

**VOICE CONTROL SYSTEM OF GOOGLE APPLICATIONS AND
IMAGE DETECTION FOR VISUALLY IMPAIRED.
A MINI PROJECT REPORT**

Submitted by

C.ABINAYA 1805066

M.DINESHKANNA 1805073

M. MANIKANDAN 1805098

D. ROOBA DEVI 1805113

in partial fulfilment for the award of the degree of

BACHELOR OF ENGINEERING

In

COMPUTER SCIENCE AND ENGINEERING



COIMBATORE INSTITUTE OF TECHNOLOGY
*(Government Aided Autonomous Institution Affiliated to Anna
University)*

**COIMBATORE-641 014 ANNA UNIVERSITY
CHENNAI 600 025 MARCH 2021**

COIMBATORE INSTITUTE OF TECHNOLOGY

(A Govt. Aided Autonomous Institution Affiliated to Anna University)

COIMBATORE – 641 014

BONAFIDE CERTIFICATE

Certified that this mini project “**VOICE CONTROL SYSTEM OF GOOGLE APPLICATIONS AND IMAGE DETECTION FOR VISUALLY IMPAIRED**” is the bonafide work of **C.ABINAYA (1805066), M.DINESHKANNA (1805073), M.MANIKANDAN(1805098), D.ROOBA DEVI(1805113)** under my supervision during the academic year 2020-2021.

Dr.G.Kousalya, M.E.,Ph.D
HEAD OF THE DEPARTMENT,
Department of CSE
Coimbatore Institute of Technology,
Coimbatore – 641 014.

Ms.A.PUNIDHA, M.E.,
ASSOCIATE PROFESSOR,
Department of CSE,
Coimbatore Institute of Technology,
Coimbatore – 641 014.

Certified that the candidates were examined by us in the project work viva-voce examination held on

Internal Examiner

External Examiner

Place:

Date:

TABLE OF CONTENTS

CHAPTER NO.	TITLE	PAGE NO
	ACKNOWLEDGEMENT	4
	ABSTRACT	5
	LIST OF FIGURES	6
	LIST OF ABBREVIATIONS	7
1	INTRODUCTION	8
	1.1 INTRODUCTION	8
2	LITERATURE SURVEY	9
	2.1 LITERATURE SURVEY	9
	2.2 INFERENCES FROM LITERATURE SURVEY	16
	2.3 SCOPE OF THE PROJECT	16
	2.4 LIMITATIONS CONSIDERED IN THE PROJECT DESIGN	16
3	SYSTEM SPECIFICATION	17
	3.1 SOFTWARE SPECIFICATION	17
	3.2 HARDWARE SPECIFICATION	17
4	SYSTEM ANALYSIS	18
	4.1 EXISTING SYSTEM	18
	4.2 PROPOSED SYSTEM	18
	4.3 METHODOLOGIES	19
5	SYSTEM IMPLEMENTATION AND RESULT	
	5.1 OVERVIEW	22
	5.2 MODULES	22
6	CONCLUSION AND FUTURE WORK	36
7	REFERENCES	37

ACKNOWLEDGEMENT

ACKNOWLEDGEMENT

We express our sincere thanks to our Secretary **Dr.R.Prabhakar** and our Principal **Dr.V.Selladurai** for providing us a great opportunity to carry out our work. The following words are a small part to express our gratitude to them. This work is the outcome of their inspiration and product of plethora of their knowledge and rich experience.

We record the deep sense of gratefulness to **Dr.G.Kousalya**, Head of the Department of Computer Science and Engineering, for his encouragement during this tenure.

We equally tender our sincere thankfulness to our project guide **Ms.A.Punidha**, Assistant Professor, Department of Computer Science and Engineering ,for her valuable suggestions and guidance during this course.

During the entire period of study, the entire staff members of the Department of Computer Science and Engineering & Information Technology have offered ungrudging help. It is also a great pleasure to acknowledge the unfailing help we have received from our friends.

It is a matter of great pleasure to thank our parents and family members for their constant support and cooperation in the pursuit of this endeavour.

ABSTRACT

ABSTRACT

People with visual impairment and blindness face a lot of challenges in their daily lives in adopting this modern era where everything is digitalised. The main aim of this project is to provide voice-based control system to access the google search engine, emails, and navigation map for blind people using speech recognition module and it is intended to provide object detection as real time assistance to identify the objects(things) around them. It has the capability of access to World Wide Web by browsing the internet, checking, sending and receiving email, listening to the content of search only by giving a voice command to the system. The visually impaired person will give the destination's name as the input to speech recognition module. The blind person receives the pronounced direction which he needs to follow to reach his destination. With help of the camera the person can detect the objects around them in their house and on the road. Such technology enhances the functioning and performance of daily living skills, thereby improves the independent living. Over time the need as well as demand of such technologies will continue to grow due to expected increase in proportion of population with visual disability particularly in low middle-income countries like India. Therefore, this service for visual disability in India is to create awareness, sensitize among eye care professionals, potential beneficiaries and their families, to make their lives also easier and to confront the newest technology.

Keywords: Voice control system, Speech recognition ,Real time assistance ,Object Detection, Visual impairment

LIST OF FIGURES

FIGURE NO.	TITLE	PAGE NO.
5.1	THE WEATHER FORECAST	23
5.2	YOU TUBE SEARCH	23
5.3	YOUTUBE RESULT PAGE	24
5.4	SENDING EMAIL	25
5.5	OUTPUT OF GMAIL PAGE	25
5.6	DETAILS OF UNSEEN MAIL	26
5.7	OUTPUT PAGE OF UNSEEN EMAIL	26
5.8	GOOGLE SEARCH	27
5.9	OUTPUT OF THE GOOGLE PAGE	28
5.10	PLAY MUSIC THROUGH VOICE COMMAND	28
5.11	DETECTING OBJECT PAGE	29
5.12	CALCULATOR	30
5.13	NEWS HEADLINES	30
5.14	TAKING PHOTOS AND VIDEOS AND STORE IN FOLDER	31
5.15	TAKING AND STORING NOTES IN FOLDER	31
5.16	GUI INTEGRATION PART	32
5.17	BATTERY RESULT	33
5.18	OUTPUT OF SIMPLE CONVERSATION PART 1	33
5.19	OUTPUT OF SIMPLE CONVERSATION PART 2	34
5.20	OUTPUT OF SIMPLE CONVERSATION PART 3	35
5.21	WIKIPEDIA	35

LIST OF ABBREVIATIONS

GUI – Graphical User Interface
VUI - Voice User Interfaces
SSML - Speech Synthesis Markup Language
pyttsx3 - Python text-to-speech library
SRE - Speech Recognition Engine
SR - Speech recognition
STT – Speech To Text
GTTS – Google Text To Speech
API – Application Programming Interfaces
SMTP – Simple Mail Transfer Protocol
IMAP – Internet Mail Access Protocol
YOLO – You Only Look Once

INTRODUCTION

CHAPTER 1

INTRODUCTION

In the present world with full of upcoming technologies, It is very difficult to understand and percept the new technologies, so think of blind person situation ,It is very strange and non understandable to him. In order to supplement the virtual assistance, this project was created in aim of serving the need to them. The main domain of this project is voice recognition and image processing . All the features will be implemented only through the voice commands given by the blind one, Then the actions are performed according to that in order to provide ease to them. Giving the tough fight in the competitive world, we provide basic features like searching google, and working with emails in professional world and entertainment stuffs like YouTube and many more features that deal with the system through their voice. The entire framework or platform behind this project is implemented in computer system. When the Blind person is walking in the road ,he finds lot of difficulties to travel along, so we try of giving some guidance by identifying the objects and the path to travel to his destination. Anyway we are encapsulating the most requirements of the blind person to attain the satisfaction behind the reason of the project.

LITERATURE SURVEY

CHAPTER 2

2.1 LITERATURE SURVEY

2.1.1 Digital Assistant for the visually Impaired

Author: Ezekiel Marvin 2020

Voice User Interfaces (VUI) is the primary or supplementary visual, auditory, and tactile interfaces that enable voice interaction between people and devices. Android Studio IDE3 is used as *Programming Environment*. *Speech* recognition module is used. Google Cloud Speech to Text API. Google Mobile Vision Text Recognition API is used for *Object Character Recognition*. Google Cloud Text to Speech API in order to convert text or Speech Synthesis Markup Language (SSML). The methodological approaches include Interactive mode and Text detection mode. For Interactive mode, the application first start up, the user will be introduced to the digital A.I assistant named DAVID in which the user will mainly interact with the user in accordance with the appropriate keyword or command. For Text detection mode the application will access the smartphone camera and initialize the text detection module. By uttering the command “scan” or “scanning” the user will ask DAVID whether a text is detected in the camera overlay. In terms of accuracy, the application is better at perceiving user command in a quiet environment. The author would like to acknowledge that DAVID is not perfect and still a work in progress. Therefore, future works are needed to improve the system, such as the capacity for DAVID to tell the user whether they needed to shift the phone upward, downward, or zoom in or zoom out in order to better detect text. Future works could also implement new features such as object detection, face detection, text translation system, and low power consumption that will be added to the application.

2.1.2 Virtual assistant for the visually impaired

Author: Iyer, V., Shah, K., Sheth, S., & Devadkar, K. 2020

The software provides a way to interact with these websites with much ease. The software invokes by the keyboard shortcut. With the use of voice commands to access and provide commands to any website. The software will read out the content of the website and then using speech to text and text to speech modules along with selenium, the software can automate any website. The methodology followed to implement three modules are Main menu, Google, Gmail, Wikipedia. The Google library of speech-to-text (Speech Recognition) for Python is used. For communicating the system's output to the user, the recognized input is played back to the user using the Python text-to-speech library (pyttsx3). The modules are written in Python and make use of Selenium for automation of the respective module and BeautifulSoup for scraping the contents of the web page. The "Script" component of each module consists of the customized code that entails the features of the website contained in the module. At present the application supports only commands given in the English language and plan to expand that and make it available in most of the daily used languages thus people from all parts of the world can access the web without any issue and like to create a uniform framework that can be plugged to any website and create a browser extension thus making it possible to toggle between the two modes easily, especially for educational websites to enable visually impaired individuals to access online courses just like the average individual.

2.1.3 A Virtual assistant for visually Impaired Persons

Author: Sultan, M. R., & Hoque, M. M. 2019

Human voice commands are given to the application in English Language. This application model enables the user to control installed applications, send and receive e-mails or perform internet browsing. The model is also capable of reading a news portal and also allows the users accessing their home appliances over voice commands. The modules are Core Processing Module, Email Module ,News Reader Module, Web Surfing Module and Home Automation Module. A microphone takes voice commands as input. Speech Recognition Engine (SRE) converts the recognized speech into text. This converted text is then fed to the Core Processing Module (CPM). After that, the CPM judges the characteristics of the command and calls the relevant process for execution or redirects to the corresponding working module. In the end, the speech

synthesizer provides speaker output. At the moment, a user has to activate a module through the CPM over voice commands. Switching a module requires to handover the control to the CPM with voice commands. They are in the development stage of designing the system to run the modules in parallel threads to take the advantage of the multi-core system.

2.1.4 Object Recognition and Classification System for Visually Impaired

Author: Rashika Joshi, Meenakshi Tripathi., Amit Kumar ,Manoj Singh Gaur 2020

The system (It is intended to be used along with normal walking cane) consists of Jetson Nano ported with the trained deep learning model and is interfaced with camera, that acts as a easy-to-use platform for object recognition, speech processing and image classification .Google colab, Jetson Nano, Mobile Net, SSD, voice Assistant are used. Hardware description- Jetson Nano: low-powered, small computer development kit from Nvidia an AI computer that allows Artificial Intelligence's ability to be brought to resource-constrained devices and capable of running in parallel numerous neural networks. Hardware and Software integration Mobile Net SSD model that we obtained after training is ported into jetson nano via Sd card. Then this module is integrated with the RPi v2 camera. The Jetson Nano Developer Kit has a connector which is compatible with RPi camera. The device recognizes the object and sends information about the object to the user through headset, rather than simply detecting the object. ongoing research on the exploration of systematic design for these complex systems will aid in better analysis leading to optimal design decisions. Also they are focusing towards the cost-effectiveness of the system in order to be implemented on a large scale.

2.1.5 Human computer interaction (HCI) based Smart Voice Email(Vmail) Application – Assistant for visually impaired users(VIU)

Author: Sherly Noel 2020

Visually impaired users to send emails which converts speech to text and text to speech. This application is not only for visually impaired users but also can be well used by the normal users. The application works through voice commands which are simple and user friendly. Once the email is composed using text to speech, the user can

verify the dictated content through text to speech conversion. The application receives the input through suitable hardware and processes the voice. During processing, the voice input received is compared with the sample voice stored in the database. The Google Web Kit API recognizes and converts voice to text. The text output gets typed in the designated text box from the voice command received. Email Compose Application Modules are User Voice, Voice Processing, Comparing voice with pre-recorded sample and Voice to text conversions: “Compose Mail” opens the email composer window. “Edit To” place cursor in To text box and type recipient email id .“Edit Subject” place cursor in subject text box and type the subject. “Edit Body” place cursor in email text area and type the dictated content. “Mail Send” send the email to the mentioned recipient. “Mail Discard” discards the composed email. “Read mail” reads the complete active mail content. Send Email and Read Email. Instead of authenticating the user by using email ID and password, enhancements like fingerprint reader can be used. Additional voice commands like reply, forward, file attachment, starred etc., can be added to enhance the interface. Semantic suggestions and auto corrections can be added to the email composer to access better. More research and enhancements can be performed using NLP algorithms to recognize local languages. The stated enhancements boost user experience and it can be productive too.

2.1.6 A Review on Voice Based E-mail System for Blind

Author: Paulus A. Tiwari, Pratiksha Zodawan 2020

Automatic speech recognizer, screen reader, text to speech to speech to text, braille console, Voice enabled interface with expansion support for signal based information and yield approaches are for the "Social Robot Maggie ".Interactive voice response (IVR) is an innovation that enables a PC to associate with people using voice and DTMF tones input through a keypad. Speech recognition (SR) is the ordered sub-field of computational linguistics (CL) that generate techniques and advancements to empower the acknowledgment and interpretation of communicated in language into text by PCs. It is also known as “automatic speech recognition “(ASR),“computer speech recognition ", or only “speech to text“(STT) using web pages. Thus we proposed a voice based authentication instead of the traditional username and password. The framework growing presently is depend just on personal machine. With the use of technically advanced smartphones, such systems and applications has a chance to be implemented

as an App in smartphones. Thus there is a scope to implement the framework in various other languages rather than implementing it only in English language.

2.1.7 A Smart personal AI Assistant for visually impaired people

Author: Shubham Melvin Felix, Sumer Kumar, and A.Veeramuthu 2020

Artificial Intelligence, Machine Learning, Image and Text Recognition. The idea is implemented through Android mobile app that focuses on voice assistant, image recognition, currency recognition, e-book, chat bot etc. The proposed system uses Artificial Intelligence to assist the visually impaired people which is all based on voice command. It also does image recognition of the photographs clicked or uses camera to recognize the objects and describes them in audio and also a chat bot to have light and friendly conversations .Object recognition-hybrid algorithm ,Image reading using RGB colours. Google Cloud API architecture mainly use chat-bot client used for the vision as well as Dialog-Flow(ID3,IG Algorithm used). In future our proposed system will be able interpret the textual description in a much better way. The Image recognition can be enhanced with much more details about the image captured through the camera. Enhancement to this system can be done by adding the features of currency recognition. The existing methodology for image and currency recognition can be done with more accuracy. Use the application by their own languages.

2.1.8 A Visually Impaired Assistant System

Author: Partha, Devashish Pradeep Khairnar2020

The proposed VI (Visually Impaired)Assistant System is developed to assist the visually impaired people have four modules which are obstacle recognition, obstacle avoidance, indoor and outdoor navigation and real-time location sharing. Image Processing, Arduino, Indoor Navigation, Speech Commands, indoor localization, visually impaired people, Obstacle detection, Ultrasonic sensors. Detection of obstacles, we have built a completely independent hardware module (smart glove) which uses proximity sensors and vibration motors to find obstacles as well as the distance between user and obstacle. Object detection module prompts the names of objects in front of the

user. Applications like google maps only provide outdoor navigation(Wi-Fi, tensor flow).Real-time location sharing can be used to track a visually impaired person in an outdoor environment by his relative.(Firebase Cloud Messaging service). Current indoor navigation system requires training for location learning and navigation to be done by an administrator. The system can be enhanced by dynamic training for indoor navigation module by tracking daily movements of other normal people. Real-time scene description can be provided by capturing images around the visually impaired person. In the current system, the path between two different places is set manually by an administrative person. The system can be improved by using the shortest path finding algorithms like Dijkstra's algorithm and representing interconnection between places using a weighted graph.

2.1.9 An AI-Based Visual Aid With Integrated Readings Assistant for the completely Blind

Author: Muiz Ahmed Khan , Pias Paul2020

In this article, we propose a novel visual aid system for completely blind individuals Hands free, wearable, low power, and compact design, mountable on a pair of eyeglasses, for the indoor and outdoor navigation with an integrated reading assistant. Complex algorithm processing with a low-end configuration. Real-time, camera-based, accurate distance measurement, which simplifies the design and lowers the cost by reducing the number of required sensors. Raspberry Pi 3 Model B+, Ultrasonic sensors, Image-to-text converter ,OpenCV, OCR engine. The current setup lacks advanced functions, such as wet-flood and staircases detection or the use of GPS and mobile communication module, the flexibility in the design leaves room for future improvements and enhancements. In addition, with the advanced machine learning algorithms and a more improved user interface, the system can further be developed and tested in a more complex outdoor environment.

2.1.10 “Smart Assistant”: A solution to facilities vision impaired individuals

Author: Pavithra Sankalpani, Iresha Wijesinghe2018

Smart phone to guide visually impaired people in indoor environments. FIDs, Bluetooth, and Wi-Fi triangulation, Navigation System. Dijkstra’s algorithm and binary tree algorithm are used to route navigation. User Localization is the concept which determines the location of a device using the strength of the Wi-Fi signal. In this research, several Wi-Fi routers were placed throughout the floor as Wi-Fi access points. Then the receiver, the smartphone can obtain the intensity of the received signal. Route Navigation - *The map should find best location*, For the routing algorithm, the researchers have used a graph-based approach -Dijkstra’s algorithm *Obstacle Detection* Smart phone camera to determine the object. The purpose of Voice Guidance is to convert the image of the optimal route to voice so the user can navigate according to the voice instructions. Make a web interface so that when the user goes to a new building, he/she can download the mapping information from a remote server to work with this system. Navigate Multiple floor Improve the system workable without internet Connectivity. The system can play a vital role in the indoor navigation of public.

2.2 INFERENCE FROM LITERATURE SURVEY:

- In some of the papers, time complexity is comparatively high.
- The honeywords generated are very random.
- The intruder is not stopped even after trying numerous times, making the site vulnerable to such attacks.

2.3 SCOPE OF THE PROJECT:

This project is used to give the visually challenged people's a moral support to adapt to the modern technologies that grew up faster and faster by some features that make them feel comfortable and make their work easy. It also provides as an enclosed package to the user with image processing functionality to guide them too

2.4 LIMITATIONS CONSIDERED IN THE PROJECT:

- It fails to operate efficiently in a noisy environment. You need to have a silent and calm environment in order to make Elaine work.
- A user may have privacy concern as personal assistant require huge amount of data and are always listening to take command, there is a concern that our data may be over heard by others. So, cloud based solution must be there which will protect the user information.

SYSTEM SPECIFICATION

CHAPTER 3

SYSTEM SPECIFICATION

3.1 SOFTWARE SPECIFICATION:

- Editors : Jupiter notebook,
visual studio code,
Open CV
- Coding Language : Python

3.2 HARDWARE SPECIFICATION:

- Operating system : Windows 10,Ubuntu
- Processor : Intel core i5
- Hard disk : 1TB
- RAM : 8GB (maximum)

SYSTEM ANALYSIS

CHAPTER 4 SYSTEM ANALYSIS

EXISTING SYSTEM:

The existing system is not covering all the features mentioned in the structure to the extent and some of the features that deals with only the android type not regarding to the system. With considering the features that are presented as atomic form not as integrated structure it is necessary as a package which can be useful for the society .The system of project we implemented is not presented as full package previously ,it is just a glimpse of it. The system does not covers the image processing implementation that we are fitting in our system in order to provide a whole sum of what we want to project it to the users. It was established for open set of users not for the particular field of people.

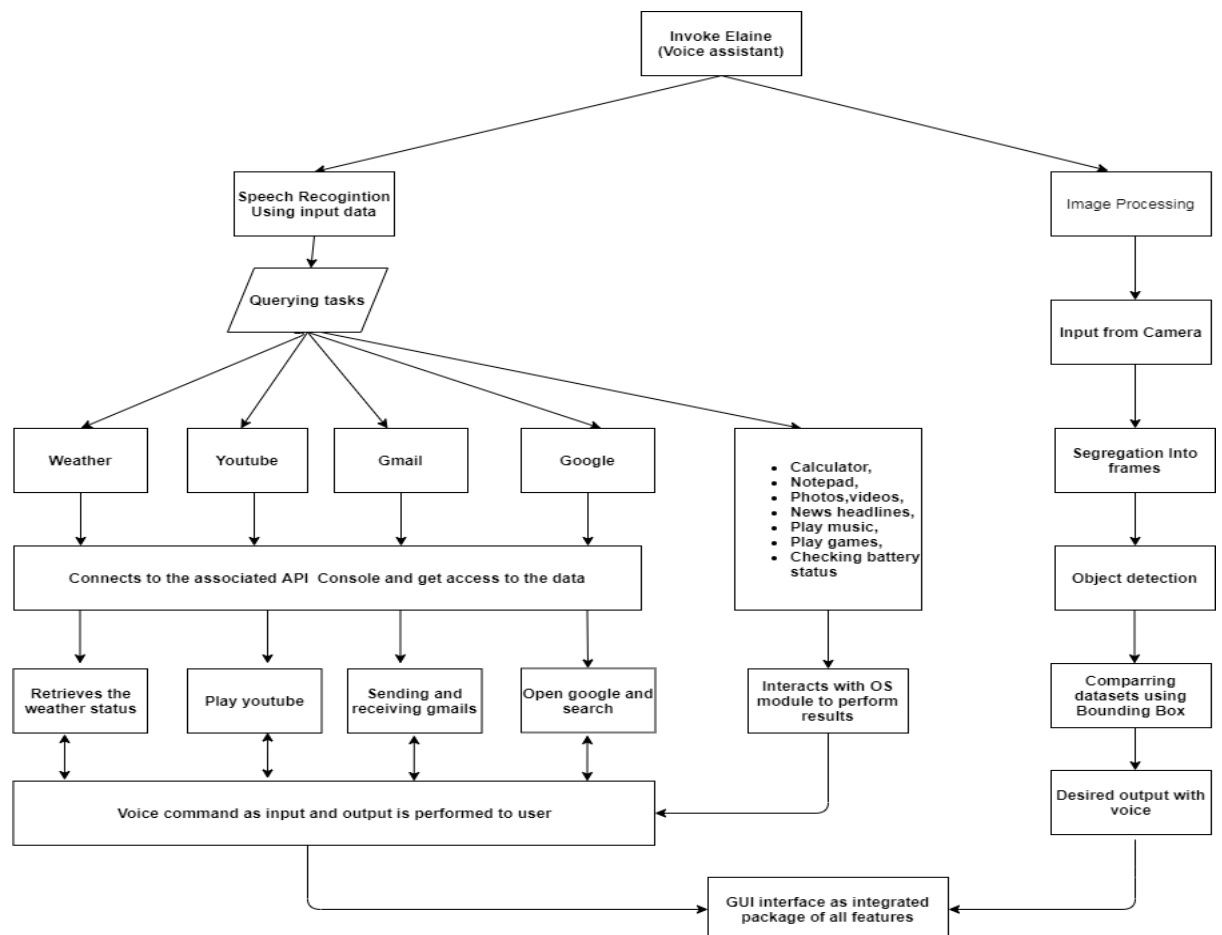
PROPOSED SYSTEM:

The proposed system tries to fill this void manner in the existing one. We build this project as a integrate one covering many features and functionalities like Google, Wikipedia, You tube, Gmail, photos and videos, Taking notes and interactable chatbot and with image processing idea to guide the blind people so that they find their way as ease and user friendly in spite of upcoming technologies that they find it difficult to adapt them .We also created this interface from the system that they lack of ,so it forms another level of framework which is completely different from them.

METHODOLOGIES:

The working models used here is python language for voice recognition and Yolo V3 algorithm dealt with image processing. We have used Google API console for every module like maps, YouTube, search API's associated with it. There are many modules like pyttsx3 to implement the speech recognition. First the voice is Percepts by the user and the system identifies using the GTTS feature and converts into word and the continued performance is done according to needs like searching ,Gmail etc. API provided on the system will let you do all the functions like interacting with other level of system like YouTube ,Gmail in order to provide a function related to that. Image processing deals with better algorithm to identify objects and guide through the user with defined level of dataset that we prescribed through the project. The system provides the reliability, intelligence and assistance to the user as entire package comprising of all the feature to deal with the blind person.

OVERALL SYSTEM DESIGN:



SPEECH RECOGNITION:

Speech recognition, or speech-to-text, is the ability for a machine or program to identify words spoken aloud and convert them into readable text. Here we use Pyttsx3 – Text to speech conversion library and GTTS - Google Text-to-Speech)library and CLI tool to interface with Google Translates text-to-speech API. This module helps to convert String text to Spoken text and can be saved as .mp3. The input voice is extracted and identify the content of command and proceed accordingly into any one of the modules.

IMAGE PROCESSING:

Image processing is a method to perform some operations on an image, in order to get an enhanced image or to extract some useful information from it. Here we use YOLO V3 algorithm to extract and identify the information from the given image. We deal with large amount of dataset to work with. First the image is divided into large set of frames. Image is fed into the system using camera as an input. After the extraction process, the object is identified and compared with existing dataset using bounding box method and then automated voice is prepared as a result.

INTERFACING WITH API:

In developing the project, we need to interact with several API consoles in order to proceed with databases that we are implementing here such as Google enterprise and other stuffs. To connect with several API'S we need to create a API key that can be used in segment of the program which is access to the entire framework of database. The key can be created by enabling the needed API's for particular process. This way we can interface the API's easily.

SEGREGATION OF FRAMES:

Here we consider an image to partially formed into large no of frames. Each frame is formed by pixel located in that, since it has unique color and property so that differentiating will able to give you the entire context. This will be helpful in identifying the object and proceed for the further operations.

OBJECT DETECTION USING YOLO:

Here we using Yolo algorithm to object detection since it is the efficient method. We obtain the image using camera and try to identify it. In this approach, a single neural network divides the image into regions and predicts bounding boxes and probabilities for each region. The neural network predicts bounding boxes and class probabilities directly from full images in one evaluation. The base YOLO model processes images in real-time at 45 frames per second.

API USED IN OUR MODULES:

GOOGLE:

Here we used **Selenium Web driver** which is a open source API's to automate web application testing in order to deal with web data store. Custom **search API** is to use to generate results based on the pattern and **BeautifulSoup** which is a web scraping html data from URL.

YOUTUBE:

Here we used **YouTube data API** to incorporate you tube functionality into our application and **search()** function is implemented to retrieve top 10 results with the given pattern .In order to end the video within a given duration we used **paify()** method.

GMAIL:

This part of module is implemented using **SMTP** and **IMAP** protocol for sending and receiving Gmail's. Two parts of process are Compose mail using receiver mail id and subject from the user as voice input, and Check inbox by retrieving the inbox and unseen mails and deliver using voice

SYSTEM IMPLEMENTATION **AND RESULT**

CHAPTER 5 **SYSTEM IMPLEMENTATION AND RESULT**

5.1. OVERVIEW

For the implementation of the system we used google APIs for accessing google applications such as you tube ,Gmail, google search etc. For image detection the algorithm used is YOLOv3, and python's OpenCV. The GUI used is

5.2. MODULES

The modules used in the system are

1. Weather forecast
2. YouTube
3. Gmail
4. Google Search
5. Play music
6. Calculator
7. News Headlines
8. Taking photos and videos
9. Storing notes
10. GUI

5.2.1 Weather forecast

Weather forecast module will give the weather stats about the location. Command for weather module is when Elaine is in listening state user have to say weather in your command for example "What is the weather now?" , then Elaine will ask user for the location ,then you have to say the location you want to have weather stats .For example "Erode". Then Elaine will give detailed stats about the temperature, weather description,

humidity and wind speed of the location. An example of the implementation is attached below.

```
Elaine is Listening...
You said: what is the weather now

Elaine is Listening...
You said: erode

-----
Weather Stats for - ERODE  || 22 Apr 2021 | 01:28:22 PM
-----
Current temperature is: 34.00 deg C
Current weather desc  : scattered clouds
Current Humidity      : 46 %
Current wind speed    : 2.57 kmph
```

Fig 5.1 The weather forecast

5.2.2 You tube

YouTube module will search and play the desired YouTube video by the user. Command for the YouTube module is when Elaine is in the listening state user have to say for example “Open YouTube and search for Indian economy” and then it will play the video for the user . An example of implementation is attached below.

```
Elaine is Listening...
You said: open youtube and search for indian economy

f-1gn89rqPM
https://www.youtube.com/watch?v=f-1gn89rqPM
```

Fig 5.2 You Tube search

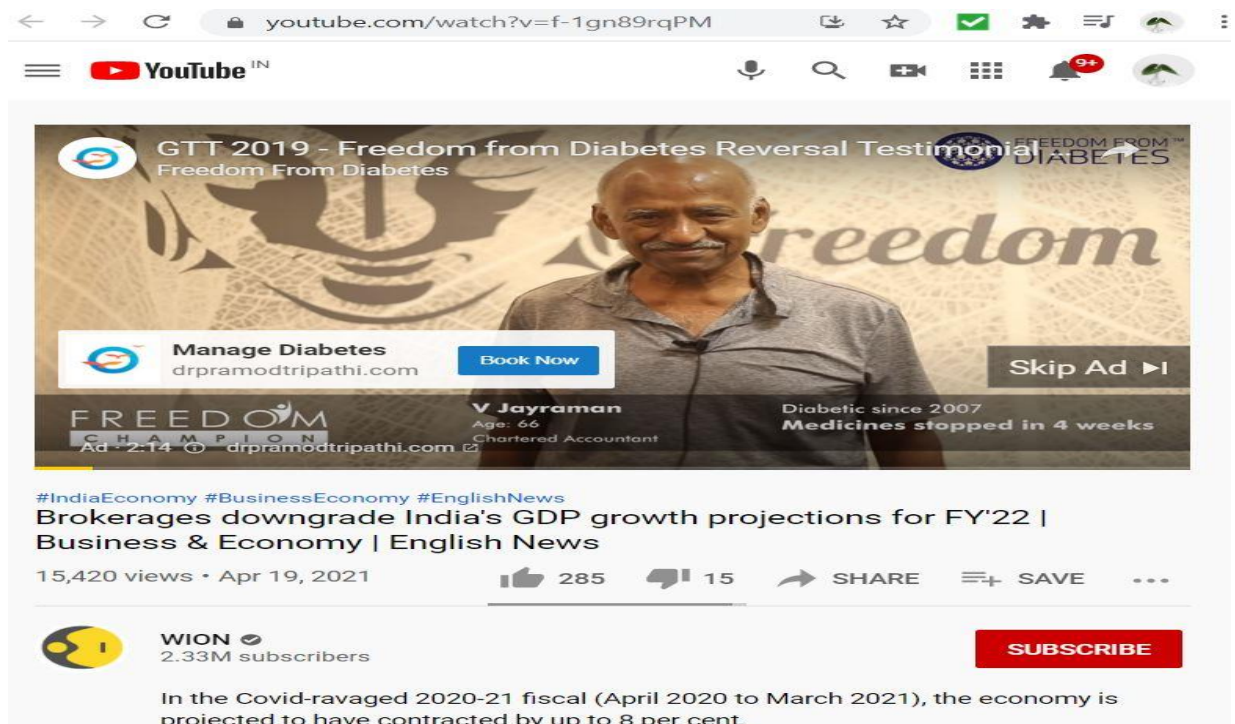


Fig 5.3 Result page

5.2.3 Gmail

Compose mail

In Gmail module users can send email . Command for compose mail is , when Elaine is in listening state users have to say “compose mail” . Elaine will say users login . Then Elaine will ask whether to compose mail or check inbox?. User have to say “compose mail”. Then Elaine will ask for receiver id and user can say the message he/she wants to send.

An example of the implementation is attached below.


```

Elaine is Listening...
You said: compose mail

You are logged In from : Rooba
Do you want to compose mail or check inbox?
Elaine is Listening...
You said: compose mail

Clearing noise...
Please tell the Name of receiver mail ID
Elaine is Listening...
You said: hello i will be there in 5 minutes

listening the message....
ok!
printing the message....
Your message:please come here as soon as possible
Receiver mail id:abinaya200116@gmail.com
Congrates! Your mail has been send.
Elaine is Listening...
You said: exit

```

Fig 5.4 Sending Email through voice

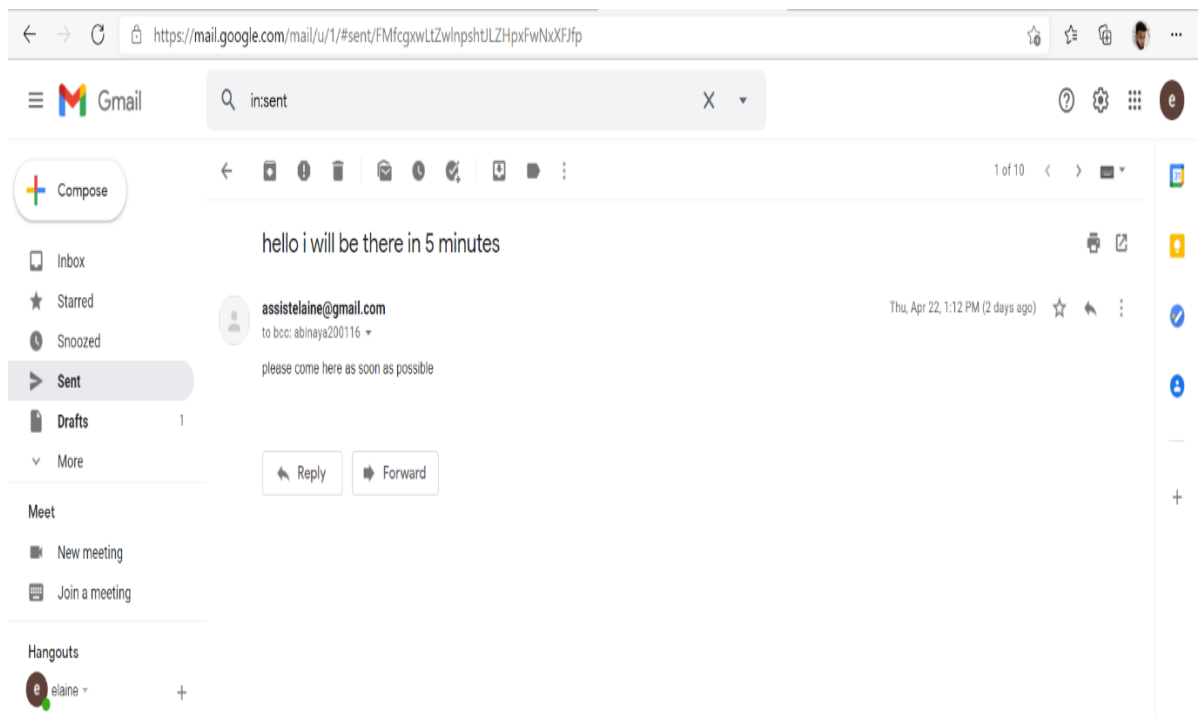


Fig 5.5 Output of Gmail Page

Check inbox

User can check their inbox. When Elaine is in listening state users have to say “to mail”. Elaine will say the user login. It will ask whether to compose mail or check inbox? User should say “check inbox”. It will say user the number of mails in the inbox and number of unread mails. It will read the unread mail to the user. An example of the implementation is attached below.

```
Elaine is Listening...
You said: to mail

You are logged In from : Rooba
Do you want to compose mail or check inbox?
Elaine is Listening...
You said: check inbox

Number of mails in your inbox :[b'20']
[b'20']
Your Unseen mail :1
subject: Google kia kar me marjun
to: "assistelaine@gmail.com" <assistelaine@gmail.com>
from: Rooba Devi Devaraj <roobadevi27@gmail.com>
date: Thu, 22 Apr 2021 13:08:26 +0530
C:\Users\Rooba\mini-project\lib\site-packages\pyglet\media\codecs\wmf.py:838
-2147417850] Cannot change thread mode after it is set
  warnings.warn(str(err))
Please be ready .. I will be there within 5 mins..
Ok
```

Fig 5.6 Details of Unseen mail

The number of unseen mail and subject and body of the message all printed through voice

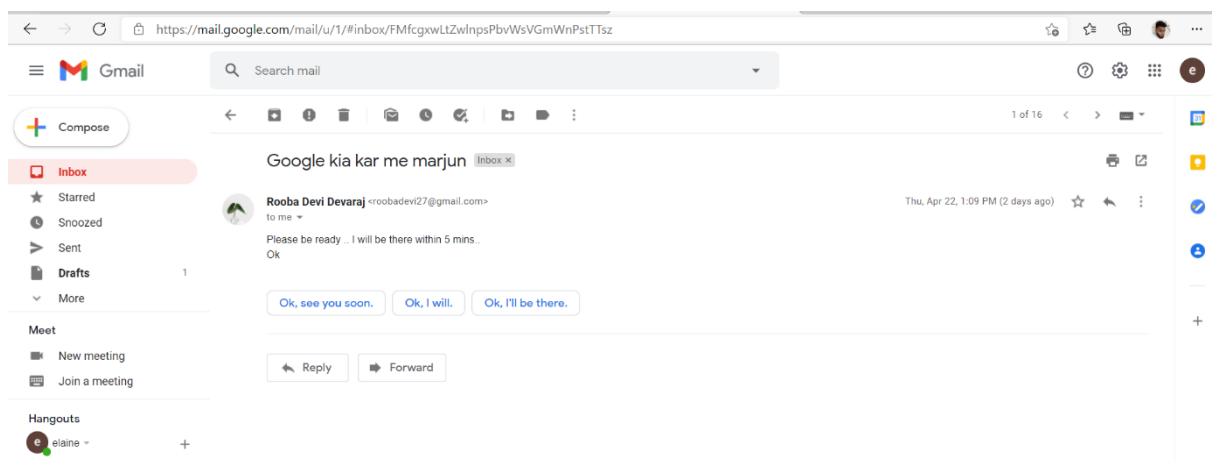


Fig 5.7 Output page of Unseen email

5.2.4 Google search

Google search module will search and read the content from google to the user. Command for this is when Elaine is in listening state user have to say for example “open google and search for indian economy” . It will search google and read the content for the user. An example of implementation is attached below.

```
Elaine is listening...
You said: open google and search for indian economy

for indian economy

[WDM] - Current google-chrome version is 90.0.4430
[WDM] - Get LATEST driver version for 90.0.4430
[WDM] - Driver [C:\Users\Rooba\wdm\drivers\chromedriver\win32\90.0.4430.24\chromedriver.exe] found i
n cache

DevTools listening on ws://127.0.0.1:62957/devtools/browser/c8b41c58-406b-4fe6-b20e-f59705b1706f
https://www.thehindu.com/business/Economy/indias-economy-to-contract-by-103-says-international-moneta
ry-fund/article32844627.ece

International Monetary Fund (IMF) logo is seen outside the headquarters building in Washington. File
| Photo Credit:
Reuters

With the country and world reeling under the impact of the coronavirus pandemic, the Indian economy i
s expected to grow at -10.3 % ( i.e., a contraction) in 2020 as per the International Monetary Fund (
IMF). Global growth is projected to be -4.4% ( i.e., a contraction in output of 4.4%) for this year ,
the IMF said with the release of its World Economic Outlook October 2020 report titled, “A Long and
```

Fig 5.8 Google search

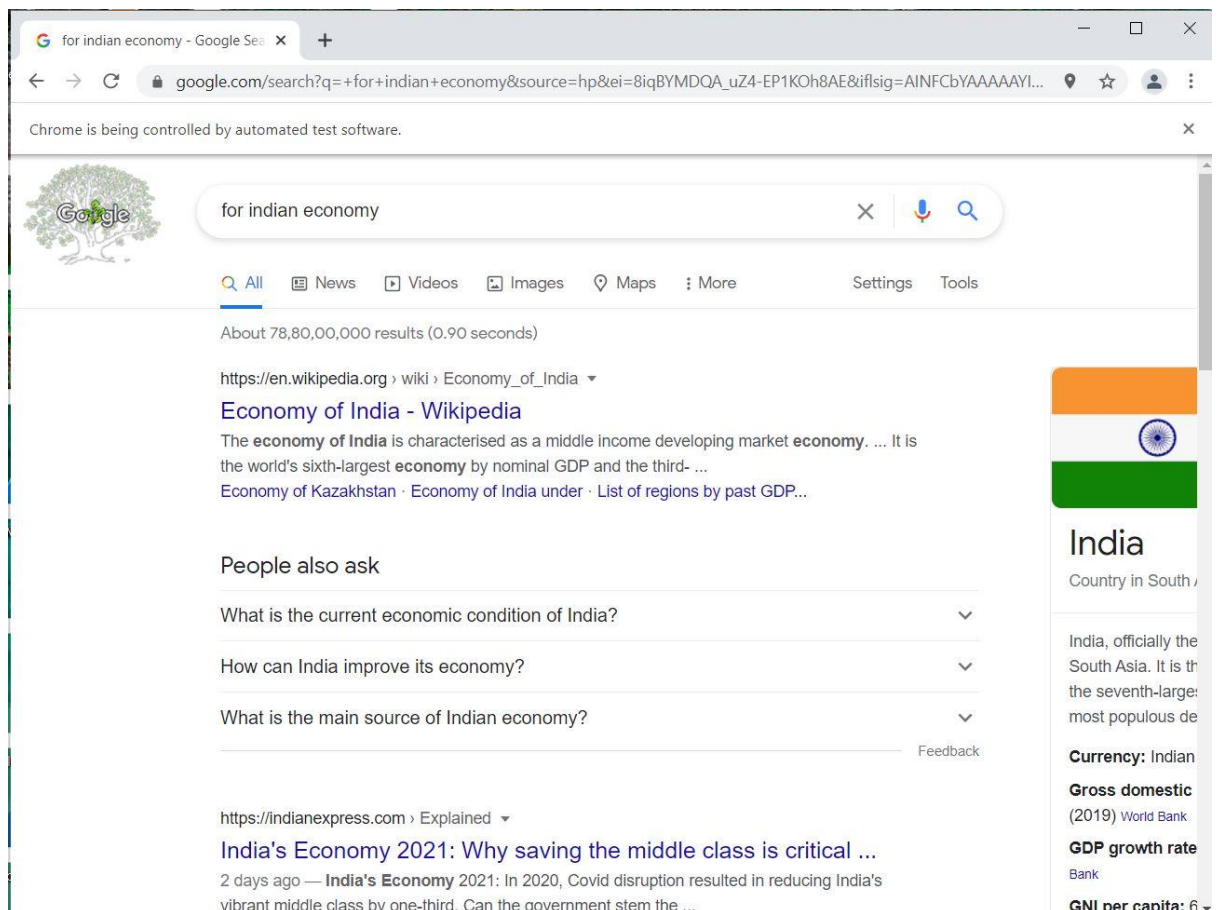


Fig 5.9 Output of the google page

5.2.5 Play music

Play music module will search and play music for the user. Command for this is when Elaine is in listening state user have to say “ Play music”. It will play some music for the user. An example of the implementation is attached below.

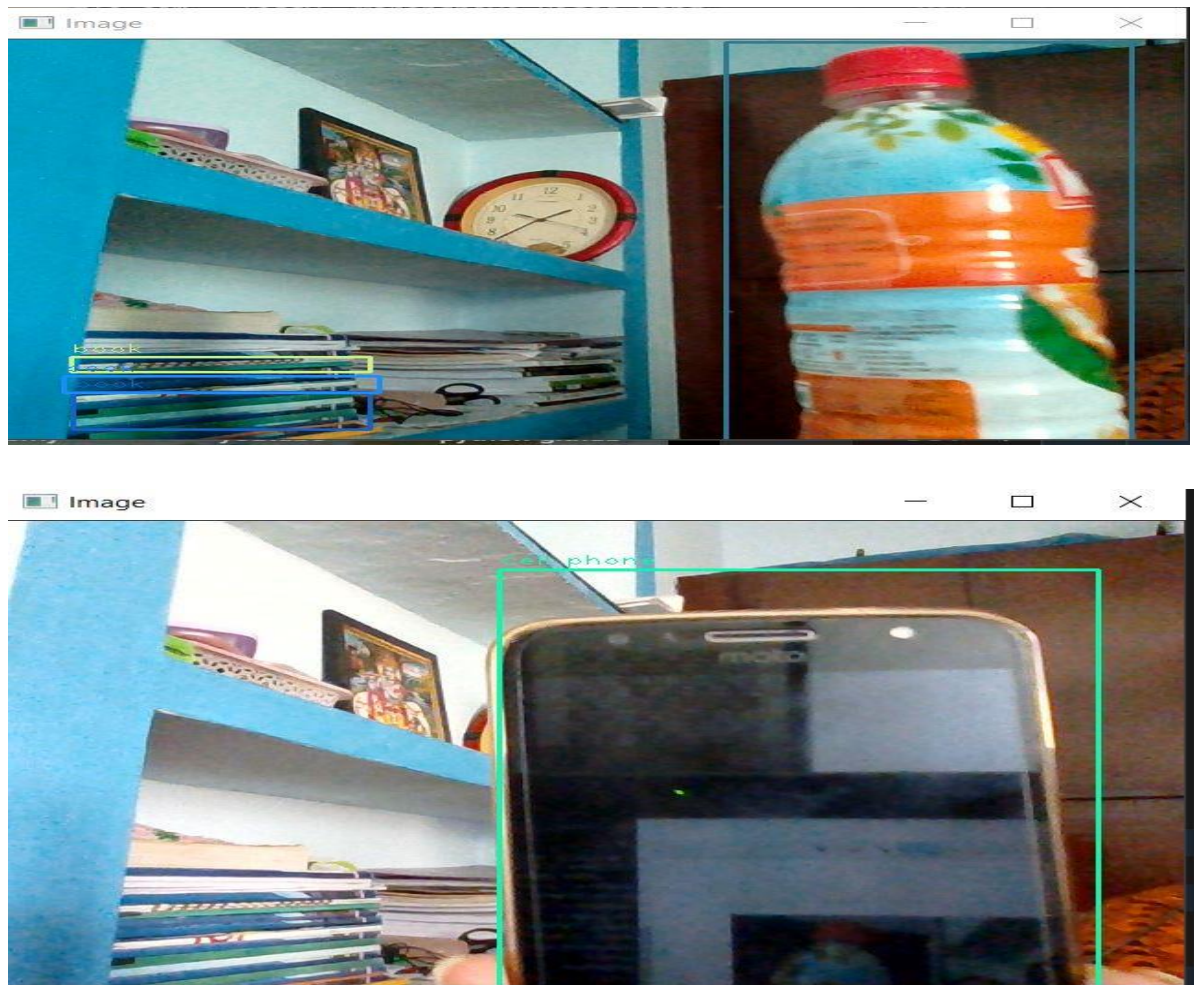
```
Elaine is Listening...
You said: play music

['classic_love scenery ost.mpeg', 'desktop.ini', 'Rock- Love Scenery 🎵 🎵 🎵 🎵 🎵 OST.mp3', 'Rock-
Love Scenery 🎵 🎵 🎵 🎵 🎵 OST.mp4', 'yt1s.com - Enjoy Enjaami Dhee ft Arivu Official Video.mp3', 'y
t1s.com - Teddy Marandhaye Video Song Arya Sayyeshaa D Imman Shakti Soundar Rajan.mp3']
Playing::::: Rock- Love Scenery 🎵 🎵 🎵 🎵 🎵 OST.mp3
```

Fig 5.10 Play Music through voice command

5.2.6 Image processing

Image processing module will detect objects Infront of the user and say it to them. Command for this is when Elaine is in listening state user have to say “start image”. Here it detects bottle, cell phone, books, and clock. An example of implementation is attached below.



Detect images and tell objects through voice. Here it detects bottle, cell phone, books, and clock

Fig 5.11 Detecting objects page

5.2.7 Calculator

Calculator module will help user to calculate. Command for this is when Elaine is in listening state user have to say “calculate” . Then it ask value to calculate. For example user can ask “4 power 12” it will say the result as “16777216”. An example of the implementation is given below.

```
Elaine is Listening...
You said: calculate

Say what you want to calculate, example: 3 plus 3
if u want to calculate power of something or cube or square say like 4 power 2
Elaine is Listening...
You said: 4 power 12

answer : 16777216
Elaine is Listening...
```

Fig 5.12 Calculator

5.2.8 News headlines

News headlines module will read the news headlines of the current day to the users.

An example of implementation is attached below.

```
===== TIMES OF INDIA =====

1. "Will Hang That Man": High Court On Anyone Obstructing Oxygen Supply - NDTV
If anyone obstructs oxygen supply, we will not spare them, the Delhi High Court said today as it heard a hospital's petition over shortage of oxygen for seriously-ill Covid patients.

2. Astronauts Aboard SpaceX Capsule Prepare For Docking At ISS - NDTV
A recycled SpaceX Crew Dragon capsule with four astronauts aboard was speeding on its way to the International Space Station, where the spacecraft is set to dock early Saturday.

3. Mi 11X Pro Pre-Orders Begin in India Today: Price, Launch Offers, Specifications - Gadgets 360
Mi 11X Pro is up for pre-order in India via Mi.com. Amazon India is only delivering essentials at the moment in some areas and, therefore, the pre-order button is not live for those pin codes. Launch offers include up to Rs. 4,000 instant discount on HDFC credit cards.

4. Udit Narayan says Shravan Rathod called him from Kumbh: 'I thought to myself that during the pandemic, why did he go' - Hindustan Times
Singer Udit Narayan has talked about how late music composer Shravan Rathod had called him from the Kumbh Mela. Shravan died on Thursday night of Covid-19 complications.
```

Fig 5.13 News Headlines

5.2.9 Taking photos and videos and store in folder

This module allows users to take photos and videos through their camera and store it locally. Command for this is user have to say “camera” to take photo and “video” to take video. An example of implementation is attached below.

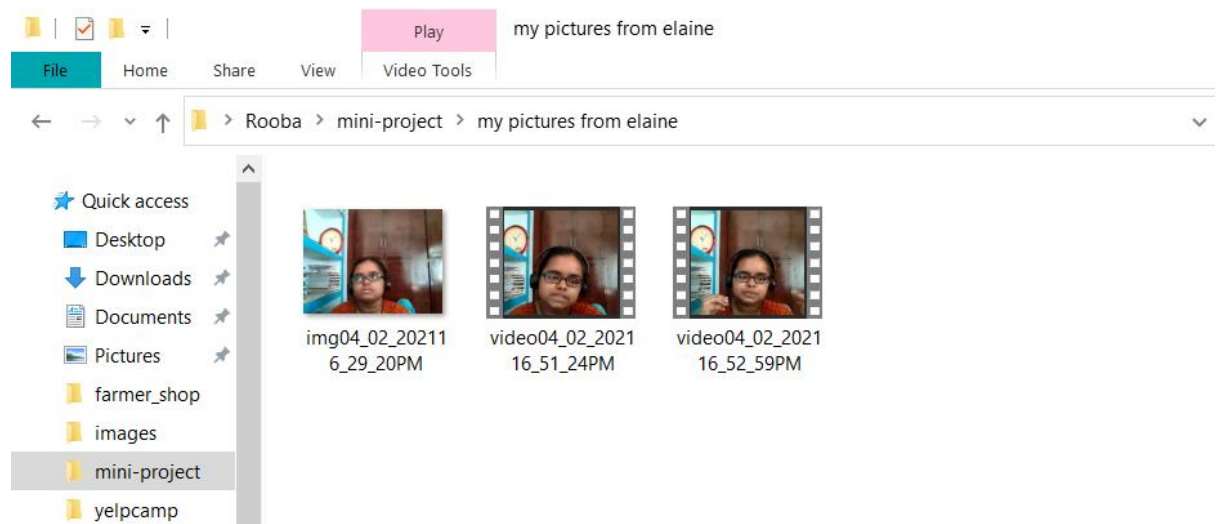


Fig. 5.14 Taking photos and videos and store in folder

5.2.10 Taking and storing notes in folder

In this module user can store their notes and can ask Elaine to read their latest notes. Command for write a note is “take note” and command for read a note is “show note”. An example of implementation is attached below.

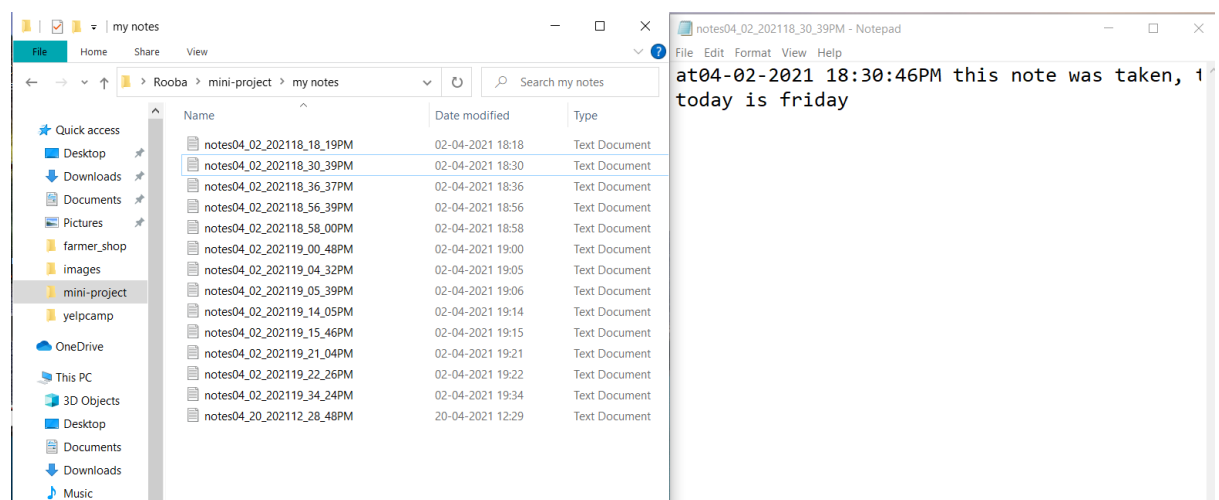


Fig 5.15 Taking and storing notes in folder

5.2.11 GUI

We developed the GUI as an integrated package of all features implemented in our project using python utility PyQt5. This comprises of every feature and executed using the voice command and get interacted with the system.

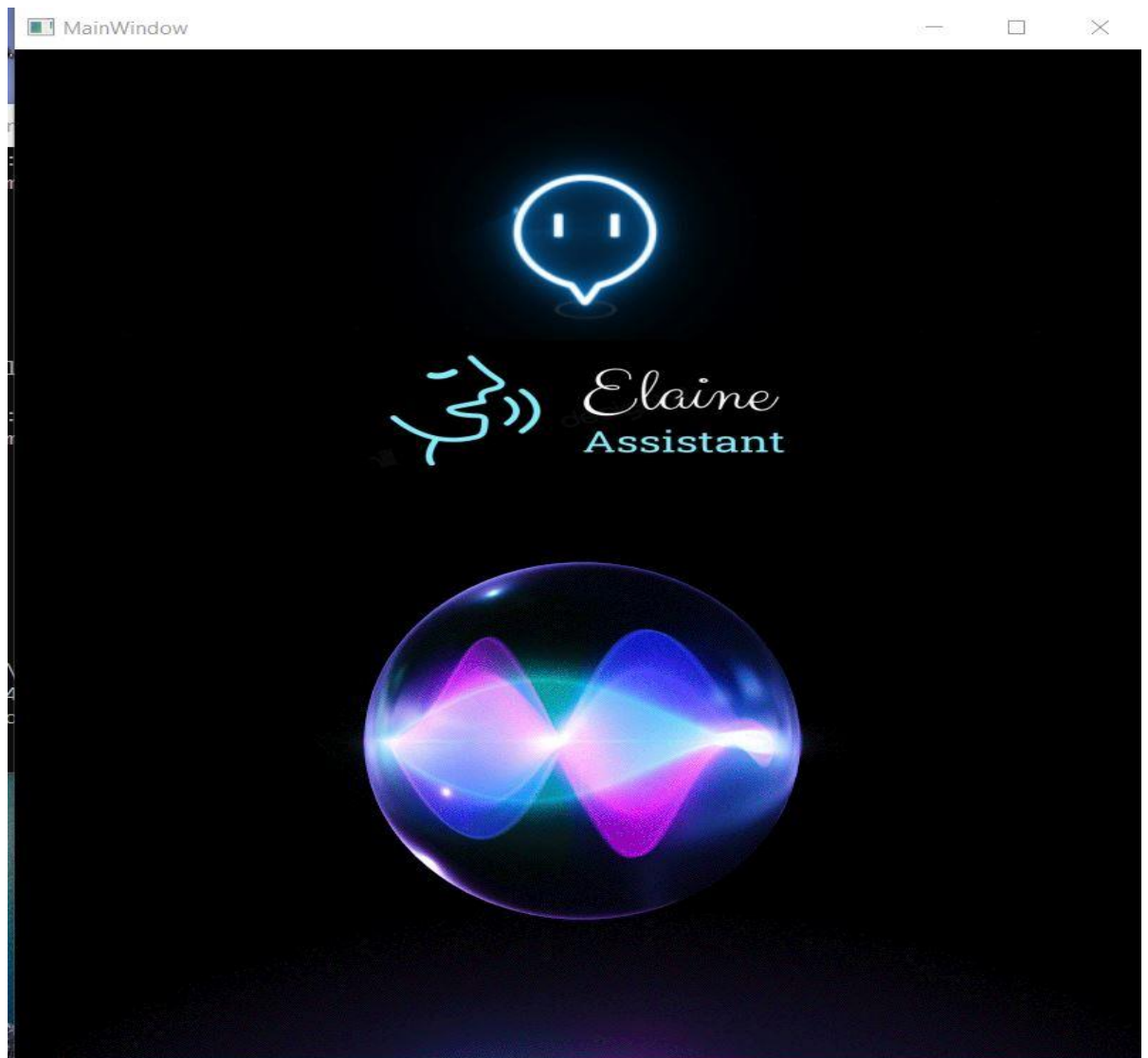


Fig 5.16 GUI integration part

5.2.12 Checking battery

System detect the battery percentage and tell through voice

```
hello from the pygame community. https://www.pygame.org
Elaine is Listening...
You said: check battery status

Your system is having 100 percent battery
Elaine is Listening...
Elaine is Listening...
```

Fig 5.17 Battery result

5.2.13 Simple commands

Elaine will respond through voice for some basic conversations like who are you?, who made you? , who I am?, games, jokes, time, date etc

```
(mini-project) C:\Users\Rooba\mini-project>python main.py
pygame 2.0.1 (SDL 2.0.14, Python 3.8.3)
Hello from the pygame community. https://www.pygame.org/contribute.html
Initializing Elaine
Checking the internet connection
All drivers are up and running
All systems have been activated
Now I am online
Elaine is ready!
Good Evening!
I am your Assistant Elaine 1 point o How can I help you?
Elaine is Listening...
You said: how are you a line

I am fine, Thank you
How are you?
Elaine is Listening...
You said: i am good

It's good to know that your fine
Elaine is Listening...
You said: thank you

its fine, check out me for more exciting things
Elaine is Listening...
You said: who made you a line

I have been created by the team VisionWalk.
Elaine is Listening...
You said: who i am

If you talk then definitely your human.
Elaine is Listening...
```

Fig 5.18 Output of Simple conversation part 1

```
Elaine is Listening...
You said: who i am

If you talk then definitely your human.
Elaine is Listening...
You said: why you came to world

Thanks to VISIONWALK. further It's a secret
Elaine is Listening...
You said: who are you a line

I am Elaine, your virtual assistant created by caring team rooba, dineshkanna, abinaya and manikandan to assist visually impaired people in everyday life.
Elaine is Listening...
You said: what is love

It is 7th sense that destroy all other senses
Elaine is Listening...
You said: what is the reason for you

I was created as a Minor project by the team VisionWalk
Elaine is Listening...
You said: what is newton's third law

noun | action and reaction are equal and opposite
Elaine is Listening...
You said: what can you do

i can able to fetch information you want by surfing google and update about news headlines, weather status, and i can play music and youtube to relax you and i help you to send and receive mails and i help you with any mathematical calculation
you please don't worry about anything!!
im here to help you with anything
Elaine is Listening...
You said: bye

Bye. Check Out me for more exciting things.. Im going offline

(mini-project) C:\Users\Rooba\mini-project>
```

Fig 5.19 Output of Simple conversation part 2

```

Elaine is Listening...
You said: tell me some jokes

What's the object-oriented way to become wealthy? Inheritance.
Elaine is Listening...
Elaine is Listening...
You said: what day is today

today isSaturday
Elaine is Listening...
You said: what is the time now

the time is 04-24-2021 21:19PM
Elaine is Listening...

```

Fig 5.20 Output of Simple conversation part 3

```

Elaine is Listening...
You said: search in wikipedia mass

Mass is both a property of a physical body and a measure of its resistance to acceleration (rate of change of velocity with respect to time) when a net force is applied.[1] An object's mass also determines the strength of its gravitational attraction to other bodies.

The SI base unit of mass is the kilogram (kg). In physics, mass is not the same as weight, even though mass is often determined by measuring the object's weight using a spring scale, rather than balance scale comparing it directly with known masses. An object on the Moon would weigh less than it does on Earth because of the lower gravity, but it would still have the same mass. This is because weight is a force, while mass is the property that (along with gravity) determines the strength of this force.

There are several distinct phenomena that can be used to measure mass. Although some theorists have speculated that some of these phenomena could be independent of each other,[2] current experiments have found no difference in results regardless of how it is measured:

```

Fig 5.21 Wikipedia

CONCLUSION AND FUTURE WORK

CHAPTER 6

CONCLUSION AND FUTURE WORK

The main aim of this project is to provide voice-based control system to access the google search engine, emails, and navigation map for blind people using speech recognition module and it is intended to provide object detection as real time assistance to identify the objects(things) around them. It has the capability of access to World Wide Web by browsing the internet, checking, sending and receiving email, listening to the content of search only by giving a voice command to the system. These can be implemented and adopted by the strategies followed in the research papers which is more efficient and suitable for our system

Future work

At present the application supports only English language. We plan to expand that and make it available in most of the daily used languages thus people from all parts of the world can access the application without any issue. And also embed our application in IOT device along with GPS to guide user navigation by connecting with maps.

But for all intents and purposes, the future of these technology is a bright one. With advances in it and technologies related to it(search process, for example) voice assistants can carry out even more complex tasks like booking tickets, etc.

REFERENCES

CHAPTER 7

- [1] E. Marvin, "Digital Assistant for the Visually Impaired," 2020 International Conference on Artificial Intelligence in Information and Communication (ICAIIIC), pp. 723-728.
- [2] Iyer, V., Shah, K., Sheth, S., &Devadkar, K. ,”Virtual assistant for the visually impaired” 2020 5th International Conference on Communication and Electronics Systems (ICCES)
- [3] Sultan, M. R., & Hoque, ,”ABYS (Always By Your Side): A Virtual Assistant for Visually Impaired Persons “ M. M. 2019 22nd International Conference on Computer and Information Technology (ICCIT)
- [4] Rashika Joshi, Meenakshi Tripathi., Amit Kumar and Manoj Singh Gaur ,”Object Recognition and Classification System for Visually Impaired” , 2020 International Conference on Communication and Signal Processing (ICCSP)
- [5] S. Noel, "Human computer interaction(HCI) based Smart Voice Email (Vmail) Application - Assistant for Visually Impaired Users (VIU)," 2020 Third International Conference on Smart Systems and Inventive Technology (ICSSIT),
- [6] P. A. Tiwari, P. Zodawan, H. P. Nimkar, T. Rotke, P. G. Wanjari and U. Samarth, "A Review on Voice based E-Mail System for Blind," 2020 International Conference on Inventive Computation Technologies (ICICT)
- [7] S. M. Felix, S. Kumar and A. Veeramuthu, "A Smart Personal AI Assistant for Visually Impaired People," 2018 2nd International Conference on Trends in Electronics and Informatics (ICOEI), 2018

- [8] . P. Khairnar, R. B. Karad, A. Kapse, G. Kale and P. Jadhav, "PARTHA: A Visually Impaired Assistance System," 2020 3rd International Conference on Communication System, Computing and IT Applications (CSCITA), 2020
- [9] M. A. Khan, P. Paul, M. Rashid, M. Hossain and M. A. R. Ahad, "An AI-Based Visual Aid With Integrated Reading Assistant for the Completely Blind," in IEEE Transactions on Human-Machine Systems
- [10] P. Sankalpani, I. Wijesinghe, I. Jeewani, R. Anooj, M. D. J. T. H. Mahadikaara and J. A. D. C. Anuradha Jayakody, "'Smart Assistant': A Solution to Facilitate Vision Impaired Individuals," 2018 National Information Technology Conference (NITC), 2018