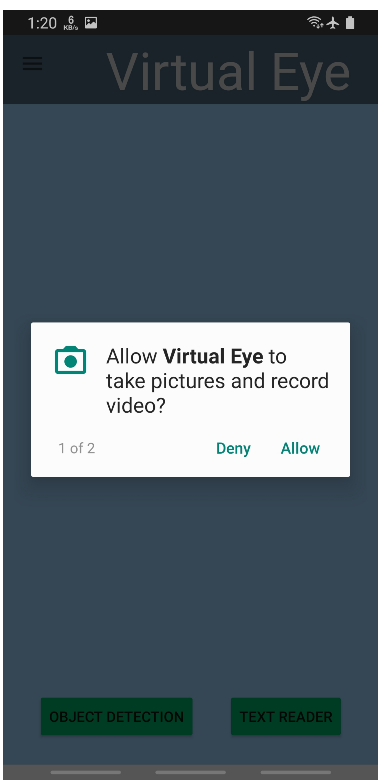

Fig. 16. Permission for Camera Access

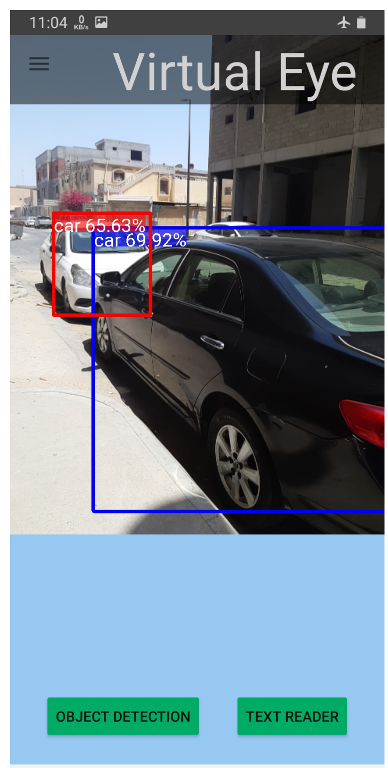

Fig. 18. Detection outside the house

Fig. 19. Detection of moving car

VI. Conclusion

People who are visually impaired are our target market, and we've found that their lives are complicated. We designed the app with features that will assist them in making decisions. By understanding what's going on around them, they will make their lives simpler. It has a user-friendly interface designed specifically for blind people, with the identification results read out loud so the user can hear them clearly. The application correctly detected artifacts, according to the results. Not to forget that the speech out feature allows visually impaired users to view written text that may contain important messages. The application is completely controlled by speech and requires no visual ability. It provides additional functionality of text recognition and colour recognition. Human Computer Interaction Design Principles such as visibility of system status, match between system and the real world, freedom and control of the user of the system, consistency and standards, recognition rather than recall, flexibility and efficiency of use have been incorporated. Also, a very simplistic yet aesthetic design has been adopted. Furthermore, a comprehensive empirical analysis has been done on 30 participants. Participants were tested on different levels and features of the application. The participants found it simple and easy to use. According to the analysis, it took them less time to learn the system. There is some scope for improvement in terms of accuracy and added features. The UI can be enhanced further.

VII. Individual Contribution

The following table summarizes:

Name	Roll Number	Contribution to the project (percentage)
Bhagyashri Bhamare	181IT111	Object Detection CNN model design, Color Recognition UI, OCR model implementation, empirical analysis, design analysis (33.33%)
Chinmayi C. Ramakrishna	181IT113	Object Detection CNN pre processing, output optimisation, Navigation Bar UI, OCR model implementation and design analysis (33.33%)
K. Keerthana	181IT221	Color Recognition module, Object Detection UI, Text Recognition UI, Text to Speech module, empirical analysis (33.33%)

TABLE V
INDIVIDUAL CONTRIBUTION

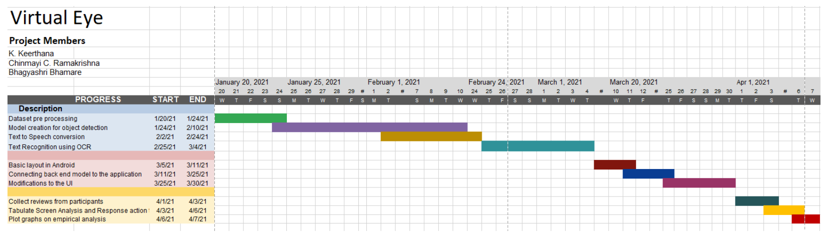

Fig. 20. Gantt Chart

References

[1] World Health Organization (WHO). 2014. Visual Impairment and Blindness. [online] available: http://www.who.int/mediacentre/factsheets/fs282/en/ [accessed 31 May 2017]

[2] Britto Neto, Laurindo & Maike, Vanessa & Koch, Fernando & Baranauskas, Maria & Rocha, Anderson & Goldenstein, Siome. (2016). A Wearable Face Recognition System Built into a Smartwatch and the Blind and Low Vision Users. 515-528. 10.1007/978-3-319-29133-8_25.

[3] Joe Louis Paul I, Sasirekha S etal "Smart Eye for Visually Impaired-An aid to help the blind people" Second International Conference on Computational Intelligence in Data Science (ICCIDS-2019). http://www.caretec.at/ColorTest_Colorino.32.0.html [accessed 31 May 2017]

[4] Mukesh Prasad Agrawal and Atma Ram Gupta "Smart Stick for the Blind and Visually Impaired People" Proceedings of the 2nd International Conference on Inventive Communication and Computational Technologies (ICICCT 2018).

[5] Milios Awad, Jad El Haddad etal "Intelligent Eye: A Mobile Application for Assisting Blind People" 2018 IEEE Middle East and North Africa Communications Conference (MENACOMM).

[6] Noboru Takagi and Yuichiro Mori "High Speed Image Retrieval Method Executable on Smartphones: Toward Vision Assistance for Blind People". 2018 IEEE International Conference on Systems, Man and Cybernetics.

[7] Akhilesh Krishnan, Deepakraj G etal "Autonomous Walking Stick For The Blind Using Echolocation And Image Processing".

[8] Moriuchi, Emi. (2019). Okay, Google!: An empirical study on voice assistants on consumer engagement and loyalty. Psychology & Marketing. 36. 10.1002/mar.21192.

[9] areTec - Products for the Blind and Visually Impaired. Colorino Talking Color Identifier and Light Probe. [online] available: http://www.caretec.at/ColorTest_Colorino.32.0.html [accessed 31 May 2017]

[10] D. Kendrick, "PenFriend and Touch Memo: A Comparison of Labeling Tools", Access World Magazine, Vol. 12, No. 9, 2011.

[11] Polish Association of the Blind. Seeing Assistant home. [online], available: http://seeingassistant.tt.com.pl/home/ [accessed 31 May 2017]

[12] https://www.haxed.dk/tablesoccer/2019/02/27/transfer-learning/

[13] https://www.geeksforgeeks.org/tesseract-ocr-with-java-with-examples/