Visvesvaraya Technological University

Jnana Sangama, Belagavi - 590018

A Project Work Phase-2 (18CSP83)

Report on

"Language Pronunciation Translator For Better Understanding"

Project Report submitted in partial fulfilment of the requirement for the award of the degree of

BACHELOR OF ENGINEERING

IN

COMPUTER SCIENCE AND ENGINEERING

Submitted by

Hamsaveena S	**1KG18CS040**
Karthik Gowda D	**1KG18CS045**
Rama Reddy B T	**1KG18CS089**
Shankar K S	**1KG18CS099**

Under the guidance of

Jayashubha J
Assistant Professor
Department of Computer Science & Engineering
KSSEM, Bengaluru-560109

DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING

K. S. School of Engineering and Management

#15, Mallasandra, off. Kanakapura Road, Bengaluru – 560109

2021 - 2022

K. S. School of Engineering and Management

#15, Mallasandra, off. Kanakapura Road, Bengaluru - 560109

Department of Computer Science & Engineering

CERTIFICATE

Certified that the Project Work Phase-II (18CSP83) entitled **"Language Pronunciation Translator For Better Understanding"** is a bonafide work carried out by:

Hamsaveena S	**1KG18CS040**
Karthik Gowda D	**1KG18CS045**
Rama Reddy B T	**1KG18CS089**
Shankar K S	**1KG18CS099**

in partial fulfilment for VIII semester B.E., Project Work in the branch of Computer Science and Engineering prescribed by **Visvesvaraya Technological University, Belagavi** during the period of April 2022 to June 2022. It is certified that all the corrections and suggestions indicated for internal assessment have been incorporated. The Project Work Phase-2 Report has been approved as it satisfies the academic requirements in report of project work prescribed for the Bachelor of Engineering degree.

.....	.....	.....
Signature of the Guide	**Signature of the HOD**	**Signature of the Principal**
[Jayashubha J]	[Dr. K Venkata Rao]	[Dr. K Rama Narasimha]

DECLARATION

We, the undersigned students of 8th semester, Computer Science & Engineering, KSSEM, declare that our Project Work Phase-II entitled "**Language Pronunciation Translator For Better Understanding**", is a bonafide work of ours. Our project is neither a copy nor by means a modification of any other engineering project.

We also declare that this project was not entitled for submission to any other university in the past and shall remain the only submission made and will not be submitted by us to any other university in the future.

Place: Bangalore
Date:11-07-2022

Name and USN	**Signature**
Hamsaveena S (1KG18CS040)	**…….**
Karthik Gowda D (1KG18CS045)	**…….**
Rama Reddy B T (1KG18CS089)	**…….**
Shankar K S (1KG18CS099)	**…….**

ACKNOWLEDGEMENT

The satisfaction and euphoria that accompany the successful completion of any task will be incomplete without the mention of the individuals, we are greatly indebted to, who through guidance and providing facilities have served as a beacon of light and crowned our efforts with success.

We would like to express our gratitude to our **MANAGEMENT,** K.S. School of Engineering and Management, Bengaluru, for providing a very good infrastructure and all the kindness forwarded to us in carrying out this project work in college.

We would like to express our gratitude to **Dr. K.V.A Balaji, CEO**, K.S. School of Engineering and Management, Bengaluru, for his valuable guidance.

We would like to express our gratitude to **Dr. K. Rama Narasimha, Principal**, K.S. School of Engineering and Management, Bengaluru, for his valuable guidance.

We like to extend our gratitude to **Dr. K Venkata Rao**, **Professor and Head**, Department of Computer Science & Engineering, for providing a very good facilities and all the support forwarded to us in carrying out this Project Work Phase-II successfully.

We also like to thank our Project Coordinators, **Mrs. R S Geethanjali, Asst. Professor, Mrs. Gargi N, Asst. Professor, Mrs. Swathi D, Asst. Professor, and Mrs. Shalini K V, Asst. Professor, Department of Computer Science & Engineering** for their help and support provided to carry out the Project Work Phase-II successfully.

Also, we are thankful to **Jayashubha J, Assistant Professor**, for being our Project Guide, under whose able guidance this project work has been carried out Project Work Phase-II successfully.

We are also thankful to the teaching and non-teaching staff of Computer Science & Engineering, KSSEM for helping us in completing the Project Work Phase-II work.

Hamsaveena S

Karthik Gowda D

Rama Reddy B T

Shankar K S

ABSTRACT

With technology getting better and advanced every single day through intensive inventions and discoveries carried out by experts all around the world, our lives are getting modified and bent more towards being dependent on technology and making the best out of it. With such a level of enhancements, it will be really illogical, if we still struggle to interact with our systems through various contemporary I/O devices (input/output). As, we need to interact with our smart systems almost every hour of the day and it is really necessary to use hands-free techniques other than the use of peripheral devices to communicate with these computer systems; techniques that are faster, easier and affordable.

To cope up with these requirements, voice assistants are being made and developed during past few years. The main goal of a voice assistant is to reduce the use of I/O devices, to make interaction with our systems easy and possible for all. Also, it helps to save the space consumed by I/O devices and the cost we need to bear for buying and maintaining them properly. In this research paper, we have proposed a new python-based voice assistant that will be able perform almost every task that we use our devices for, just by listening to voice commands. It will be affordable for any class of the society and at some points better than the prevalent voice assistants available in the market. We have also provided a framework that shows the flow of control of the system. Finally, an extensive comparison with the existing systems proved the worthiness of our project.

TABLE OF CONTENTS

LIST OF FIGURES

Chapter 1

INTRODUCTION

1.1 Overview

Machine replication of human functions like reading is an ancient dream. However, over the last five decades, machine reading has grown from a dream to reality. Visually impaired people report numerous difficulties with accessing printed text using existing technology, including problems with alignment, focus, accuracy, mobility and efficiency. We present a smart device that assists the visually impaired and travellers which effectively and efficiently reads paper-printed text. The proposed project uses the methodology of a camera based assistive device that can be used by people to read Text document.

The framework is on implementing image capturing technique in an embedded system based on Raspberry Pi board. The design is motivated by preliminary studies with visually impaired people, and it is small-scale and mobile, which enables a more manageable operation with little setup. In this project we have proposed a text read out system for the travellers and visually challenged. The proposed fully integrated system has a camera as an input device to feed the printed text document for digitization. Speech is probably the most efficient medium for communication between humans.

Text extraction from image is one of the complicated areas in digital image processing. It is a complex process to detect and recognize the text from image. It's possible of computer software can provide extracted text from image using most complicated algorithm. So it can't be use anywhere in this existing environment. Here different types of language translators are available such as voice based translator, keyboard based translator etc. But those translators are not easy to use.

The purpose of this work is to demonstrate that a tight dynamical connection may be made between text and interactive visualization imagery. The Android device camera can prove this type of extraction and also the algorithm will easily implemented using java language. Millions of mobile users in this world and they always have mobile in their hand, so simply they can capture the image to extract the text.Optical Character Recognition (OCR)

is a technology for recognizing text in images, such as scanned documents and photos. We must have taken a photo of a text just because we are too lazy to take notes or type the text because taking photos takes less time than taking notes. Fortunately, in today's smartphone, we can directly apply OCR so that we can directly copy the text we took before without having to write or retype it. In Python, we can do that too just using a few lines of code. One of the OCR tools that are often used is Tesseract. Tesseract is an optical character recognition engine for various operating systems.

It was originally developed by Hewlett-Packard as proprietary software. To extract the text from image we use optical character recognition technique (OCR). Optical character recognition has become one of the most successful applications of technology in the field of pattern recognition and artificial intelligence. Optical character Recognition (OCR) is a process that converts scanned or printed text images, handwritten text into editable text for further processing. Speech synthesis is the artificial synthesis of human speech. A Text-To-Speech (TTS) synthesizer is a computer-based system that should be able to read any text aloud, whether it was directly introduced in the computer by an operator or scanned and submitted to an Optical Character Recognition (OCR) system.

1.2 Purpose of the project

The system supports multiple language recognition and translation by making use of advanced Cloud based services and Speech Recognition System. Quite obviously it encourages localization of languages. Thus the system is beneficial for almost everyone as it saves time in typing the text and makes translation independent of linguists.

People with disabilities like physical impairments often find it difficult in sending a text message or typing a statement in general. The system supports them by allowing them to speak their text through microphone and performing the conversion from audio to text. This system also helps people connect with any languages especially Indian Language from their own language.

1.3 Scope of the project

We are probably living in the worst time in our modern society where people refuse to help each other.We aim to make people with disabilities and travellers to find the best independent way of understanding different local languages by translating into required language.We aim at developing a interface that supports people with disabilities and travellers by allowing them to speak their text through microphone and performing the conversion from audio to text. This system also helps people connect with any languages especially Indian Language from their own language

1.4 Definition

1.4.1 TKINTER

Tkinter is the standard GUI library for Python. Python when combined with Tkinter provides a fast and easy way to create GUI applications. Tkinter provides a powerful object-oriented interface to the Tk GUI toolkit. Here we had used Buttons and Labels For user interaction.

1.4.2 Pytesseract

Python-tesseract is an optical character recognition (OCR) tool for python. That is, it will recognize and "read" the text embedded in images. Python-tesseract is a wrapper for Google's Tesseract-OCR. It is also useful as a stand-alone invocation script to tesseract, as it can read all image types supported by the Python Imaging Library, including jpeg, png, gif, bmp, tiff, and others, whereas tesseract-ocr by default only supports tiff and bmp. Additionally, if used as a script, Python-tesseract will print the recognized text instead of writing it to a file.

1.4.3 GTTS

gTTS (Google Text to Speech): A Python interface for Google's *Text to Speech* API. Create an *mp3* file with the **gTTS** module or **gtts-cli** command line utility. It allows unlimited lengths to be spoken by tokenizing long sentences where the speech would naturally pause.

1.4.4 Pil

The Python Imaging Library (PIL) adds image processing capabilities to your Python interpreter. This library supports many file formats and provides powerful image processing and graphics capabilities.

1.4.5 Python

Python is commonly used for developing websites and software, task automation, data analysis, and data visualization. Since it's relatively easy to learn, Python has been adopted by many non-programmers such as accountants and scientists, for a variety of everyday tasks, like organizing finances.

"Writing programs is a very creative and rewarding activity," says University of Michigan and Coursera instructor Charles R Severance in his book *Python for Everybody*. "You can write programs for many reasons, ranging from making your living to solving a difficult data analysis problem to having fun to helping someone else solve a problem."

Chapter 2

LITERATURE SURVEY

2.1 Image to Speech Conversion for Visually Impaired:

Asha G. Hagargundet et al carried out a work and they concluded that the basic framework is an embedded system that captures an image, extracts only the region of interest (i.e. region of the image that contains text) and converts that text to speech. It is implemented using a Raspberry Pi and a Raspberry Pi camera. The captured image undergoes a series of image pre-processing steps to locate only that part of the image that contains the text and removes the background. Two tools are used convert the new image (which contains only the text) to speech. They are OCR (Optical Character Recognition) software and TTS (Text-to-Speech) engines. The audio output is heard through the raspberry pi's audio jack using speakers or earphones.

Visual impairment is one of the biggest limitation for humanity, especially in this day and age when information is communicated a lot by text messages (electronic and paper based) rather than voice. The device we have proposed aims to help people with visual impairment. In this project, we developed a device that converts an image's text to speech. The basic framework is an embedded system that captures an image, extracts only the region of interest (i.e. region of the image that contains text) and converts that text to speech. It is implemented using a Raspberry Pi and a Raspberry Pi camera. The captured image undergoes a series of image pre-processing steps to locate only that part of the image that contains the text and removes the background. Two tools are used convert the new image (which contains only the text) to speech. They are OCR (Optical Character Recognition) software and TTS (Text-to-Speech) engines. The audio output is heard through the raspberry pi's audio jack using speakers or earphones.

2.2 Smart Reader for Visually Impaired People Using Raspberry PI:

D.Velmurugan et al carried a work and they concluded that this work proposes a smart reader for visually challenged people using raspberry pi. This paper addresses the integration of a complete Text Read-out system designed for the visually challenged. The system consists of a webcam interfaced with raspberry pi which accepts a page of printed text. The OCR (Optical Character Recognition) package installed in raspberry pi scans it into a digital document which is then subjected to skew correction, segmentation, before feature extraction to perform classification. Once classified, the text is readout by a text to speech conversion unit (TTS engine) installed in raspberry pi. The output is fed to an audio amplifier before it is read out. The simulation for the proposed project can be done in MATLAB.

The simulation is just an initiation of image processing ie., the image to text conversion and text to speech conversion done by the OCR software installed in raspberry pi. The system finds interesting applications in libraries, auditoriums, offices where instructions and notices are to be read and also in assisted filling of application forms. Though there are many existing solutions to the problem of assisting individuals who are blind to read, however none of them provide a reading experience that in any way parallels that of the sighted population. In particular, there is a need for a portable text reader that is affordable and readily available to the blind community. Inclusion of the specially enabled in the IT revolution is both a social obligation as well as a computational challenge in the rapidly advancing digital world today.

With the humongous amount of texts and written media available today it has been increasingly necessary to make these available to people from all walks of life specially catering to the visually impaired people, thus devising a system which can assist in this task is of prime importance. The proposed device aims to solve this predicament by using Raspberry Pi module B+ to convert printed and hand written into easily accessible playable speech. It also stores speech which can be replayed at a time deemed suitable. The main focus of this paper is to develop a smart reader system which coverts hand written and printed text to speech. It has been observed that the scanned texts were converted into easily audible speech heard via the speaker with high efficiency.

2.3 Camera based Text to Speech Conversion, Obstacle and Currency Detection for Blind Persons:

J. K. R. Sastry et al carried out a work the main object of this paper is to present an innovated system that can help the blind for handling currency. Methods/Statistical Analysis: Many image processing techniques have been used to scan the currency, remove the noise, mark the region of interest and convert the image into text and then to sound which can be heard by the blind. The entire system is implemented by using Raspberry Pi Micro controller based system. In the proto type model an IPR sensor is used instead of camera for sensing the object. Findings: In this paper a novel method has been presented using which one can recognize the object, mark the interesting region within the object, scan the text and convert the scanned text into binary characters through optical recognition.

A second method has been presented using which the noise present in the scanned image is eliminated before characters are recognized. A third method that can be used to convert the recognised characters into espeech through pattern matching has also be presented. Applications: An embedded system has been developed based on ARM technology which helps the blind persons to read the currency notes. All the methods presented in this paper have been implemented within an embedded application. The embedded board has been tested with different currency notes and the speech in English has been generated that identify the value of the currency. Further work can be done to generate the speech in different other both National and International Languages.

A camera based reader helps blind persons to read labels on the products and other handheld devices in their day by day lives1 . To differentiate the object from heavy backgrounds and other surroundings, an effective motion based method is used to define Region of Interest (ROI) in the camera view. In the obtained ROI text, localization and recognition are done2 . The printed context or scanned image is converted into computer recognition format by using Optical Character Recognition (OCR) so that it can increase the speed of operation3 . In model identification, all the characters are bounded and detached and then the end character image is directed to a pre-processor for removing the noise. All the characters are compared with a database of identified characters which are assembled together to form initial text pattern. The output is then given to the e-speak engine to convert text to speech and this output is given to blind users through earphones. The obstacle in the process is identified by using PIR sensor.

Chapter 3

PROBLEM IDENTIFICATION

3.1 Problem Statement

"To design and develop a system that supports multiple language recognition and translation by making use of advanced Speech Recognition System"

People with disabilities like physical impairments often find it difficult in sending a text message or typing a statement in general. The system supports them by allowing them to speak their text through microphone and performing the conversion from audio to text. This system also helps people connect with any languages especially Indian Language from their own language. The system basically uses two techniquesSpeech Recognition and Translation.

3.2 Project Scope

This project is intended to be used by visually impaired and travellers , so that they can easily understand any language by converting the text to their required language.The scope of this project includes conversion of image into text and further text into speech.

The traditional way of reading a book is through a printed copy. It's not possible to carry a physical copy of a book everywhere, that's when e-books came in handy. Printed books as well as e-books affect the eyes. Visual impairment is one of the biggest limitations for humanity, especially in this day and age when information is received by reading a lot of books (electronic and paper based). A study performed on "Listening or Reading: what's more effective?" shows us that people tend to understand more while listening to book rather than reading it. Considering the result our website converts hardcopy of a book into mp3 audible format. On this platform user just needs to upload the image of pages which will be converted into an audible form. The end user will not have to worry about eye problems. This web based platform converts the book into a saga which you can listen and thus carry the book anywhere by a few clicks.

Chapter 4
GOALS AND OBJECTIVES

4.1 Project Goals

The goal of this project is to present a smart device that assists the visually impaired and travelers which effectively and efficiently reads paper-printed text. It has wide range of features and functionality like providing a camera based assistive device that can be used by people to read Text document. This would help visually impared and travellers to reduce the problems in understanding different languages. This system will help its users in difficult situation. This interface is highly sensitive and easy to handle.

4.2 Project Objectives

- The assistant asks the user for input and keeps listening for commands. The time for listening can be set according to user's requirement.
- User interface is created to represent the image to text converter for computer-based system. In this it gives the components that are used for implementing the image to text converter.
- This assistant can be customized to have either male or female voice according to user’s requirement.
- The current version of the assistant supports features like Text translation, text synthesis and text extraction.

Chapter 5

SYSTEM REQUIREMENT SPECIFICATION

5.1 Software Requirements Analysis

A software requirements definition is an abstract description of the services, which the system should provide, and the constraints under which the system must operate. It should only specify only the external behaviour of the system and is not concerned with system design characteristics. It is a solution, in a natural language plus diagrams, of what services the system is expected to provide and the constraints under which it must operate.

Software Requirements

- Python 3.8
- HTML
- Windows / MAC OS / Ubuntu OS

Libraries Used:

- Urllib3
- Tk
- Playsound
- gTTs

5.2 Hardware Requirements Analysis

Hardware Requirements Analysis is to define and analyze a complete set of functional, operational, performance, interface, quality factors, and design, criticality and test requirements.

Hardware Requirements

- Processor: Intel Processor (32 bit).
- Ram: **128MB** or More.
- Hard Disk: Minimum 200MB.
- Computer System

Chapter 6

METHODOLOGY

6.1 Working-Flow of Application

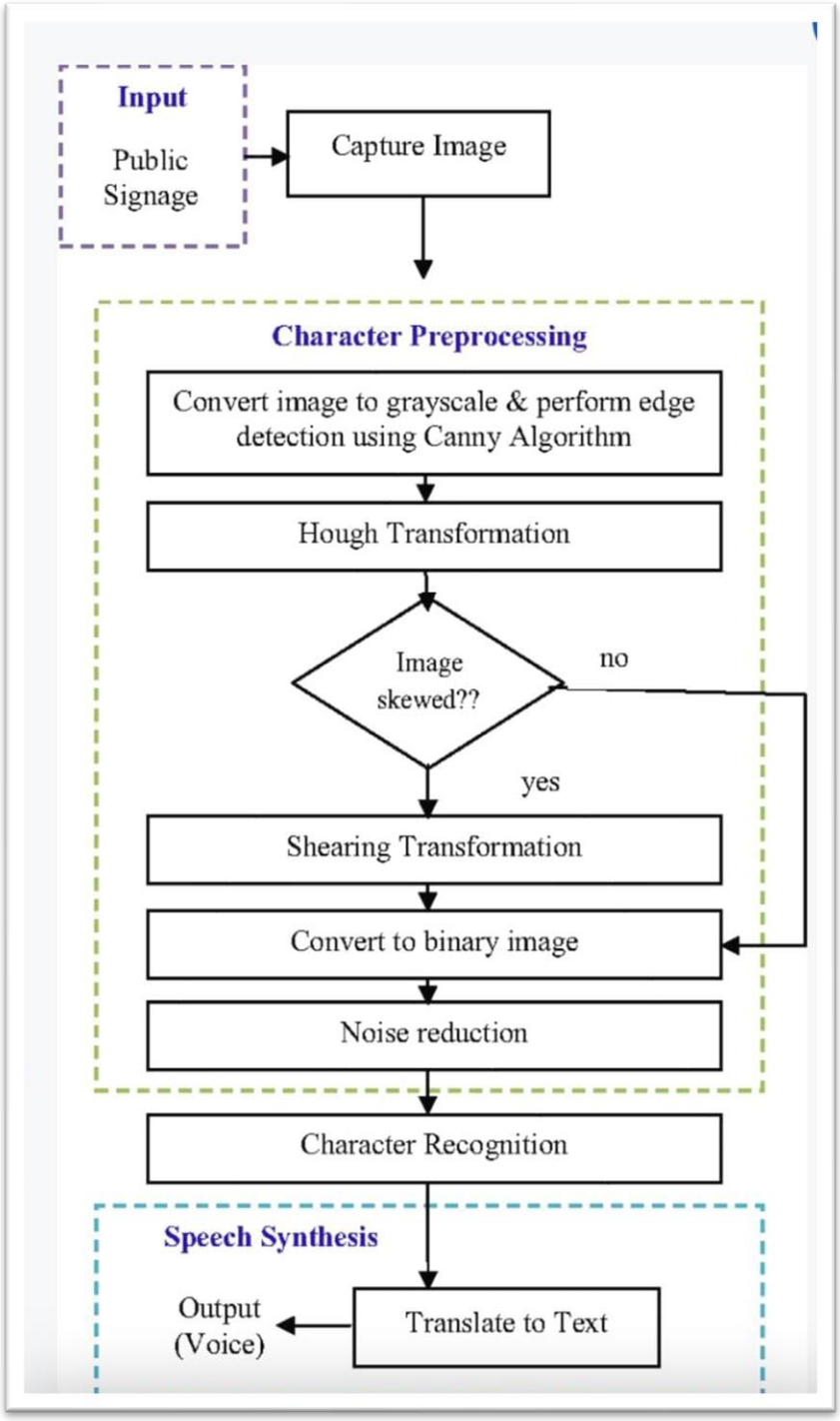

Fig 6.1.1: Flow chart of the Image converter

`6.2 Working

➢ Image to Text :

- Insert an image in Jpeg format.
- Extract the text in the image and give an output in text format.

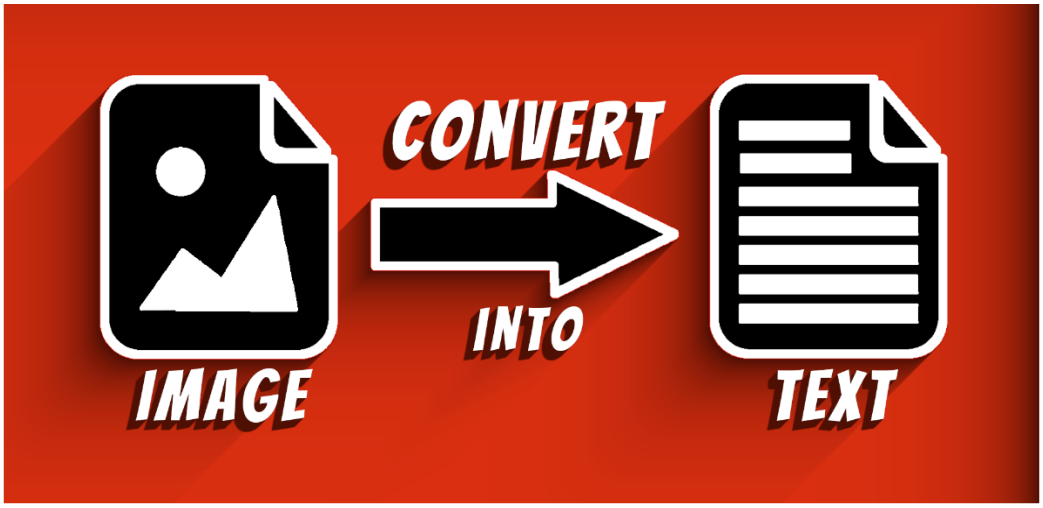

Fig 6.2.1: Image to text converter

➢ Image to Speech:

- Insert an image in Jpeg format.
- Extract the text in the image and give an output in text format.

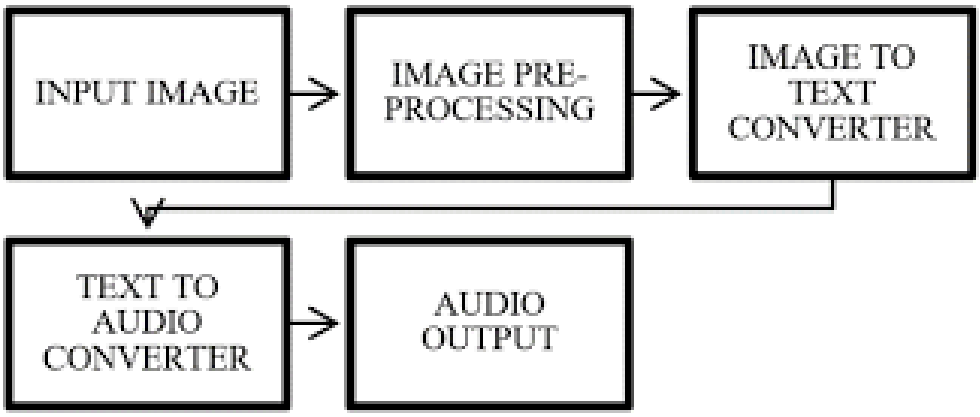

Fig 6.2.2: Image to speech converter

➢ Text to Speech:

- Insert an text in text box.
- Extract the text in the text box and give an output in audio format.

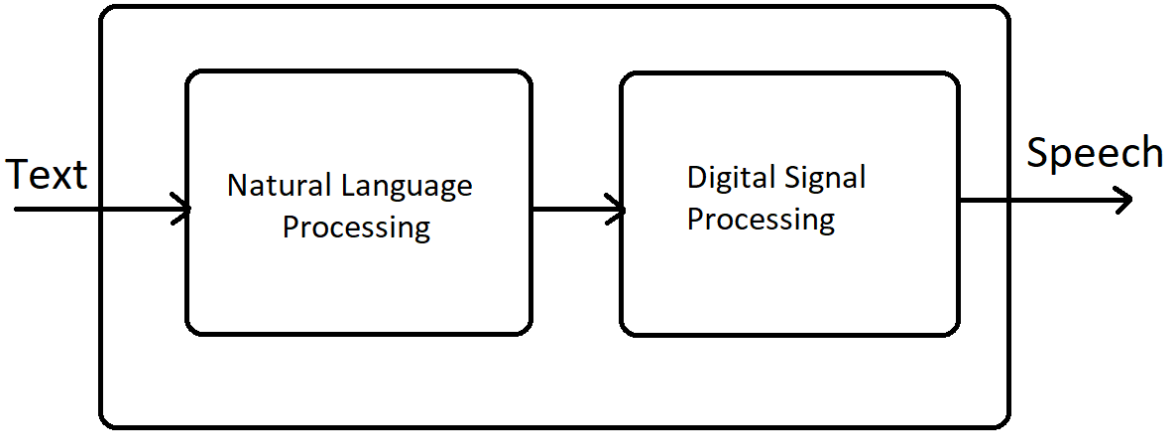

Fig 6.2.3: Text to Speech converter

Chapter 7

IMPLEMENTATION

```
#importing the Libraries
from gtts import gTTS          # used for converting text to speech
import PIL                     # Python Imaging Library
import gtts                    # Google's text to Speech API
import pytesseract             # used for image to text conversion using OCR
from tkinter import filedialog  # Used to provide GUI open/save feature
from tkinter import *
from PIL import Image,ImageTk   # used for handling image type file
import pyperclip
pytesseract.pytesseract.tesseract_cmd = r'C:\Program Files\Tesseract-OCR\tesseract.exe'
from playsound import playsound
import time
import os
try:
    from PIL import Image
except ImportError:
    import Image

import pytesseract
import cv2
##from picamera import PiCamera
from time import sleep
import pyttsx3
from translate import Translator

class clsTranslate():

    def translateText(self, strString, strTolang):
        self.strString = strString
        self.strTolang = strTolang
        translator = Translator(to_lang=self.strTolang)
        translation = translator.translate(self.strString)
        return (str(translation))

#
```

Fig 7.1: Converter.py

```
lbl3.place(x=160, y=550)
txt3 = Entry(window,width=40  ,bg="white" ,fg="black",font=('Helvetica', 20 , ' bold '))
txt3.place(x=390, y=575)
#Defining the funtions
def PicTexts():
    window.filename =  filedialog.askopenfilename()
    text4=txt3.get()
    # provides a dialog box for asking file to open and returns it's path
    img= PIL.Image.open(window.filename)      # opening image type file
    result= pytesseract.image_to_string(img)   # converting image to text
    objTrans=clsTranslate()
    result= objTrans.translateText(result, text4)
    res = result
    pyperclip.copy(res)
    message.configure(text= res)
    if(result==""):
        res = "Sorry!! Nothing recogonized"
        message.configure(text= res)

def picSpeechs():
    text3=txt3.get()
    window.filename =  filedialog.askopenfilename()
    img= PIL.Image.open(window.filename)
    result= pytesseract.image_to_string(img)









    if(result==""):
        res = "Sorry!! Nothing recogonized"
        message.configure(text= res)


    objTrans=clsTranslate()
    strTranslatedText= objTrans.translateText(result, text3)

    print(strTranslatedText)
```

Fig 7.2: Converter.py

```
lbl3.place(x=160, y=550)
txt3 = Entry(window,width=40  ,bg="white" ,fg="black",font=('Helvetica', 20 , ' bold '))
txt3.place(x=390, y=575)
#Defining the funtions
def PicTexts():
    window.filename =  filedialog.askopenfilename()
    text4=txt3.get()
    # provides a dialog box for asking file to open and returns it's path
    img= PIL.Image.open(window.filename)      # opening image type file
    result= pytesseract.image_to_string(img)   # converting image to text
    objTrans=clsTranslate()
    result= objTrans.translateText(result, text4)
    res = result
    pyperclip.copy(res)
    message.configure(text= res)
    if(result==""):
        res = "Sorry!! Nothing recogonized"
        message.configure(text= res)

def picSpeechs():
    text3=txt3.get()
    window.filename =  filedialog.askopenfilename()
    img= PIL.Image.open(window.filename)
    result= pytesseract.image_to_string(img)









    if(result==""):
        res = "Sorry!! Nothing recogonized"
        message.configure(text= res)


    objTrans=clsTranslate()
    strTranslatedText= objTrans.translateText(result, text3)

    print(strTranslatedText)



    res= gTTS(strTranslatedText)               # converting text to speech  .... Internet required
##    window.filename =  filedialog.asksaveasfilename()
    # provides a dialog box for asking file to save and returns it's path
    res.save('text1.mp3')     # inbuit audio saving function
    res = "Saved"
    file="text1.mp3"
    # import required module
    message.configure(text= strTranslatedText)
    time.sleep(2)
    playsound('text1.mp3',False)
    time.sleep(1)
    os.remove(file)
```

Fig 7.3: Converter.py

```
    objTrans=clsTranslate()
    strTranslatedText= objTrans.translateText(result, text3)

    print(strTranslatedText)



    res= gTTS(strTranslatedText)              # converting text to speech  .... Internet required
##    window.filename =  filedialog.asksaveasfilename()
    # provides a dialog box for asking file to save and returns it's path
    res.save('text1.mp3')     # inbuit audio saving function
    res = "Saved"
    file="text1.mp3"
    # import required module
    message.configure(text= strTranslatedText)
    time.sleep(2)
    playsound('text1.mp3',False)
    time.sleep(1)
    os.remove(file)


def TextSpeechs():
    text4=txt3.get()
    textInp= (txt2.get())
    objTrans=clsTranslate()
    textInp= objTrans.translateText(textInp, text4 )
    res= gTTS(textInp)
##    window.filename =  filedialog.asksaveasfilename()
    res.save('welcome.mp3')
    file ='welcome.mp3'
##    res.save(window.filename+ '.mp3')
    res = "Saved"
    message.configure(text= textInp)
    playsound('welcome.mp3',False)
    time.sleep(2)
    os.remove(file)
```

Fig 7.4: Converter.py

Chapter 8

RESULTS & SNAPSHOTS

8.1 RESULTS:

The proposed method successfully detects the text regions in most of the images and is quite accurate in extracting the text from the detected regions. Based on the experimental analysis that we performed we found out that the proposed method can accurately detect the text regions from images which have different text sizes, styles and color. Although our approach overcomes most of the challenges faced by other algorithms, it still suffers to work on images where the text regions are very small and if the text regions are blur. Below is the word-confidences of the words that we retrieve after performing the optical character recognition on the image which is tested in the experimental analysis section of this paper.

- People with disabilities like physical impairments often find it difficult in sending a text message or typing a statement in general.
- The system supports them by allowing them to speak their text through microphone and performing the conversion from audio to text.
- This system also helps people connect with any languages especially Indian Language from their own language.
- The system basically uses two techniques Speech Recognition and Translation.

8.2 SNAPSHOTS:

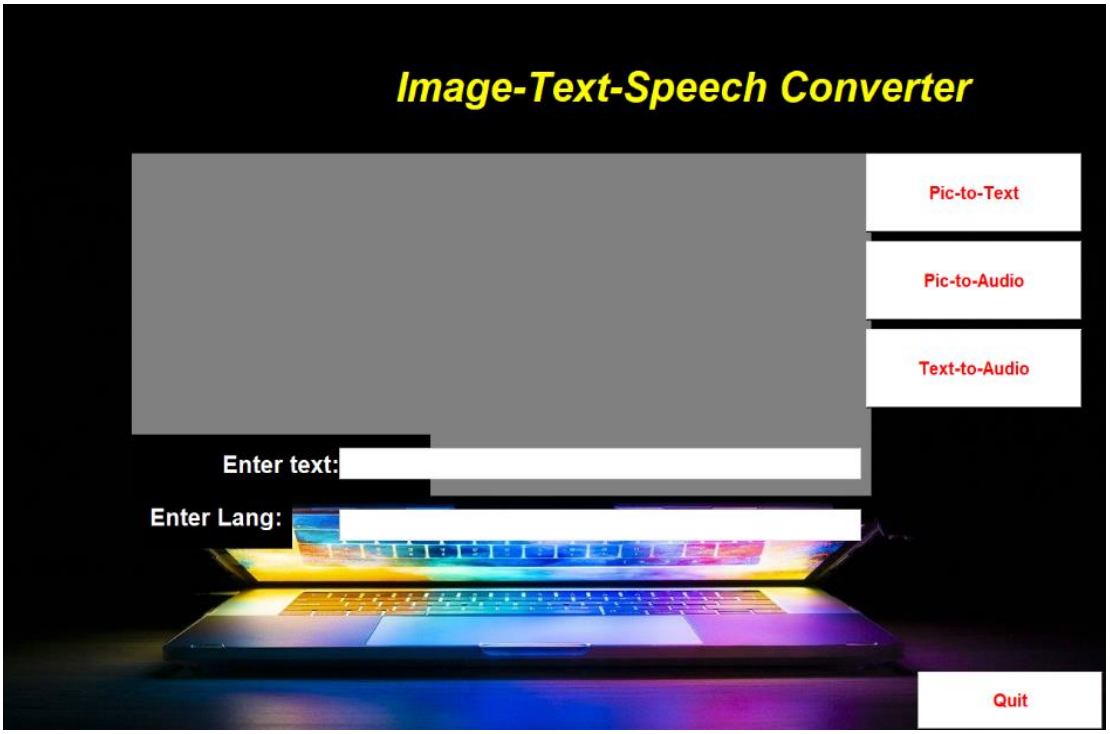

Fig 8.2.1: Clicking Image to convert it to text

Fig 8.2.2: Text to Audio Converter

Fig 8.2.3: Image converted to text

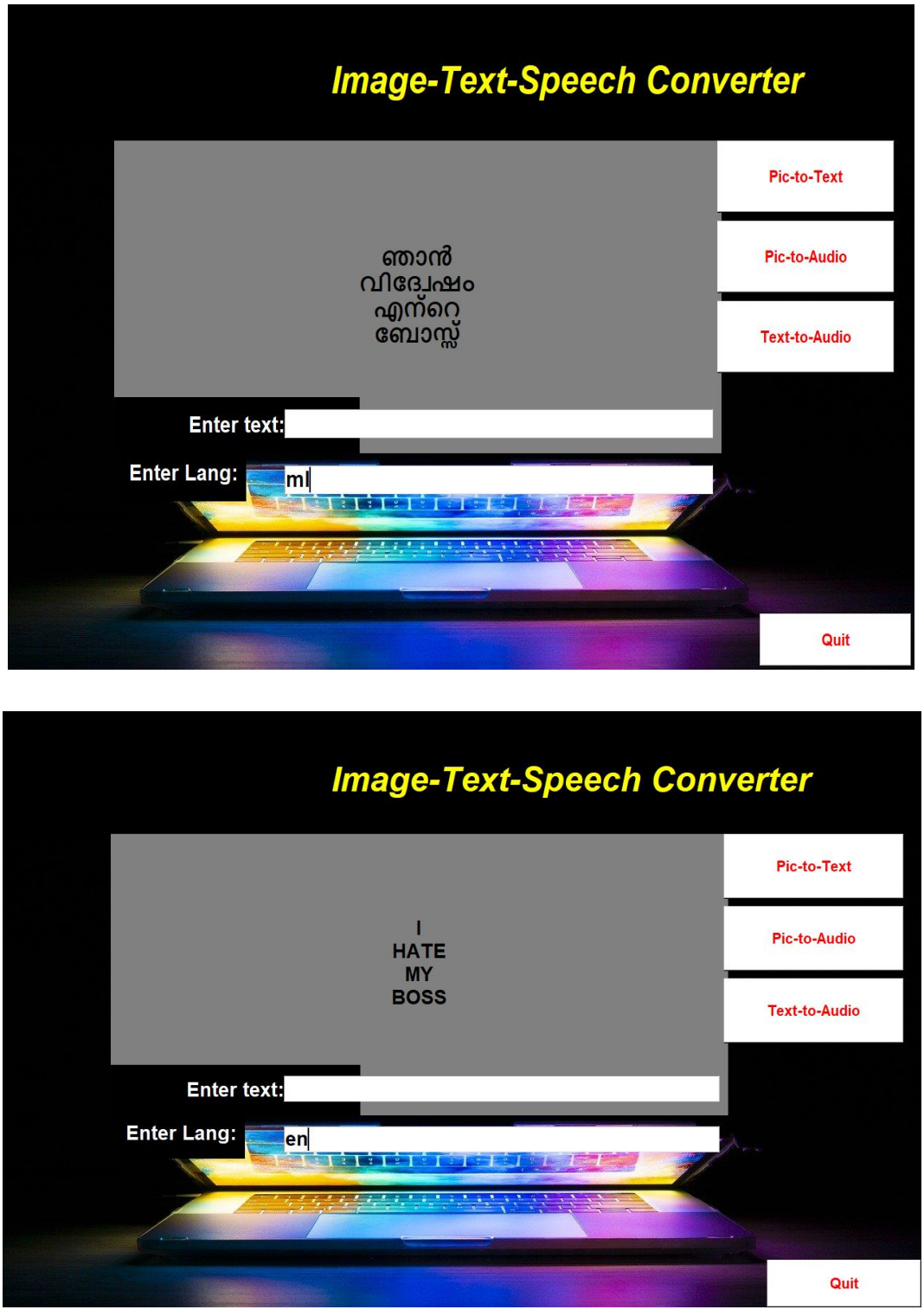

Fig 8.2.4: Pic to Text Converter

Chapter 9

APPLICATIONS

The basic approach in this project is to help traveller's and visually impaired to easily understand different languages. This project is a short term measurement for the present day problems and difficulties faced by travellers to understand language while travelling to different places. It gives an interface where the users can input the image to convert it onto text. Further this text can be converted into text in required language and even can be converted into speech in desired language. Our intention is to provide you with fastest and simplest way to convert languages into desired understandable language. This features for both visually impaired and travellers and making it an ultimate tool for all. The proposed design will deal with issues faced by travellers to understand different languages while travelling to different places and will help to solve them with technologically sound equipment's and ideas.

- Can be used for the better understanding of different languages.
- Can be used by travellers and visually impaired.
- Can be used for the safety of physically challenged people.

Chapter 10

CONCLUSIONS & FUTURE WORK

CONCLUSIONS:

- People with disabilities like physical impairments often find it difficult in sending a text message or typing a statement in general.
- The system supports them by allowing them to speak their text through microphone and performing the conversion from audio to text.
- This system also helps people connect with any languages especially Indian Language from their own language.
- The system basically uses two techniques Speech Recognition and Translation.

FUTURE WORK:

- The product provides an Optical Character Recognizer tool so for instance it is only able to capture textual data, due to this various images throughout the book are left uncaptured.Hence the audiobook doesn't involve any diagrams or images present in the book.
- Only one image can be uploaded, so if there are more pages the end user needs to upload them one by one which consumes more time.
- The whole text to speech conversion depends on internet connectivity, with slow internet connection the speech conversion might take long to generate or may even end up breaking the code.
- GTTS will need an internet connection to convert the text into an audio. So it can be slow then other offline APIs

Chapter 11

CONTRIBUTION TO SOCIETY AND ENVIRONMENT

Machine replication of human functions like reading is an ancient dream. However, over the last five decades, machine reading has grown from a dream to reality.Visually impaired people report numerous difficulties with accessing printed text using existing technology, including problems with alignment, focus, accuracy, mobility and efficiency.We present a smart device that assists the visually impaired and travelers which effectively and efficiently reads paper-printed text.

The proposed project uses the methodology of a camera based assistive device that can be used by people to read Text document. The framework is on implementing image capturing technique in an embedded system based on Raspberry Pi board.In this project we have proposed a text read out system for the traveler's and visually challenged. The proposed fully integrated system has a camera as an input device to feed the printed text document for digitization. Speech is probably the most efficient medium for communication between humans.

REFERENCES

[1] Keerthana S, Meghana H, Priyanka K, Sahana V. Rao, Ashwini B "Smart Home Using Internet of Things", proceedings of Perspectives in Communication, Embedded - systems and signal processing, 2017

[2] Sutar Shekhar, P. Sameer, Kamad Neha, Prof. Devkate Laxman, "An Intelligent Voice Assistant Using Android Platform", IJARCSMS, ISSN: 232-7782, 2017

[3] Rishabh Shah, Siddhant Lahoti, Prof. Lavanya. K, "An Intelligent Chatbot using Natural Language Processing", International Journal of Engineering Research , Vol.6 , pp.281-286, 2017

[4] Ravivanshikumar ,Sangpal, Tanvee, Gawand, Sahil Vaykar, "JARVIS: An interpretation of AIML with integration of gTTS and Python", proceedings of the 2019 2nd International Conference on Intelligent Computing, Instrumentation and Control Technologies (ICICICT), Kanpur, 2019.

[5] Dhiraj Pratap Singh, Deepika Sherawat, Sonia, "Voice activated desktop assistant using Python", proceedings of High Technology Letters, ISSN: 1006-6748, 2020.

1), pp.823- 826, March 2014.

APPENDIX-I

CERTIFICATES OF PAPER PRESENTED (if any)

APPENDIX-II

Journal Published Paper attachment (if any)