VIVEKANAND EDUCATION SOCIETY'S INSTITUTE OF TECHNOLOGY

Department of Computer Engineering

Project Report On

SAHARA - A STANCHION FOR THE VISUALLY IMPAIRED

In partial fulfillment of the Fourth Year, Bachelor of Engineering (B.E.) Degree in Computer Engineering at the University of Mumbai Academic Year 2017-2018

Submitted by

Amey Pimple (D17C, R.No.48)

Dhaval Bagal (D17B, R.No.03)

Rahul Devadiga (D17A, R.No.20)

Shreya Patil (D17C, R.No.46)

Project Mentor

Mrs. Sharmila Sengupta

(2019-20)

VIVEKANAND EDUCATION SOCIETY'S INSTITUTE OF TECHNOLOGY

Department of Computer Engineering

Certificate

This is to certify that ***Amey Pimple, Dhaval Bagal, Rahul Devadiga and Shreya Patil*** of Fourth Year Computer Engineering studying under the University of Mumbai have satisfactorily completed the project on "***SAHARA - A STANCHION FOR THE VISUALLY IMPAIRED***" as a part of their coursework of PROJECT-II for Semester-VIII under the guidance of their mentor ***Prof. Sharmila Sengupta*** in the year 2019-2020.

This project report entitled ***Sahara - A Stanchion for the Visually Impaired*** by ***Dhaval Bagal*** is approved for the degree of ***Bachelor of Engineering in Computer Engineering***.

Programme Outcomes	Grade
PO1,PO2,PO3,PO4,PO5,PO6,PO7,PO8,PO9, PO10, PO11, PO12, PSO1, PSO2	

Date:
Project Guide:

NATIONAL ASSOCIATION FOR THE BLIND, INDIA

(Registered under Societies Registration Act XXI of 1860 and Bombay Public Trusts Act 1950)

ADVOCACY CELL

11 KHAN ABDUL GAFFAR KHAN ROAD, WORLI SEAFACE, MUMBAI 400 030.
Telephone: 022-24935365/5370/2539 Ext. 221/242/244
Email: advocacy@nabindia.info • Website: www.nabindia.org

President
Mr. Hemant Takle

Honorary Secretary General
Mr. Satya Kumar Singh

Ref: ADV:SVK/2019/034

September 17, 2019

Dr. Nupur Giri
H.O.D. – Computer Department,
Vivekanand Education Society's Institute of Technology,
Chembur,
MUMBAI – 400 074.

Dear Sir,

We at NAB, India are happy to meet the students of your College along with Prof. Sharmila Sengupta, at our office on September 17, 2019. During our discussions, to identify problems faced by the blind, we have realised a need for a Hardware Device to identify Indian Currency Notes.

We have requested the team comprising of: Prof. Sharmila Sengupta (Mentor), Mr. Dhaval Bagal, Mr. Amey Pimple, Mr. Rahul Devadiga and Ms. Shreya Patil from your College, to undertake the research of such Hardware and make a full proof proto-type.

Your co-operation in this endeavour will be highly appreciated.

Regards,

Yours truly,

SUHAS V. KARNIK
HON. SECRETARY

svk/jd

Project Report Approval
For
B. E (Computer Engineering)

This project report entitled ***Sahara - A Stanchion for the Visually Impaired*** by ***Dhaval Bagal*** is approved for the degree of ***Bachelor of Engineering in Computer Engineering***.

Internal Examiner

External Examiner

Head of the Department

Principal

Date:
Place:

Declaration

We declare that this written submission represents our ideas in our own words and where others' ideas or words have been included, we have adequately cited and referenced the original sources. We also declare that we have adhered to all principles of academic honesty and integrity and have not misrepresented or fabricated or falsified any idea/data/fact/source in our submission. We understand that any violation of the above will be cause for disciplinary action by the Institute and can also evoke penal action from the sources which have thus not been properly cited or from whom proper permission has not been taken when needed.

Dhaval Bagal (D17B, R.No.03)

Rahul Devadiga (D17A, R.No.20)

Shreya Patil (D17C, R.No.46)

Amey Pimple (D17C, R.No.48)

Date:

ACKNOWLEDGMENT

We are thankful to our college **Vivekanand Education Society's Institute of Technology** for considering our project and extending help at all stages needed during our work of collecting information regarding the project.

It gives us immense pleasure to express our deep and sincere gratitude to Associate Professor **Mrs. Sharmila Sengupta** (Project Guide) for her kind help and valuable advice during the development of project synopsis and for her guidance and suggestions.

We are deeply indebted to Head of the Computer Department **Dr. (Mrs.) Nupur Giri** and our Principal **Dr. (Mrs.) J.M. Nair**, for giving us this valuable opportunity to do this project.

We express our hearty thanks to them for their assistance without which it would have been difficult in finishing this project synopsis and project review successfully.

We convey our deep sense of gratitude to all teaching and non-teaching staff for their constant encouragement, support and selfless help throughout the project work. It is great pleasure to acknowledge the help and suggestion, which we received from the Department of Computer Engineering.

We wish to express our profound thanks to all those who helped us in gathering information about the project. Our families too have provided moral support and encouragement at several times.

Department of Computer Engineering
Course Outcomes for B.E Project

Learners will be to,

Course Outcome	Description of the Course Outcome
CO 1	Able to apply the relevant engineering concepts, knowledge and skills towards the project.
CO2	Able to identify, formulate and interpret the various relevant research papers and to determine the problem.
CO 3	Able to apply the engineering concepts towards designing solution for the problem.
CO 4	Able to interpret the data and datasets to be utilized.
CO 5	Able to create, select and apply appropriate technologies, techniques, resources and tools for the project.
CO 6	Able to apply ethical, professional policies and principles towards societal, environmental, safety and cultural benefit.
CO 7	Able to function effectively as an individual, and as a member of a team, allocating roles with clear lines of responsibility and accountability.
CO 8	Able to write effective reports, design documents and make effective presentations.
CO 9	Able to apply engineering and management principles to the project as a team member.
CO 10	Able to apply the project domain knowledge to sharpen one's competency.
CO 11	Able to develop professional, presentational, balanced and structured approach towards project development.
CO 12	Able to adopt skills, languages, environment and platforms for creating innovative solutions for the project.

Abstract

Around 12 million persons in India suffer from blindness. They face challenges in performing daily activities. Major problems faced by visually impaired are recognizing currency, reading text and color identification, etc. With the advent of new technologies like Machine learning, Artificial Intelligence and rapid improvement in processing it has become possible to bridge the gap between the visually impaired people and the world.

The aim of this project is to assist the visually impaired in his routine chores. Sahara acts as an aid for dealing with such challenges. It is a handy, portable, internet-less device that comes along with an android application to be installed in the phone. Thus the device acts as a server with which the Sahara android application interacts. The device acts as an accessory for a smart-phone. It boasts features like currency recognition and counting, text-reading, summarizing a text excerpt to get the major essence of its content, color identification, supermarket bill reading, business card reading etc. It provides a constant performance no matter which kind of android device it is connected to. Bringing in the psychology factor again, a blind user will definitely not have any problem in shelling out few thousands for a normal android smartphone along with Sahara which would cost around INR 2K-2.5K (considering mass production)

Sahara also gains an upper hand when it comes to the user experience. The power bank connected device can be simply put in a bag and the user can utilise its features just from his phone. Apart from this, the Sahara android application provides a voice-based and a tap-based interface that gets all the jobs done without any hassle.

Index

Chapter 1: Introduction

1.1 Introduction

Around 12 million persons in India suffer from blindness. They face challenges in performing daily activities. Major problems faced by visually impaired are recognizing currency, reading text and color identification, etc. With the advent of new technologies like Machine learning, Artificial Intelligence and rapid improvement in processing it has become possible to bridge the gap between the visually impaired people and the world.

1.2 Motivation

While browsing some videos on YouTube, we stumbled upon the advertisement of Microsoft's recent innovation - Seeing A.I, a smart spectacle that aids a blind person in various activities. Being aware of the problems faced by the blind today, we decided to develop something similar that would be accessible to a larger part of the blind population. After cogitating for a few days, we decided to develop an A.I based system that would empower the blind and bridge the gap between him and the corporate world. In order to seek guidance regarding the same, we approached the National Association for the Blind, India (NAB). Upon having a talk with Mr. Suhas Karnik, Mr. Arvind Narvekar and Mrs. Pallavi Kadam, we got to know that currently currency recognition has become a major challenge for the visually impaired in India owing to the recent demonetisation. Also, existing devices like OrCam cost a fortune and are unaffordable to the common man. The conclusion then was to develop a device that would assist a blind in his routine chores which are important to him.

1.3 Problem Definition

Issues faced by the blind population today are multitudinous. Technology has addressed many challenges faced by the blind. However, every technology comes with its own pros and cons.

The major two technologies in lead today are smartphone-applications and cloud-based systems. The price of an android phone varies with its architecture and processor speed; expensive phones possess a high-end processor while less expensive phones have a relatively low-end processor. Taking into consideration the psychology factor, no blind person would buy an expensive phone just for an application that assists him in his quotidian jobs. Today,

there are many android applications that provide various features to aid the visually impaired. However, all of these applications are cloud-technology based and hence compulsorily require a constant internet connection. Apart from this, they also need a monthly or an annual subscription in order to avail their facilities.

1.4 Existing Systems

1.4.1 OrCam

The most advanced wearable assistive technology for the blind and visually impaired, provides independence by conveying visual information, audibly. It reads text, recognizes faces, identifies products and more.

1.4.2 Microsoft Seeing A.I

Seeing AI is a Microsoft research project that brings together the power of the cloud and AI to deliver an intelligent app, designed to help you navigate your day.

1.5 Lacunas in the Existing Systems

- Exorbitant prices.
- Internet traffic due to usage of cloud.
- Internet dependency due to usage of cloud
- Monthly or Annual subscriptions for existing products.
- Offline android applications cannot perform heavy Machine Learning/Deep Learning tasks.
- Heterogeneity of smart phones in case of android applications make it difficult to provide a common android based platform.

1.6 Relevance of the Project

There is a dire need to focus on problems faced by the blind and contrive a way out to deal with them. The proposed system will help the visually impaired to perform currency recognition, text recognition, color identification, text summarization, etc. By avoiding certain unnecessary features like navigation, scene description which require extensive

computation, the final end product aims to make itself available at an affordable cost that would have high utilisation and would reach a larger mass.

1.7 Methodology Employed

1.7.1 Agile Methodology

Agile software development refers to software development methodologies centered round the idea of iterative development, where requirements and solutions evolve through collaboration between self-organizing cross-functional teams. The ultimate value in Agile development is that it enables teams to deliver value faster, with greater quality and predictability, and greater aptitude to respond to change.

The project was developed using Agile methodology with the first deliverable being a stand-alone slit-based hardware device for currency recognition. Later on, after multiple iterations, the final deliverable is a device that connects with the mobile phone and provides various functionalities to the blind.

Chapter 2: Literature Survey

2.1 Literature Survey

2.1.1 NAB files petition against RBI

The Bombay High Court on Thursday sought to know from the Reserve Bank of India (RBI) the reason behind changing from time to time the size and other features in currency notes and coins.

The query was posed by a division bench of Chief Justice Pradeep Nandrajog and Justice NM Jamdar while hearing a petition filed by the National Association of the Blind (NAB).

NAB petition claimed that new currency notes and coins issued by the RBI posed difficulty for visually impaired persons in terms of identification and distinguishing them.

"We want to know from the RBI what is the compulsion to keep changing the features like size and so on in the currency notes," Chief Justice Nandrajog said.

The court added that no other country in the world keeps changing size and features of their currency notes so often.

The court was informed that in March this year the RBI had issued new coins with special features to help visually impaired persons distinguish them.

The court directed the apex bank to file its affidavit within a period of six weeks on the issue.

The petition has sought for directions to the central bank to include distinctive features in the new coins and currency notes.

https://www.news18.com/news/india/why-keep-changing-features-of-currency-notes-coins-bombay-high-court-asks-rbi-2254757.html

2.1.2 RBI launches MANI

The Reserve Bank of India (RBI) has launched a mobile application called ‘Mobile Aided Note Identifier’ or ‘MANI’. The ‘MANI’ would help visually challenged persons to identify the denomination of currency notes, and thereby facilitating their day to day transactions. The app was launched in the presence of RBI Governor Shaktikanta Das in the national capital.

“Indian banknotes contain several features which enable the visually impaired (colour blind, partially sighted and blind people) to identify them, viz., intaglio printing and tactile mark, variable banknote size, large numerals, variable colour, monochromatic hues and patterns. Technological progress has opened up new opportunities for making Indian banknotes more accessible for the visually impaired, thereby facilitating their day to day transactions,” the RBI said in a statement.

https://www.financialexpress.com/industry/technology/rbi-launches-mani-app-to-help-visually-challenged-identify-currency-notes-how-to-download-it/1810477/

2.2 Research Papers

2.2.1 Smart device for visually impaired people

R. Kasthuri, B. Nivetha, S. Shabana, M. Veluchamy and S. Sivakumar, "Smart device for visually impaired people," 2017 Third International Conference on Science Technology Engineering & Management (ICONSTEM), Chennai, 2017, pp. 54-59.

The objective of this paper is to guide unsighted people with smart device using an Android Phone. The application helps the user to open any app as well as to call any contact through voice commands. Users can command a mobile device to do something via speech.

2.2.2 Image recognition for visually impaired people by sound

K. G. Krishnan, C. M. Porkodi and K. Kanimozhi, "Image recognition for visually impaired people by sound," 2013 International Conference on Communication and Signal Processing, Melmaruvathur, 2013, pp. 943-946.

The aim of this paper is to represent a method where a blind person can get information about the shape of an image through speech signal. The novelty of this paper is to convert the image to sound using the methodology of edge detection.

2.2.3 Voice Assistance for Visually Impaired People

A. KARTHIK, V. K. RAJA and S. PRABAKARAN, "Voice Assistance for Visually Impaired People," 2018 International Conference on Communication, Computing and Internet of Things (IC3IoT), Chennai, India, 2018, pp. 465-468.

This paper describes a device to help those people to read environmental messages, words, letters used in postal letters, daily newspapers, and so on to cope up with the social life.The main aim of this proposal is to overcome the above problem by the utilization of the Raspberry pi & OCR sensor for automatic recognition of the environmental messages and by utilization of TTS, the environmental messages are converted to the speech or audio for better and easy interaction with the society.

2.2.4 Visual Assistance for Blind Using Image Processing

B. Deepthi Jain, S. M. Thakur and K. V. Suresh, "Visual Assistance for Blind Using Image Processing," 2018 International Conference on Communication and Signal Processing (ICCSP), Chennai, 2018, pp. 0499-0503.

The proposed system aims to create a wearable visual aid for visually impaired people in which speech commands are accepted from the user. Its functionality addresses identification of objects and sign boards.

2.2.5 Currency recognition system for visually impaired: Egyptian banknote as a study case

N. A. Semary, S. M. Fadl, M. S. Essa and A. F. Gad, "Currency recognition system for visually impaired: Egyptian banknote as a study case," 2015 5th International Conference on Information & Communication Technology and Accessibility (ICTA), Marrakech, 2015, pp. 1-6.

The basic techniques utilized in the proposed system include image foreground segmentation, histogram enhancement, region of interest (ROI) extraction and finally template matching based on the cross-correlation between the captured image and our data set.

2.3 Patent Search

2.3.1 Method and system for image recognition for aiding the visually impaired (US8050484B2)

A method for tracking paper currency in a holder, includes: scanning paper currency deposited or removed from a holder; determining the total number of each individual denomination of paper currency contained within the holder based on the scanned paper currency deposited and removed from the holder; recording the total number of each individual denomination of paper currency; determining the total value of the paper currency within the holder; outputting the denomination of paper currency when the paper currency is scanned during depositing or removal from the holder; outputting the recorded number of each individual denomination of paper currency and the total value of the currency within the holder; and wherein the recorded number of each individual denomination of paper currency and the total value of the currency within the holder is dynamically tabulated based on the scanning of paper currency deposited or removed from the holder.

2.3.2 Portable dollar note denomination identifier for blind (CN101814128A)

The invention discloses a portable dollar note denomination identifier for the blind in the technical field of embedded image processing and identification. The identifier adopts image processing as the core technology, informs a user of an identification result through voice, and is particularly suitable for the blind users to carry for judging the dollar note denomination at each place. The identifier consists of a set-top box and note positioning mechanism, an image sensing part, an image processing part, a man-machine interface and a power supply. The dollar note is placed in the positioning mechanism, a key is pressed by the user, and a digital signal processor controls the CMOS sensor to acquire a dollar note image. An identification program in the digital signal processor compares the image with the pre-stored dollar image so as to acquire the denomination of the dollar note. The set-top box and note positioning mechanism, the image sensing part, the image processing part, the man-machine interface and the power supply are integrated in a set-top box of 80*80*20

(mm). The identifier has the remarkable characteristics of portability and voice prompt, and has the advantages of high identification precision, high identification speed and low cost.

2.3.3 Portable blind aid device (US8606316B2)

A blind aid device including enabling a blind person to activate the blind aid device; capturing one or more images related to a blind person's surrounding environment; detecting moving objects from the one or more images captured; identifying a finite number of spatial relationships related to the moving objects; analyzing the one or more images within the blind aid device to classify the finite number of spatial relationships related to the moving objects corresponding to predefined moving object data; converting select spatial relationship information related to the one or more analyzed images into audible information; relaying select audible information to the blind person; and notifying the blind person of one or more occurrences predetermined by the blind person as actionable occurrences.

Chapter 3: Requirements Gathering

3.1 Definition of Requirements Gathering

Requirements elicitation (also known as Requirements Gathering or Capture) is the process of generating a list of requirements (functional, system, technical, etc.) from the various stakeholders (customers, users, vendors, IT staff, etc.) that will be used as the basis for the formal Requirements Definition.

3.2 Functional Requirements

- **Currency Recognition:** The system should accurately detect the denominations for the currencies under real-time conditions.
- **Currency Counting:** If multiple notes are presented before the system, it should identify all of them and return the total count of the detected denominations
- **Text Recognition:** The system should utilise the OCR engine to accurately convert an image into text.
- **Color identification:** Leveraging HSV color space, the system should identify the color of the patch in the image.
- **Text Summarization:** Given a lengthy text, the system must speak out the essence of it.
- **Business Card Reading:** Given a business card, the system should return all the entities within it like Email-Id, Contact no.s, Address, etc.
- **Bill Reading:** Given a supermarket bill, the system should read out the product names and its associated amount.

3.3 Non-Functional Requirements

- **Portability and Compactness:** A blind user needs a portable and a compact device which he can handle easily in crowded places. The invention is both portable and compact with dimensions of 9cm x 6cm x 3cm.
- **Internet-free:** Even though resembling a client-server architecture model, the device is completely internet-free.
- **Independency:** Apart from the power bank and the smart-phone the device is not dependent on any other component.

- **Heterogeneity:** The device is compatible with any android smart phone provided the smart phone has a camera (above 8MP) and Mobile hotspot facility.
- **Openness:** Other features which currently are not a part of Sahara, can be easily added to Sahara. Thus the invention is open to addition of features.
- **Cost-effectiveness:** The device comes at a cost of around INR 2K - 2.5K and the android application is absolutely free.
- **Synthesizability:** The invention provides feedback at every user operation making him/her aware of the current state of the system.
- **Low Response Time:** The invention provides a delay of not more than 3 seconds.
- **Ease of Operation and Simplicity:** The invention gives a voice-based and tap-based interface, where the user can get the operation done by giving voice and screen-tap inputs.

3.4 Constraints

- Sufficient Illumination.
- Power-bank as a power supply for the device.
- Mobile hotspot to be switched on.
- Decent Camera (above 8MP)
- Decent Speaker and Microphone in the smart phone
- Proper orientation of images containing text

3.5 Hardware, Software, Technologies, Tools available

3.5.1 Available Hardware

- Arduino UNO
- Raspberry Pi Zero
- Raspberry Pi 3
- Raspberry Pi 4
- Jetson Nano

3.5.2 Available Algorithms

- RCNN based Object Detection
- Faster RCNN based Object Detection

- YOLO Object Detection

3.5.3 Available Technologies/Tools

- Python
- Java
- C++

3.6 Hardware, Software, Technologies, Tools used

3.6.1 Hardware used - Raspberry Pi 4

- On account of the chosen algorithms for processing an image, the processor specifications of Arduino UNO and Raspberry Pi Zero are on the lower side. They are unable to handle the load which is concomitant of using Neural Network based algorithms
- Although Raspberry Pi 3 is capable of handling the load owing to the chosen algorithms, the response time is somewhere around 4-5 secs which is unacceptable in the real world.
- Jetson Nano is the best alternative, however it is priced exorbitantly and needs to be imported. It is priced at $99, which if used will raise the price of the end product.
- Raspberry Pi 4 is the optimum solution when it comes to both performance and price of the product

3.6.2 Algorithm used - YOLO Object Detection

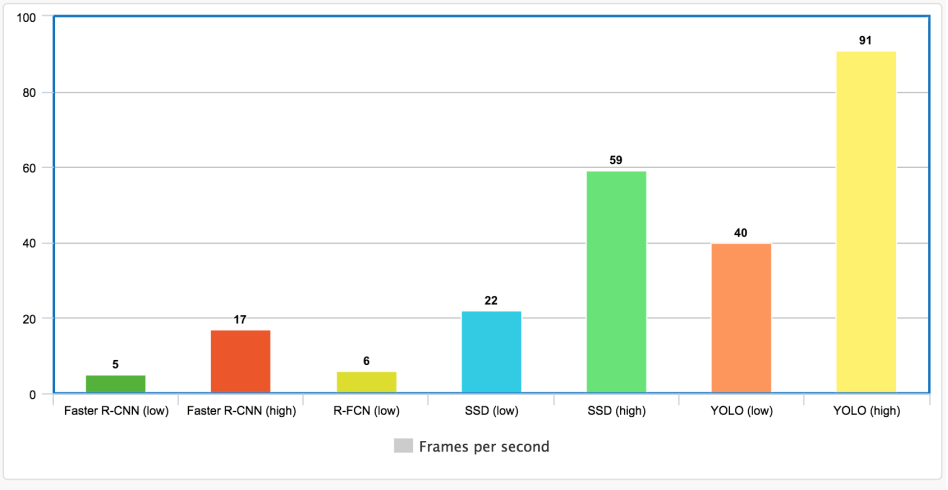

Fig. 3.5.2.1 FPS Comparison of various Object Detection Algorithms

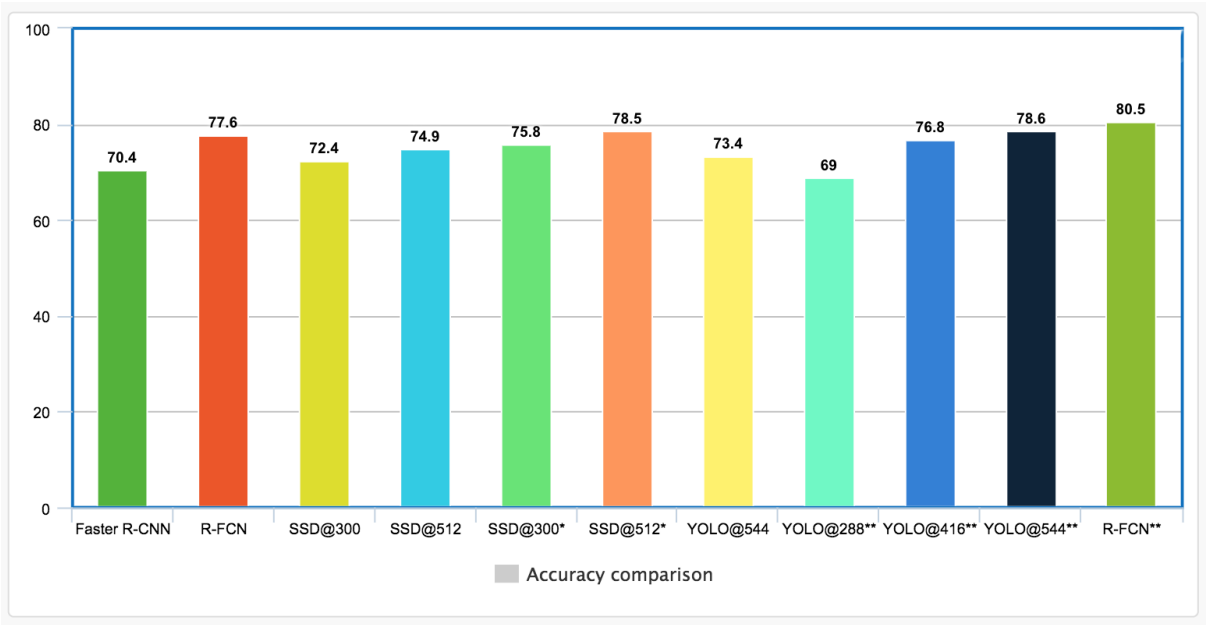

Fig. 3.5.2.2 Accuracy Comparison of various Object Detection Algorithms

3.6.3 Technologies/Tools used - Python

Python is easy to use and has inbuilt optimised A.I libraries.

Chapter 4: Proposed Design

4.1 Block Diagram

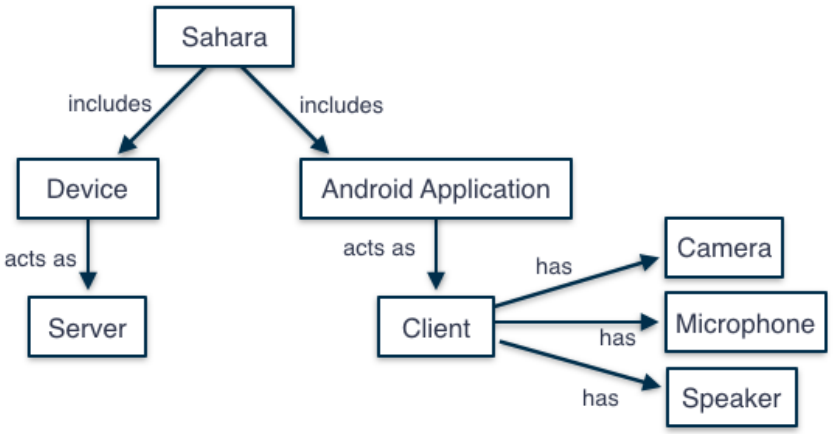

Fig. 4.1.1 Block diagram of the system

4.2 Modular Design

4.2.1 Text Summarization

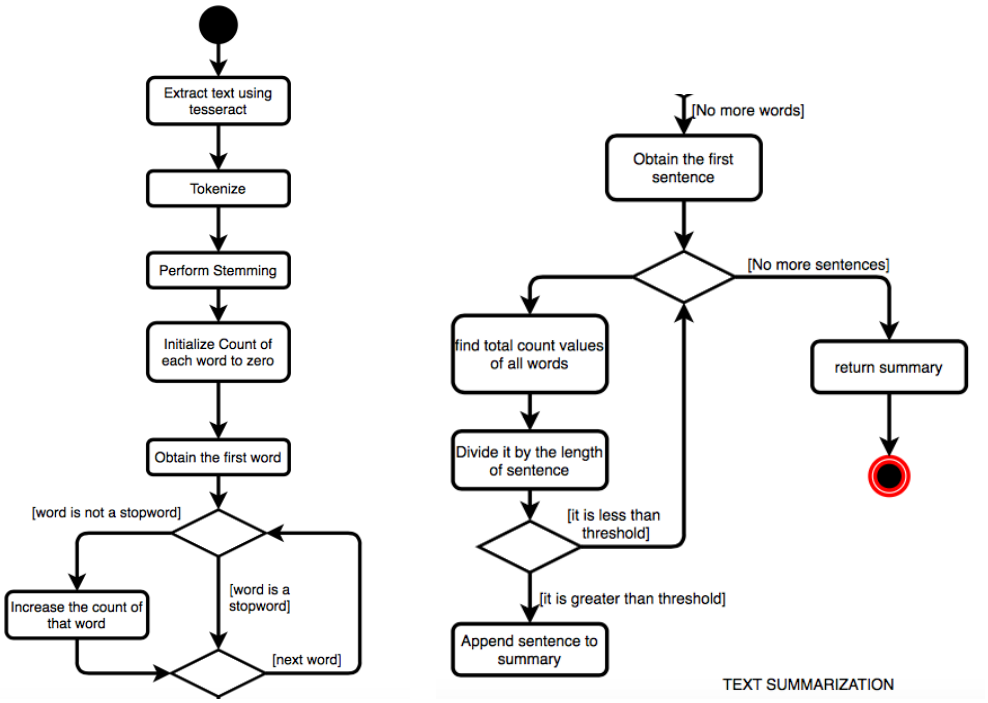

Fig. 4.2.1.1 Activity Diagram - Text Summarization

4.2.2 Currency Recognition and Counting

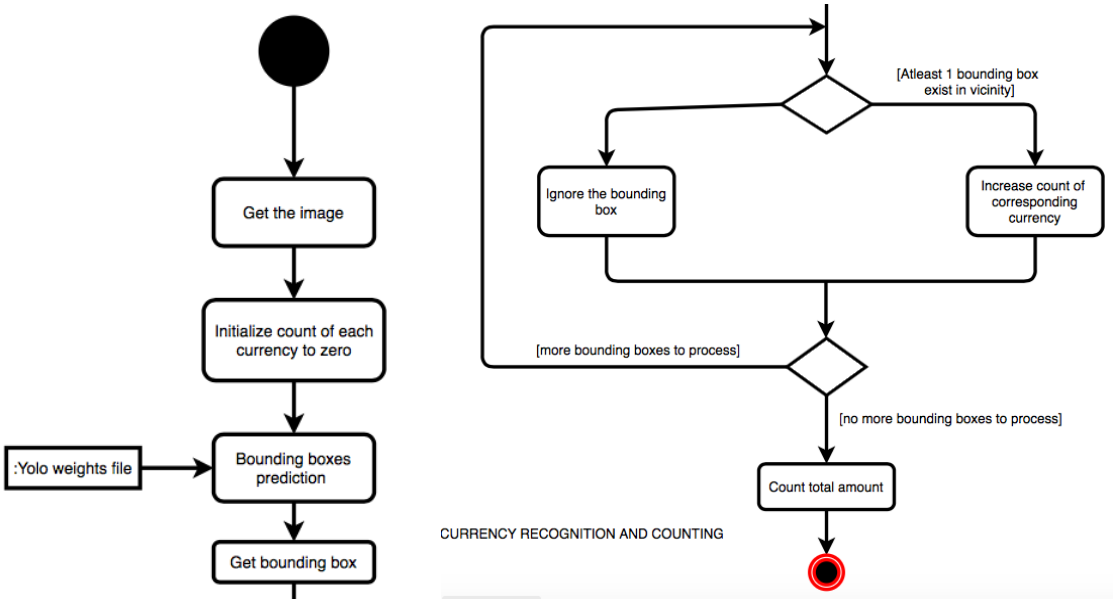

Fig. 4.2.2.1 Activity Diagram - Currency Recognition and Counting

4.2.3 Color Identification

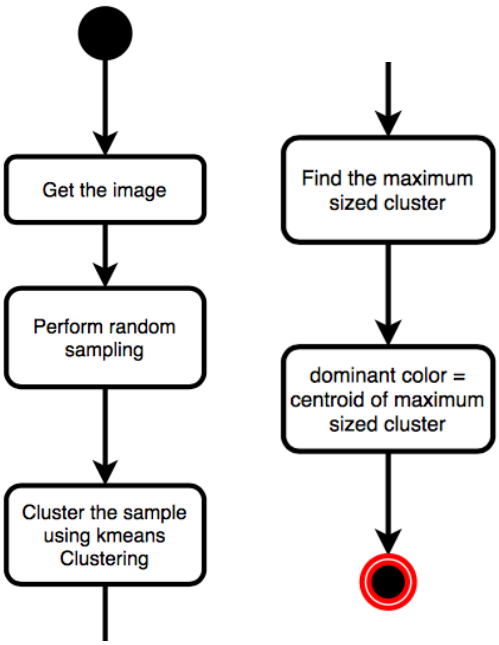

Fig. 4.2.3.1 Activity Diagram - Color Identification

4.3 Detailed Design

4.3.1 Use Case Diagram

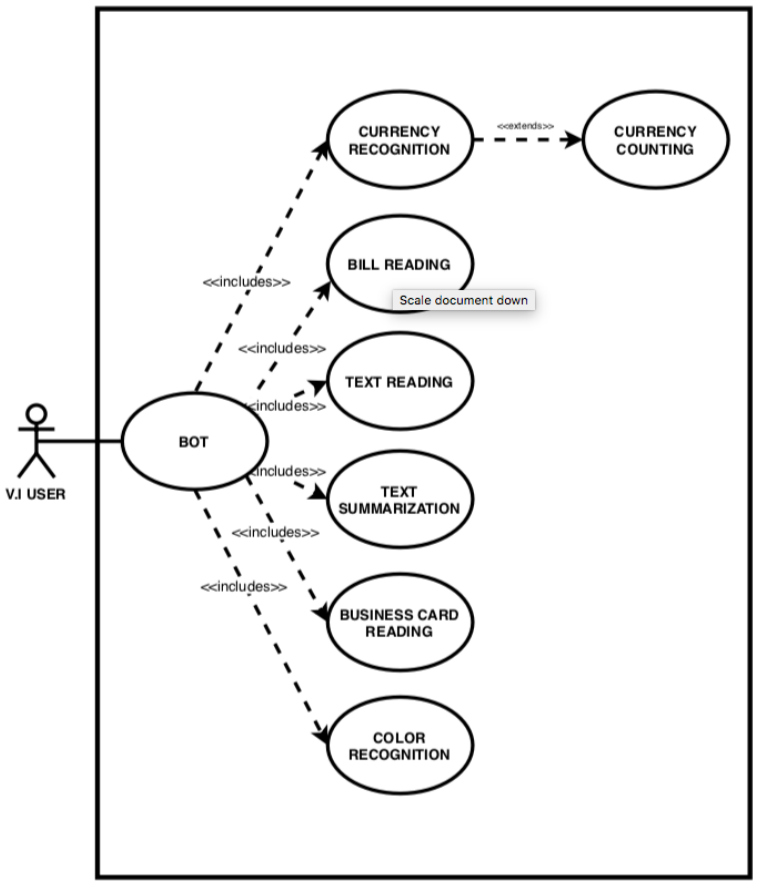

Fig. 4.3.1.1 Use Case Diagram

4.3.2 Data Flow Diagram

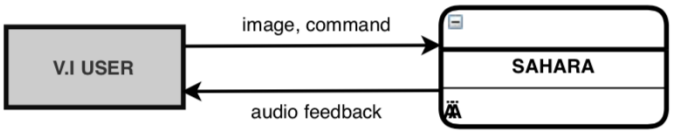

Fig. 4.3.2.1 DFD - Level 0

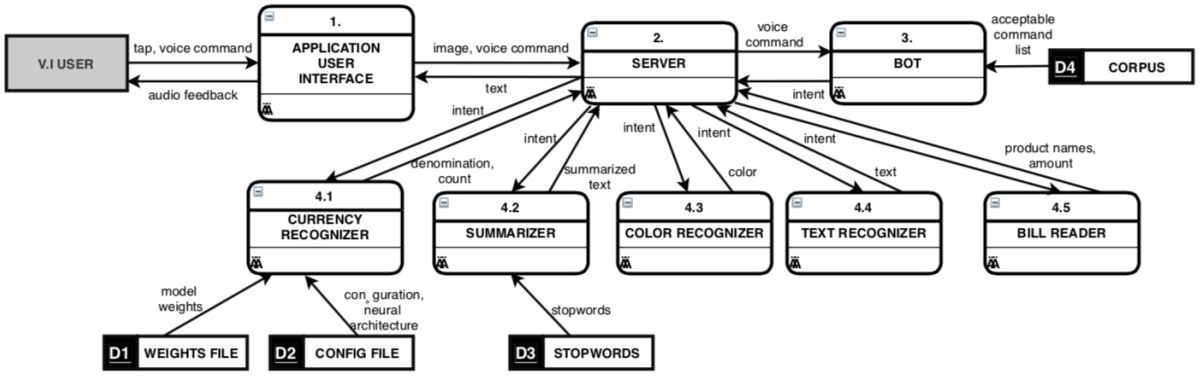

Fig. 4.3.2.2 DFD - Level 1

4.3.3 Class Diagram

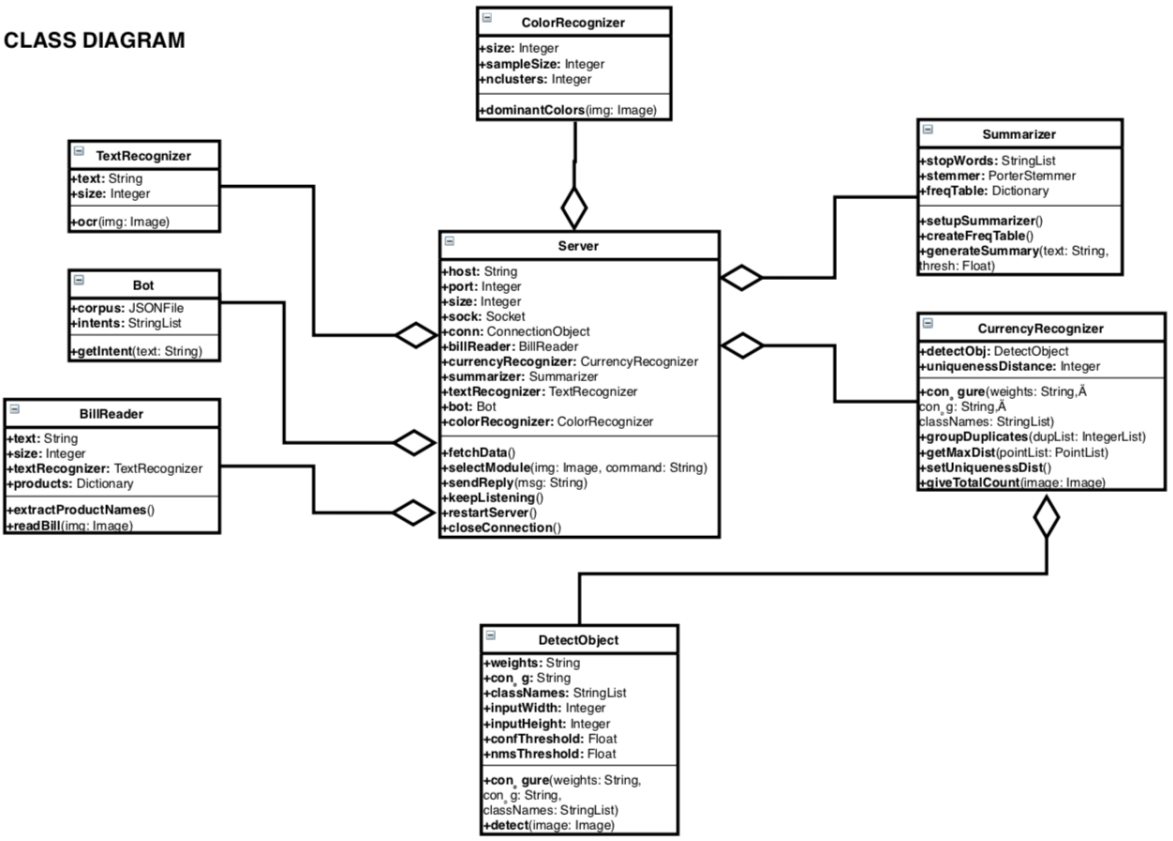

Fig. 4.3.3.1 Class Diagram

4.3.4 Sequence Diagram

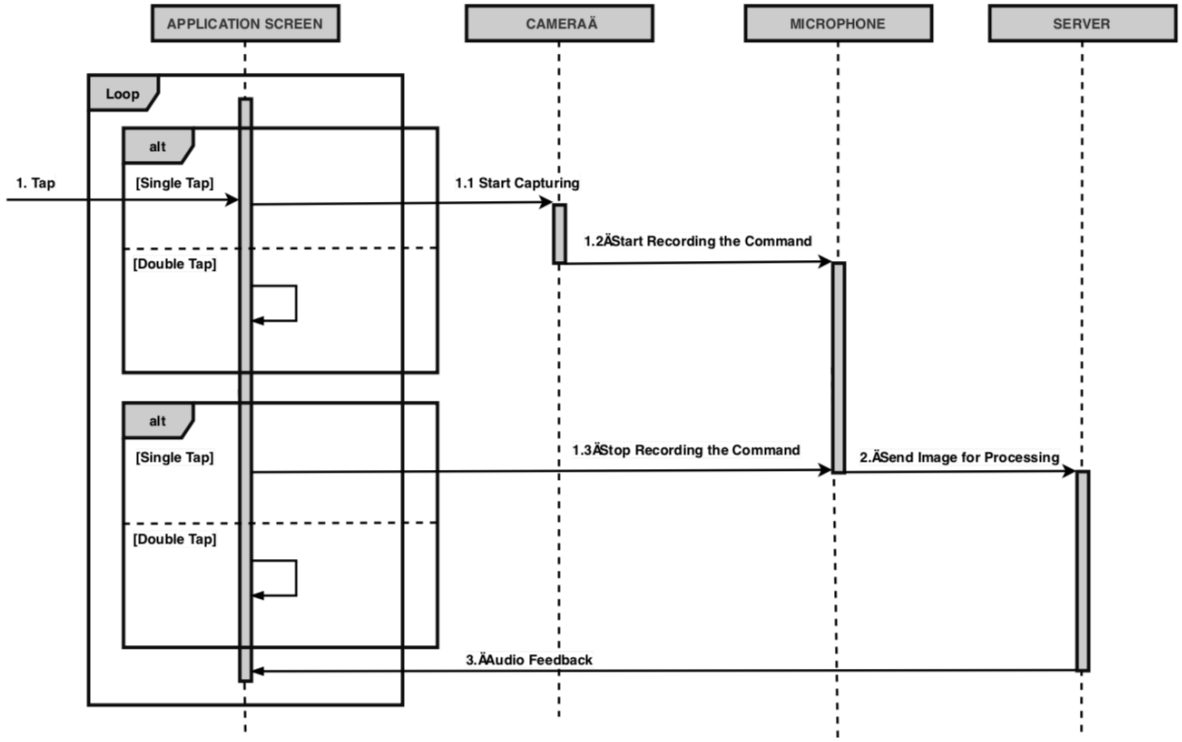

Fig. 4.3.4.1 Sequence Diagram

4.3.5 Connection Diagram

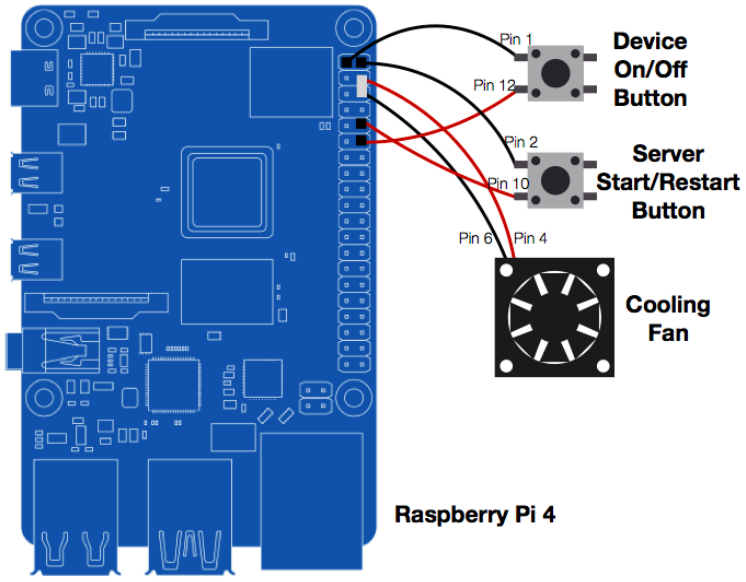

Fig. 4.3.5.1 Connections

4.3.6 3D Model

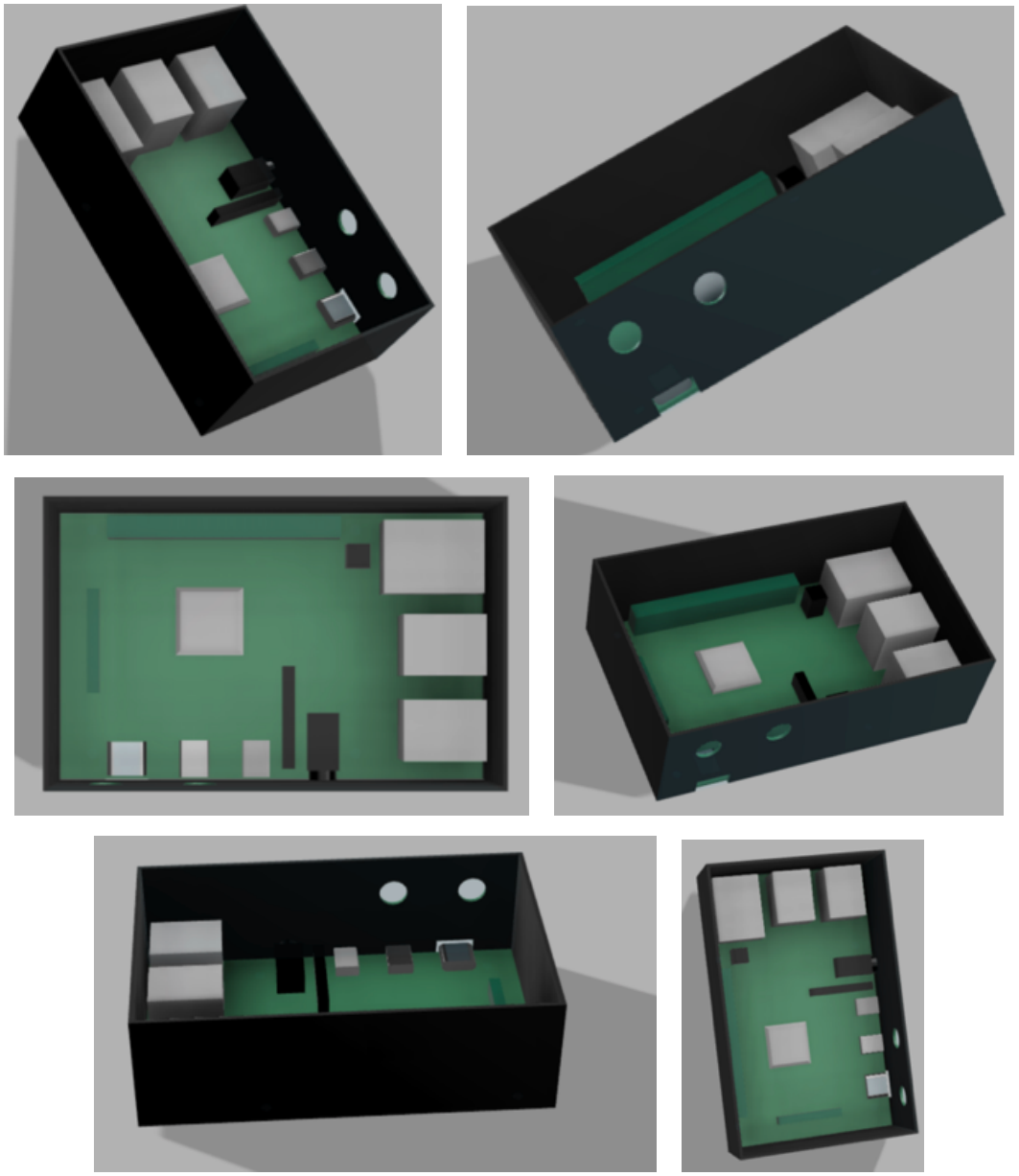

Fig. 4.3.6.1 3D Views

4.4 Project Scheduling and Tracking

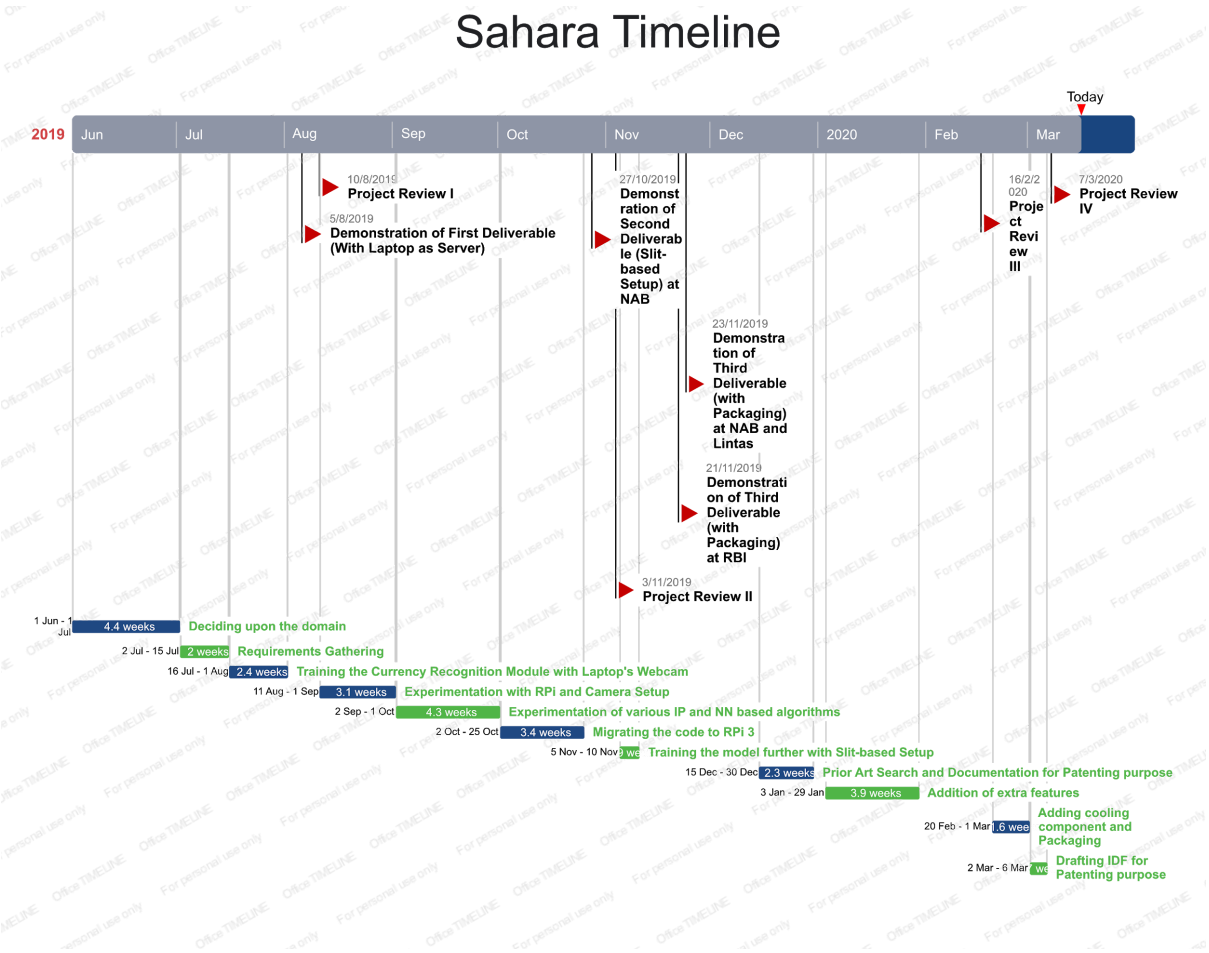

Fig. 4.4.1 Project Timeline

Chapter 5: Implementation Details

5.1 Implementation Details

5.1.1 Android Application

5.1.1.1 Description

The Sahara android application requires the mobile hotspot to be switched on. Once the hotspot is switched on, the device (server) is started and it gets automatically connected to the mobile's hotspot. Note that the internet need not be switched on. Once the device (server) is connected to the phone, the user gets an audio feedback stating that the server is now connected and is ready to function. The connection between the device and the phone obeys the TCP/IP networking protocol. The Sahara android application has just one screen containing only the camera preview. There are no buttons, no menu bars, no tool-bars for interaction. The interaction between the user and the application is solely based on single and double screen-taps and voice input.

5.1.1.2 Functioning

- A single tap anywhere on the screen captures the image and the application starts listening to the voice command of the user.
- On the second single tap, the application stops listening. It then converts the user's voice input into text using Google's speech-to- text engine.
- This text along with the captured image is bundled into a packet which is then sent to the device (server) over the network (established on account of the mobile hotspot) using TCP/IP networking protocol.
- Once the packet reaches the device (server), it processes the same and returns a text output.
- This text output is again bundled into a packet and sent to the application over the network.
- Upon receiving the packet, the application uses Google's text-to- speech engine to synthesize audio feedback from the received text.
- A double tap on the screen is used to cancel the operation in case a user commits a mistake and realises it.

5.1.2 Device

5.1.2.1 Currency Recognition and Counting

A dataset of 700 images per currency is trained using YOLO object detection algorithm which results in a file containing the final weights of YOLO's neural network. This weights file is then utilised by the OpenCV's 'dnn' module which further extracts and stores the weights from it. Given a test image, the python program (for currency recognition) runs YOLO algorithm using the weights extracted by the 'dnn' module and predicts the bounding boxes for the english denominations in a currency. Every bounding box is associated with a class. Thus the denomination of the currency in the image is the class of the bounding box that is detected.

The system allows the user to hold multiple currencies in hand in a deck- of-cards fashion and returns the individual denomination along with the total count of the currencies in hand. This feature again uses YOLO and OpenCV's dnn module for detecting the bounding boxes around the individual denominations. The bounding box with zero bounding boxes in its vicinity (decided by the uniqueness threshold) is eliminated. The remaining bounding boxes represent the denominations of the individual currencies.

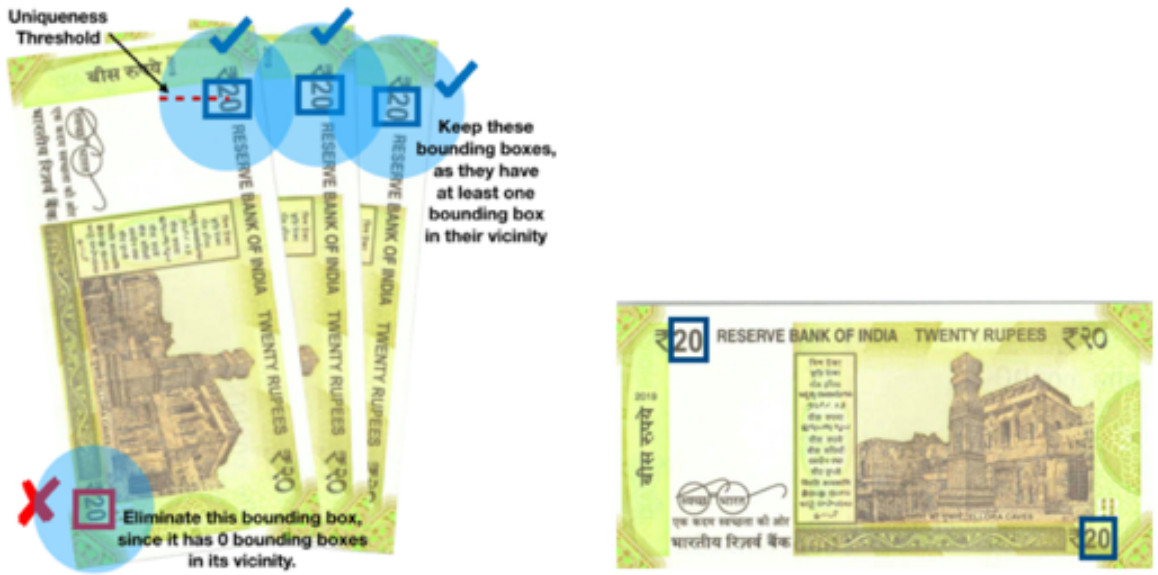

Fig. 5.1.2.1.1 Diagramatic Representation of working of Currency Recognition and Counting

5.1.2.2 Basic Text Reading

This module uses python's Pytesseract library to extract the text from the input image. Pytesseract is an open-source OCR (Optical Character Recognition) engine. However, it comes with a limitation that the orientation of the text should be proper enough for its detection. The system does preprocessing on the extracted text to determine the count of words which actually make sense (words which are present in the dictionary). If this count, exceeds the system-defined threshold, then the text is accepted, else the user is given a prompt to reorient the text.

5.1.2.3 Bill Reading

The Basic Text Reading module returns the text extracted from the input image. Upon receiving the text, the python's regex (Regular Expressions) library converts the plain text into records splitted by a new line. If the number of records with FOUR floating point numbers (Qty, MRP, Rate, Amount) exceeds the system-defined threshold, then the image can be considered to be a supermarket bill and can be processed further. Once the system determines that the image is of a supermarket bill it applies various permutations (4P2) and checks if the multiplication of two floating point numbers matches one of the four floating point numbers. Based on this, the module finds the mapping of the four floating point numbers to Qty, MRP, Rate and Amount. The remaining text apart from these 4 numbers is considered to be the product name.

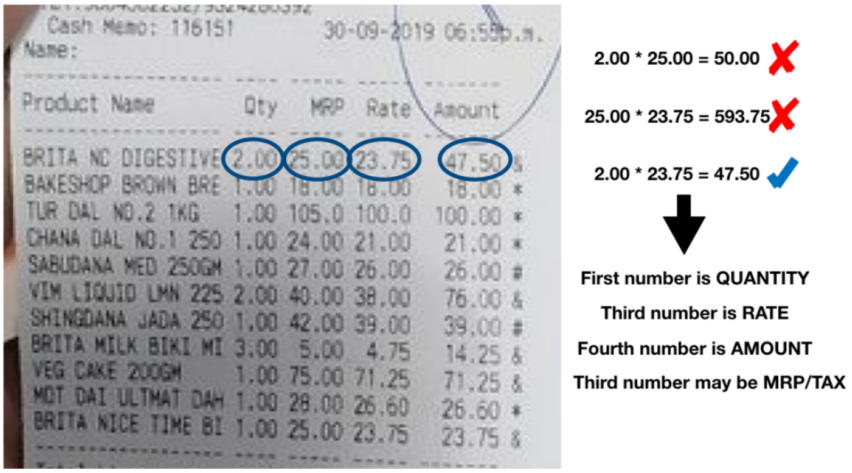

Fig. 5.1.2.3.1 Diagramatic Representation of working of Bill Reading

5.1.2.4 Text Summarization

The Basic Text Reading module returns the text extracted from the input image. The general text summarization algorithm is then applied on this extracted text to give a succinct output.

This feature is useful when the user doesn't want the system to read the entire text, rather wants to know just the essence of it. Given the input, the text is sentence-tokenized and word-tokenized followed by stopword removal (removal of words like a, an, is, the, etc). A frequency table is then generated from the preprocessed text which stores the frequency of every word in it. The frequency table generation is then followed by sentence scoring wherein each sentence in the input text is assigned a score. For every word in the sentence, its frequency is added together. The resulting sum from all of the words in that sentence is then normalized by dividing it with the word count which results in the sentence score. Sentences with scores lesser than a system-defined threshold are eliminated, while those which exceed it are included in the resultant summary.

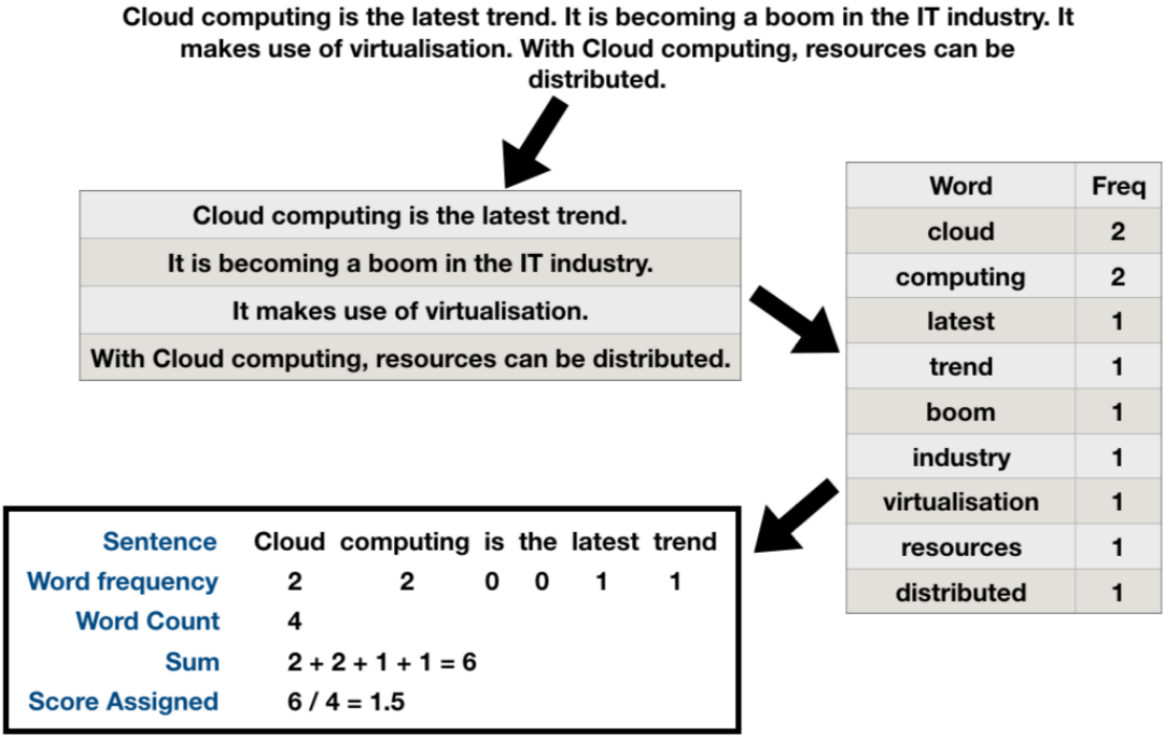

Fig. 5.1.2.4.1 Diagramatic Representation of working of Text Summarization

5.1.2.5 Business Card Reading

The Basic Text Reading module returns the text extracted from the input image. SpaCy's pretrained Entity detection model is used to extract important information from the text like email-id, mobile number, address, etc.

5.1.2.6 Color Identification

Color identification uses K-means clustering at its core. Given an image, random sampling is performed to improve the speed-up factor and these randomly selected pixels are then clustered by the K-means algorithm using python's sklearn library. The maximum sized cluster is representative of the dominant color in the image and its centroid gives the dominant color in the image.

5.1.2.7 Functioning of Server

- Server receives the packet sent by the Sahara android application.
- It separates the image and the text that was packaged together.
- The text is sent to the Bot module which returns the intent i.e the module to which the image should be sent for processing.
- Upon determining the module, the server sends the image to that module for processing.
- The module processes the image and returns the result as a text.
- This text is then sent to the Sahara android application which generates audio feedback.

Chapter 6: Testing

6.1 Definition of Testing

Software testing is a process, to evaluate the functionality of a software application with an intent to find whether the developed software met the specified requirements or not and to identify the defects to ensure that the product is defect free in order to produce the quality product.

6.2 Types of Testing

6.2.1 White Box Testing

It is also called as Glass Box, Clear Box, Structural Testing. White Box Testing is based on applications internal code structure. In white-box testing, an internal perspective of the system, as well as programming skills, are used to design test cases. This testing is usually done at the unit level.

6.2.2 Black Box Testing

It is also called as Behavioral/Specification-Based/Input-Output Testing. Black Box Testing is a software testing method in which testers evaluate the functionality of the software under test without looking at the internal code structure.

6.2.3 Grey Box Testing

Grey box is the combination of both White Box and Black Box Testing. The tester who works on this type of testing needs to have access to design documents. This helps to create better test cases in this process.

6.3 Type of testing considered

The product is tested using the Grey Box testing approach. On account of the usage of built-in libraries for development of the system, there is less room for internal code errors.

Thus functionalities provided by the system should be tested under rigorous conditions to check the performance of the system under those conditions. However, for deriving the test cases for testing the functionality, there is a need to look into the internal logic used and the design of the system. Thus Grey Box testing is a perfect fit for testing the system.

6.4 Various Test Case Scenarios

- Changing the size of the input image captured with the Camera
- Changing the orientation of the images containing text concerning modules involving text recognition
- Illumination variation
- Handling currencies with natural tendency

6.5 Inference drawn from the tests

- The input size of the image should be 500 x 500. At this resolution, the accuracy given by PyTesseract is the highest. Also YOLO network takes a square shaped input image. Thus the input image must have 1:1 Width:Height ratio.
- PyTesseract is sensitive to orientation of the image and gives the best accuracy if the text imprinted object is held straight.
- Illumination variation doesn't have much effect on the functioning of the system on account of the usage of Auto-Flash mode.
- The system gives fair accuracy when the currencies are held according to natural tendency.

Chapter 7: Conclusion

7.1 Limitations

- Currently the system requires external power supply.
- The device is dependent on the smartphone.
- User training is required.

7.2 Conclusion

Sahara is a portable, internet-free device which acts as a server for processing the data. Sahara can be used with any smartphone having a decent camera. Thus buying Sahara along with a decent smartphone is quite pocket-friendly.

7.3 Future Scope

- Use of Solar panels to make the device sustainable.
- Use of Jetson Nano and a high-end camera to make the device completely independent.
- Document classification will eliminate the need for the user to give voice commands in order to select the module to be called.
- Voice-based calculator can be an add-on.
- Semantics in text recognition can be captured.

References

- https://docs.google.com/document/d/1hHucV3ZM-ejsyf12HvrByb6VUIa3GNO_2fJIYYzL7Kg/edit#

- https://www.financialexpress.com/industry/technology/rbi-launches-mani-app-to-help-visually-challenged-identify-currency-notes-how-to-download-it/1810477/

- *R. Kasthuri, B. Nivetha, S. Shabana, M. Veluchamy and S. Sivakumar, "Smart device for visually impaired people," 2017 Third International Conference on Science Technology Engineering & Management (ICONSTEM), Chennai, 2017, pp. 54-59.*

- *K. G. Krishnan, C. M. Porkodi and K. Kanimozhi, "Image recognition for visually impaired people by sound," 2013 International Conference on Communication and Signal Processing, Melmaruvathur, 2013, pp. 943-946.*

- *A. KARTHIK, V. K. RAJA and S. PRABAKARAN, "Voice Assistance for Visually Impaired People," 2018 International Conference on Communication, Computing and Internet of Things (IC3IoT), Chennai, India, 2018, pp. 465-468.*

- *B. Deepthi Jain, S. M. Thakur and K. V. Suresh, "Visual Assistance for Blind Using Image Processing," 2018 International Conference on Communication and Signal Processing (ICCSP), Chennai, 2018, pp. 0499-0503.*

- *N. A. Semary, S. M. Fadl, M. S. Essa and A. F. Gad, "Currency recognition system for visually impaired: Egyptian banknote as a study case," 2015 5th International Conference on Information & Communication Technology and Accessibility (ICTA), Marrakech, 2015, pp. 1-6.*

- https://patents.google.com/patent/US8050484B2/en?oq=US8050484B2

- https://patents.google.com/patent/CN101814128A/en?oq=CN101814128A
- https://patents.google.com/patent/US8606316B2/en?oq=US8606316B2

Appendix

1. RBI MOM

Reserve Bank of India, Amar Building

November 21, 2019

Timing: 11.00 am - 12.30 pm

Location: Department of Currency Management

Members Present :

1. Mrs. Sharmila Sengupta
2. Mr. Dhaval Bagal
3. Ms. Shreya Patil

Members Apologies :

1. Mr. Amey Pimple
2. Mr. Rahul Devadiga

Minutes :

Today's meeting was concerned with the demonstration of the product "VIVENAB" to the RBI.

The subsequent points were discussed:

1. Following improvements were suggested by the Chief General Manager-in-Charge Shri Ajay Michyari:
    i) Fixing a battery on to the device.
    ii) Reducing the dimensions of the device and making it more compact.
    iii) Either increasing the dimensions of the slit or switching to a real-time approach.
    iv) Making provisions for currencies of other countries as a future scope.
    (v) Extending the device for hearing impaired people.

2. Other than this, Shri Ajay Michyari and his associates further suggested focussing more on the business aspects of the developed product.

3. The talk also included the ways in which the developed product can be taken further.
4. It concluded that an app would involve no cost and therefore is an easier alternative to a device. The difficulties faced by vision impaired to handle smartphones were explained to them, but did not create an impact in the discussions.

Minutes Prepared By: Mrs. Sharmila Sengupta

2. RBI Gate Pass

G30

Reserve Bank of India

Amar Building - Mumbai

Visitor Pass K21

Pass No:	309
Name:	BEGAL DHAVAL ULHAS +2
Organization	VIVEKANAND EDU SOC INS. TECH
Purpose:	OFF
To Meet:	ADMIN
Designation	AM
Department	Dept of Currency Mgmt
Floor #	4
Date-Time:	21/11/19 11:03AM
Valid upto	**21/11/2019** **Time Out**
Remarks	

टिप्पणी / **Note**

1. यह प्रवेशपत्र हस्तांतरणीय नहीं है। उसे जब भी मांगा जाए, प्रस्तुत करना आवश्यक है।

   The pass is not transferable & must be produced whenever demanded.

2. इस प्रवेशपत्र पर जिस अधिकारी से मिलें उनके हस्ताक्षर करवाएं तथा जाते समय इसे लौटा दें।

   Have this pass signed by the officer visited & return while leaving.

3. यदि किसी व्यक्ति को यह प्रवेशपत्र मिले तो कृपया इसे क्षेत्रीय निदेशक, भारतीय रिज़र्व बैंक, मुंबई 400 001 को भेज दें।

   Anyone finding this pass is requested to send it to the Regional Director, Reserve Bank of India, Mumbai - 400 001.

   टेलीफोन / Telephone : 022-2266 0500, 2267 6950

3. NAB MOM

National Association for the Blind, Worli Seaface.

November 23, 2019

Timing: 12.00 pm - 2.30 pm

Location: NAB Conference Room (Ground floor)

Members Present :
Mrs. Sharmila Sengupta (VESIT)
Ms. Shreya Patil (VESIT)
Mr. Rahul Devadiga (VESIT)
Mr. Amey Pimple (VESIT)
Mr. Dhaval Bagal (VESIT)
Mr. Amar Jain (Lawyer)
Mr. Harish Kotian (Assistant General Manager Of RBI)
Mr. Sumanta Ganguly (Lintas)
Ms. Tania Dey (Lintas)
Mrs. Aparna Ambre (Lintas)
Mr. Vignesh Mani (Lintas)
Mr. Maaz Ansari (Lintas)
Mr. Suhas Karnik (NAB)
Mrs. Pallavi Kadam (NAB)
Mr. Arvind Narvekar (NAB)
Mr. Mangesh (Client-service department-NAB)

Minutes :

Today's meeting was concerned with the demonstration of the product "VIVENAB" to Lintas (Ad.Agency).

The subsequent points were discussed:

1. Following improvements were suggested by the team:
    i) Fixing a battery on to the device.
    ii) Reducing the dimensions of the device and making it more compact.
    iii) Cheques, Invoice, Bills Reader, etc (numerical part) along with currency recognition.
    iv) Connectivity with the phone.
    v) A slit wherein the note would be slid entirely and image stitching could be incorporated.
    vi) Using a technology similar to a sliding pen for recognizing the currency.

2. Other than this, some of the questions asked were:
    - What would be the height of the end product?
    - Who would take the product further in the future?
    - Is any design team present in the college to design the product?

3. Also, a suggestion was put forward to develop a wearable platform instead of a slit based device for the user interaction is of great value when it comes to a product for the visually impaired.

4. Talking about the cost factor, the team from Lintas advised to not ponder much over the cost since, on account of the mass production, the price of the product will drop considerably. This is because, during production, profuse optimizations can be done.

5. Lintas Ad. Agency has put forward a helping hand for the following things:
    - Designing process of the product.
    - Help from a technical team w.r.t any aspects of the end product.
    - Funding if required from YES Bank, Axis Bank, etc.

6. Few existing products were discussed:
    - Look it up - Google Lens
    - Ibill (US Currency Detector) developed in Baroda.
    - Go-pro USB Camera (360 degrees).

7. Following suggestions were posited by Mr. Maaz:
    - Using an image sensor instead of the camera since it would reduce the dimensions of the device considerably.
    - Developing an SoC for the concerned product, since raspberry pi cannot be taken further in production.
    - Instead of training models and burning them into the hardware, developing a learning agent that would learn and adapt itself slowly to the needs of the blind.
    - A rudimentary SoC would be an 8085 with the program burnt in it, however, the only issue is the processing speed.
    - Designing an MVP (Most Viable Product) focussing only on the currency recognition part.
    - Drilling into the component level reduces the cost of production considerably. For e.g: An INR 3.2k costing Raspberry Pi 4 would not cost more than INR 500 when manufactured.
    - Developing an SoC with only the components that are required by the algorithms and chipping of all the other unnecessary segments.

Minutes Prepared By: Mrs. Sharmila Sengupta

---

4. Demonstration at NAB