EYESHARP

Project Reference No.: *47S_BE_2004*

College ***:*** *Vivekananda College Of Engineering and Technology, Puttur, Dakshina Kannada*
Branch ***:*** *Department of Computer Science and Engineering*
Guide(s) ***:*** *Dr. Jeevitha B. K.*
Student(S) ***:*** *Mr. H. Dheeraj K.*
Mr. Abhiramdev N.
Mr. Ashwin Narayanan
Mr. Keerthi G.

Keywords

Visual impairment, text-to-speech conversion, accessibility, education, inclusivity, user- friendly design, Google API, audio processing, wearable device, portability, audio output

INTRODUCTION

In our lives, there are many people who are suffering from different diseases or handicap. According to NCBI more than one fourth of persons aged 50 years and above are visually impaired in India. These people need some help to make their life easier and better. The main goal of “EyeSharp” is to help blind people and people who have vision difficulties by introducing a new technology that makes them able to read the typed text. These glasses are provided with technology to scan any written text and convert it into audio text. Also, it can translate words from English to other languages using Google API. The goal of “EyeSharp” is helping those people in different life aspects. For example, these glasses effectively helpful in the education field. Blind people and people with vision difficulties can be able to read, study and learn everything from any printed text images. “EyeSharp” encourage blind people or people with vision difficulties to learn and succeed in many different fields.

This project combines elements of assistive technology, audio processing, and user- friendly design to create a solution that empowers blind individuals in their educational and everyday pursuits. By transforming written text into spoken words, the spectacle opens up new avenues for learning, information

consumption, and independence for thevisually impaired.

This project not only aligns with the principles of inclusivity and accessibility but also contributes to bridging the gap between the sighted and non-sighted communities. Through this endeavor, we aim to create a tangible and impactful solution that enhances the quality of life for blind individuals, fostering a more inclusive and equitable society.

OBJECTIVES

The “EyeSharp” project are to develop a technology that assists blind individuals and those with vision difficulties in reading typed text, thereby enhancing their quality of life. The project aims to create glasses equipped with the capability to scan written text and convert it into audio, making printed material accessible. Additionally, the technology will integrate Google API to translate text from English to other languages, broadening its usability for non-English speakers. The project seeks to support the educational pursuits of visually impaired individuals by enabling them to read, study, and learn from printed text images. By promoting independence and allowing visually impaired individuals to consume information and navigate their daily activities through audio-translated text, the project fosters inclusivity and accessibility. Combining elements of assistive technology, audio processing, and user-friendly design, “EyeSharp” aims to create a practical and impactful solution that improves the quality of life for blind individuals. Furthermore, the project encourages learning and success by providing the necessary tools and resources for visually impaired people to excel in various fields, thereby contributing to a more equitable and inclusive society.

METHODOLOGY

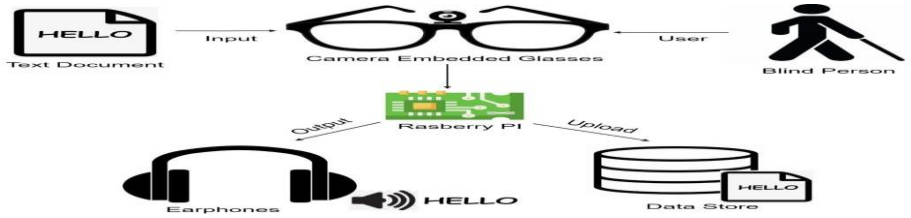

Fig 4.1: Methodology diagram

The Fig. 4.1 shows methodology diagram. The conversion of text from an image to speech involves several complex stages. Initially, users upload an image file containing textual content to the software platform. Subsequently, Optical Character Recognition (OCR) technology is employed to scrutinize the image, recognizing and deciphering the text it contains, converting it into a machine-readable format. Following OCR, the recognized text undergoes extraction from the image data, effectively isolating and organizing the textual information for further manipulation.

Once the text is extracted, it is passed through a Text-to-Speech (TTS) conversion function. Here, the written text is synthesized into speech, creating an audio file that faithfully represents the original textual content. This process utilizes sophisticated algorithms to generate natural-sounding speech, ensuring a seamless and intelligible audio output.

Finally, the resulting audio file is presented to the user, providing a means to listen to
access and comprehend the information visually depicted in the image through auditory means, catering to diverse needs and preferences. The process seamlessly integrates image processing, text recognition, and speech synthesis technologies to offer users a comprehensive and accessible method for interacting with textual content.

4.1 MODULES

A module is similar to a code library, comprisng a file with a collection of functions intended for inclusion in an application. Within our project, various modules serve distinct purposes.

4.1.1 Open CV

OpenCV, which stands for Open-Source Computer Vision Library, stands out as a versatile and widely-utilized module for tasks related to image and video processing. It offers a comprehensive array of tools and algorithms tailored for various computer vision applications, including object detection, feature extraction, image segmentation, and more. With support for multiple programming

languages such as Python, C++, and Java, OpenCV caters to a diverse community of developers and researchers. Its extensive documentation, coupled with active community support and cross-platform compatibility, enhances its appeal across different fields, including robotics, augmented reality, medical imaging, surveillance, and autonomous vehicles. From basic image manipulation to advanced techniques like deep learning-based object recognition, OpenCV's capabilities address a wide spectrum of needs.

4.1.2 Pyttsx3

Pyttsx3 stands out as a Python library designed to facilitate text-to-speech conversion, empowering developers to transform textual content into audible speech. Its compatibility extends across various platforms such as Windows, macOS, and Linux, making it exceptionally adaptable for cross-platform development needs. Offering a spectrum of voices and customizable settings, Pyttsx3 enables developers to fine-tune the speech output according to their specific preferences and requirements. Its user- friendly nature and seamless integration with Python scripts ensure a straightforward setup process, allowing developers to dive into their projects with minimal hassle.

4.1.3 Pytesseract

Pytesseract stands out as a Python interface for Google's Tesseract-OCR Engine, empowering developers to conduct Optical Character Recognition (OCR) on images. This tool streamlines the process of extracting text from images, finding applications in document analysis, text extraction, and automated data entry. Notably, Pytesseract boasts support for multiple languages and the capability to recognize diverse fonts and writing styles, which broadens its utility. Moreover, it accommodates images in various formats like JPEG, thereby catering to a wide array of needs. The seamless integration of Pytesseract with Python facilitates its adoption into existing workflows and scripts, accelerating the development of OCR solutions. Its reliability and precision render Pytesseract indispensable for tasks spanning document digitization to automated text analysis in diverse industries.

4.1.4 Socket

The "socket" module in Python is a powerful communication tool that helps create client-server applications. It allows developers to connect processes running on the same or different machines, allowing data to be exchanged across the network. Developers can implement TCP and UDP protocols via "sockets" and offer the option of TCP or UDP. This model provides flexibility in managing network operations by supporting both synchronous and asynchronous programming paradigms. Synchronous socket operations improve the performance of applications that handle multiple connections by allowing execution of asynchronous operations while blocking operations until completion. In general, the "socket" module serves as the building block of Python web programming, allowing developers to create a variety of web applications.

4.1.5 Importing the Necessary Libraries

Table 4.1 shows the imported necessary libraries. Importing the necessary library in Python involves bringing external modules or packages into your code environment to access their functionalities. It enables you to leverage pre-existing code and functionality to accomplish specific tasks without having to rewrite them

from scratch.

```
import cv2
import time
import pyttsx3
import pytesseract
from PIL import Image
import os
from pydrive.auth import GoogleAuth
from pydrive.drive import GoogleDrive
import socket
```

Table 4.1: Libraries imported

4.2 MATERIALS

4.2.1 Raspberry Pi 3B+

The Raspberry Pi 3B+ stands out as a highly adaptable sin-gle-board computer, celebrated for its compact design and impressive functionalities. Sporting a 1.4GHz quad-core ARM Cortex-A53 processor and 1GB of RAM, it delivers a substantial performance upgrade over its predecessors. Its integrated WiFi and Blue-tooth capabilities facilitate seamless communication with other devices.

4.2.2 Spectacles

Innovative wearable devices known as "spectacles for blind individu-als" have been developed to improve visual perception and assist with navigation for those with visual impairments. These devices are equipped with cameras and sensors that capture and process visual data. This information is then translated into auditory or tactile feedback, helping the wearer better understand their surroundings.

4.2.3 Headphones

Headphones are sound devices designed to be worn over the ears, delivering audio directly to the user. They are available in different styles, such as over-ear, on-ear, and in-ear. These devices are widely utilized for activities like listen-ing to music, watching videos, or making phone calls, providing a personalized audio experience without causing disturbance to others.

4.2.4 Power Bank

A power bank that works with the Raspberry Pi is a portable battery designed to provide the required power for the Raspberry Pi's operation. These power banks

usually include USB outputs that can power the Raspberry Pi's micro USB or USB-C port.

RESULTS

Fig 6.1: Initial design of the project

The Fig 6.1 shows initial design of the project, you can see the layout design of a prototype structure intended to aid visually impaired individuals in reading. This structure includes crucial elements like a Raspberry Pi, power supply, and headphones,all intricately connected to support the project's functionality.

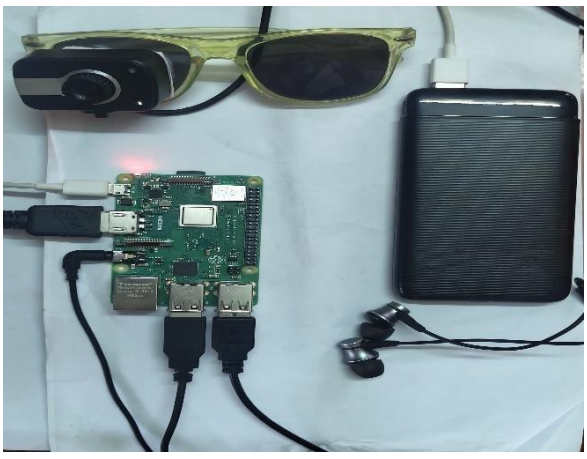

Fig 6.2: Complete design of the project

The Fig. 6.2 shows complete project's structure, as per the requirement has been finalized to include essential components such as a Raspberry Pi, headphones, a webcam mounted securely, and a power bank enclosed within a protective bag.

These elements are strategically arranged and connected to ensure optimal functionality. Tailored for the

visually impaired, the addition of a webcam introduces innovative possibilities, such as visual recognition or assistance, expanding the project's capabilities.

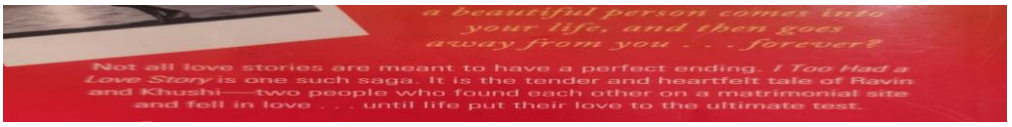

Fig 6.3: Text given for camera

The Fig 6.3 shows a picture taken of the image shown to the web camera of the hardware device. The text is as visible as possible because the page is well-lit and positioned within the frame. The page may have a variety of fonts, sizes, and styles that correspond to different textual parts, including headings, bullet points, and paragraphs.

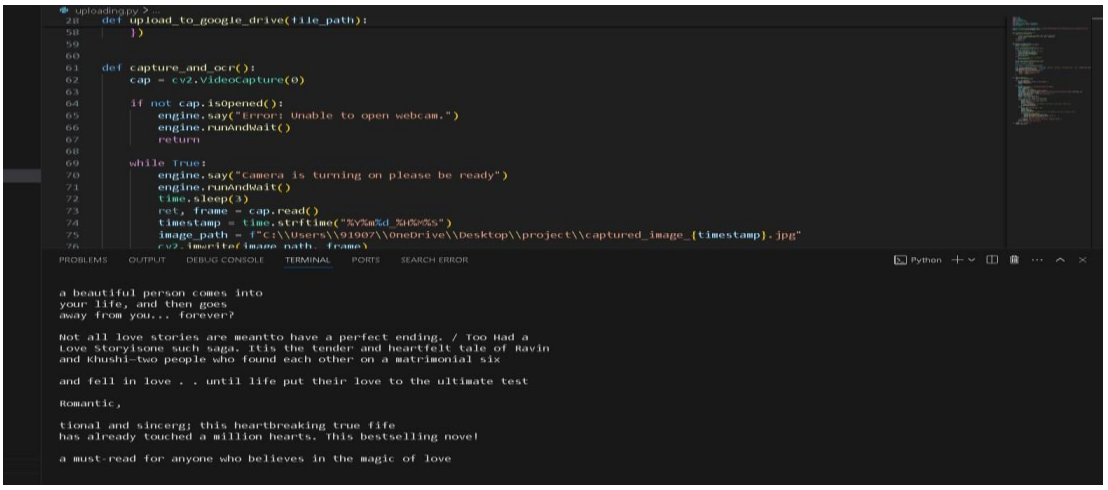

Fig 6.4: Recognized texts

The Fig 6.4 shows the identified text after optical character recognition (OCR) processing, which is shown in a terminal window. The identified text closely matches, with very few inconsistencies or mistakes, the information shown in Image 1. Every single line is presented with style, preserving the page's original design and organization. Next, text-to-speech technology is used to convert the identified text to audio format. After being converted, the audio file is smoothly played by headphones that are attached

`

to the gadget. Users can access the page's content from printed text to audio informationimproves accessibility and convenience.

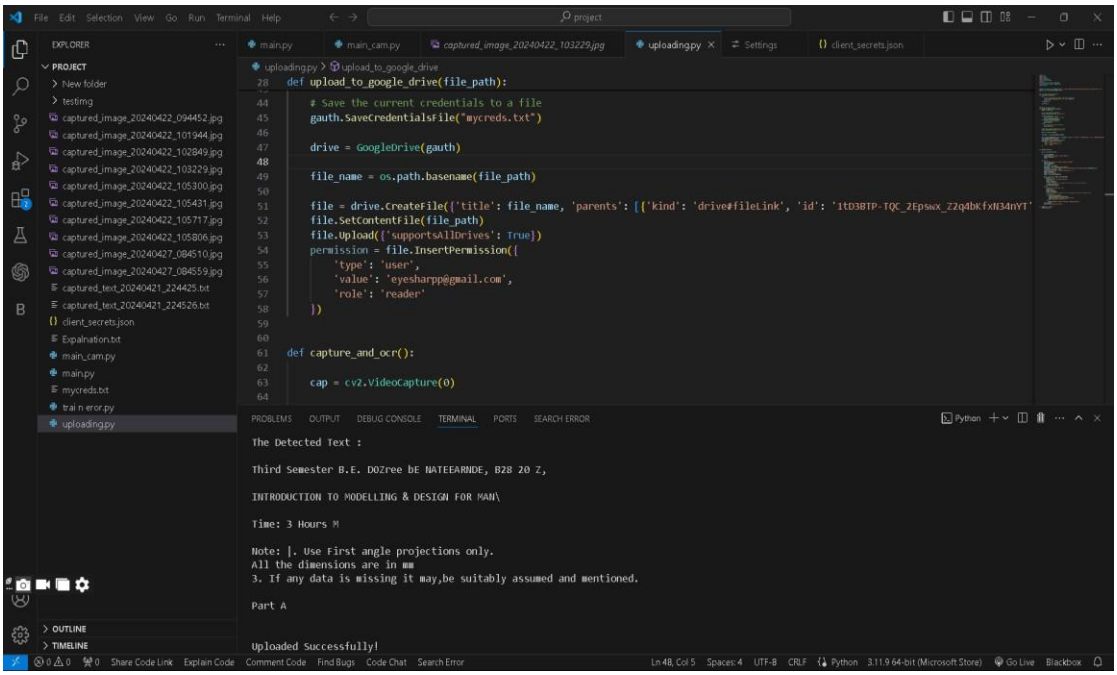

Fig 6.5: Uploading success

The Fig 6.4 shows uploading text to google drive, a terminal window displays the successful upload of a captured photo to the Google Drive account linked to eyesharpp@gmail.com. The terminal output likely includes details such as the file name or path, confirmation messages indicating successful transmission, and possibly a timestamp of the upload. The upload process occurs only if the captured text is clear if not, the program prompts the user to "Hold it properly," indicating that the image is unclear. In such cases, the image is automatically deleted from the local disk and not uploaded to Google Drive. Additionally, if the upload is attempted without an internet connection, a message informs the user that the process has been skipped or halted due to the absence of network connectivity. This message may be accompanied by an error notification explaining the issue.

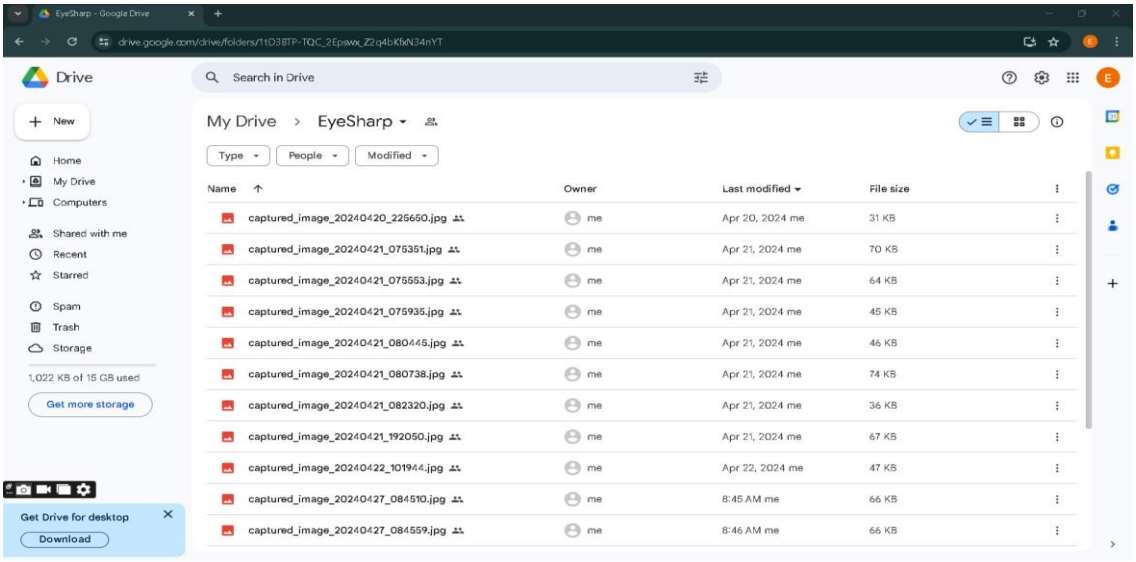

Fig 6.6: Drive folder

The Fig 6.5 shows folders present in the google drive, the Google Drive web interface is depicted, featuring a folder labeled "EyeSharp" within the user's Drive account. Inside this folder, a collection of accurately captured images is showcased, organized neatly either in a grid or list view.

CONCLUSION AND SCOPE FOR FUTURE ENHANCEMENT

7.1 Conclusion

The creation of an IoT device tailored for visually impaired individuals represents a significant advancement in improving their reading experience and overall quality of life. By incorporating state-of-the-art technology, such as text-to-speech conversion and tactile feedback mechanisms, the device closes the divide between printed text and auditory comprehension. This empowers users with greater independence and access to information. Rigorous user testing and feedback loops have honed the device's usability and functionality, ensuring it meets the varied needs of the visually impaired community. Moreover, its compact and portable design enables seamless integration into daily routines, whether at home, school, or work. Through IoT connectivity, users

can effortlessly access a wide range of digital content, further enhancing their reading experience. This initiative not only addresses a crucial societal need but also underscores technology's potential to promote inclusivity and equal opportunity. As we continue to refine and expand upon this innovation, we are committed to fostering a more accessible and inclusive world for all.

7.2 Scope for Future Enhancement

The "EyeSharp" project, designed as an IoT device for individuals with visual impairments, demonstrates significant promise. However, there exist several avenues for further development and enhancement:

- Enhancing the device's functionality and reliability involves expanding its capabilities to seamlessly integrate with a higher power supply. This ensures optimal performance for the Raspberry Pi without sacrificing speed, making it a promising avenue for improvement.
- Improving the system's abilities to smoothly incorporate a more robust power supply is a promising route to enhance the Raspberry Pi's functionality and reliability, guaranteeing optimal performance without compromising speed.

- Designing the system to accommodate reading in multiple languages offers great promise, enhancing its usefulness and accessibility for a wide range of users who require assistance with reading tasks.

- Creating a system design that seamlessly converts text into multiple languages through user-friendly button options offers significant potential for enhancing versatility and meeting diverse linguistic needs effortlessly.

INNOVATION IN THE PROJECT

The EyeSharp project is a pioneering initiative designed to assist individuals who are blind or have visual impairments by leveraging advanced technology to enhance their quality of life. According to the NCBI, more than one-fourth of persons aged 50 years and above in India are visually impaired, highlighting the critical need for innovative solutions in this area. EyeSharp aims to address this need by introducing smart glasses equipped with cutting-edge technology that can scan and convert any written text into audio, thereby enabling the visually impaired to access printed information effortlessly.

The EyeSharp glasses utilize state-of-the-art optical character recognition (OCR) technology to detect and read typed text, which is then converted into audio output. This functionality is further enhanced by integration with the Google API, allowing the glasses to translate text from English into various other languages, thus broadening their usability and appeal. This transformative capability empowers users by providing them with the tools to read, study, and learn from any printed material, significantly enhancing their educational opportunities and everyday experiences.

EyeSharp is meticulously designed to be user-friendly and accessible, ensuring that even those with limited technical skills can benefit from its features. The project embodies the principles of inclusivity and accessibility, striving to bridge the gap between sighted and non-sighted communities. By offering a practical and impactful solution, EyeSharp enhances the independence and information accessibility for visually impaired individuals, fostering a more inclusive and equitable society.