PROJECT REFERENCE ID: 46S_BE_0038

a) **Title of the project**: Multi-Lingual OCR, TTS And Google Assisted Eyewear For Visually Impaired Users

b) **Name of the College & Department**: Dept. of Electronics and Telecommunication, Dayananda Sagar College of Engineering, Bengaluru.

c) **Name of the Students & Guide**

Disha Jayaprakash Thotappala (dishajayaprakash94@gmail.com)

Diya Deb (diya.deb2506@gmail.com)

Robin K Mishra (robinkmishra2016@gmail.com)

Guide: Dr. Vinod B Durdi (vinodurdi-tc@dayanandasagar.edu)

d) **Keywords**: Assistive device, Image Recognition, Arduino, Visually Impaired People, Smart Glass

e) **Introduction / background**:

The use of language plays a vital role in human life, enabling knowledge transmission through books and records. Scientific papers are crucial for passing down theories and research to future generations. Special care and educational institutions exist for individuals with specific needs, but not all levels require specialized care. Many visually impaired individuals lack access to specialized schools, resulting in home-based learning. "Smart Glasses" aim to assist the blind and visually impaired by providing technology to convert written text into audio. These glasses have various applications, especially in education. The project utilizes camera-based "Smart Glasses" inspired by the needs of visually impaired individuals. Optical character recognition (OCR) technology is employed to convert images of text into machine-coded text, while text-to-speech synthesis converts the retrieved text into audible speech. The project proposes a novel text identification framework using connected component analysis and MSER methods for letter extraction and word recognition.

These research papers focus on various aspects of optical character recognition (OCR), translation, and assistive technologies for visually impaired individuals. The papers propose innovative approaches and hardware systems using Raspberry Pi for capturing and translating printed documents, segmenting, and recognizing Devanagari text, converting images to audio through OCR and voice technology, and utilizing OCR and text-to-speech synthesis for audio output. Additionally, the papers explore wearable visual assistance systems, object detection, tracking, and voice assistants for the blind, as well as integrating Google Assistant and Amazon Alexa on Raspberry Pi for smart speaker prototypes. These studies contribute to improving accessibility and enhancing the independence of visually impaired users.

f) **Objectives**:

- Build a system that is geared at helping people who are visually impaired.
- By contrasting the efficacy of multiple algorithms, choose the optimal algorithm for identifying the Region of Interest in an image.

- Create a prototype with a camera that can be used to take pictures while being placed on eyewear.
- Develop a system that can use the Google Speech module to convert text contained in an image (English/Hindi) into speech in any desired language.
- Include Google Assistant, which can help blind people with a range of daily tasks.

g) **Methodology** (about 30 lines) (materials, methods, details of work carried out, including drawings, diagrams etc).

Materials: Raspberry Pi Model 3B+, Raspberry Pi Camera module V2 (8 MP), Bluetooth speakers, buttons, monitor, keyboard, mouse.

Details of work carried out:

1. Raspberry Pi Setup: Started by setting up a Raspberry Pi board, connecting it to a display module, and installing Raspbian OS/

2. Camera Module: Attach a RPi camera module V2 to the Raspberry Pi. This camera will capture images of text for OCR processing.

3. OCR Software: Install an OCR library, Tesseract on the Raspberry Pi. Configure the software to process images and extract text from them.

4. Text Processing: Once the OCR software recognized the text, it will save the extracted text in a text format.

5. Text-to-Speech Conversion: Installed eSpeak text-to-speech synthesis library or software module on the Raspberry Pi.

6. Audio Output: Connected a small Bluetooth speaker Pi to provide audio output for the synthesized speech.

7. User Interface: Design and implement a user interface that allows users to trigger the OCR and TTS functionalities. This is achieved using buttons connected to the Raspberry Pi.

8. Google Assistant: Set up the Assistant on the Pi.

8. Integration: Combine all the components (Raspberry Pi, camera module, OCR software, TTS software, audio output, Google Assistant, and user interface) into a compact and wearable form factor, such as attaching them to a pair of glasses or creating a separate wearable device. This is the work that is remaining.

h) **Results and Conclusions**

Based on our comparisons amongst Canny Edge Detector, Sobel Edge Detector, Scharr Operator, Laplacian Operator; the Canny edge detection algorithm is often considered the most effective for OCR applications. It provides accurate edge localization, is robust against noise, and produces thin, continuous edges.

After successfully developing the individual components of our project, we have reached a crucial stage where we need to integrate them into a single cohesive unit to create a functional prototype. Integration plays a vital role as it brings together all the separate parts and ensures they work together seamlessly to achieve the desired outcome.

The integration process involves combining the hardware and software components that we have developed so far. On the hardware side, it includes assembling the various physical components such as the Raspberry Pi, camera module, display unit, and any other necessary components. On the software side, integration involves merging the different software modules that control the individual components. This includes integrating the OCR algorithm, edge detection algorithm, text-to-speech synthesis, and any other algorithms or functionalities that are part of the project.

Once the integration process is complete, we will have a working prototype that demonstrates the full functionality of our project. This prototype will showcase the seamless integration of the various components, allowing users to capture images, perform OCR, and have the text converted to speech. It will serve as a proof of concept and a foundation for further refinements and enhancements.

j) **Scope for future work**:

The future scope for a system aimed at making life easier for visually impaired users with real-time OCR and connectivity to the Google Assistant is promising. Here are some potential expansions and enhancements that could be considered:

- Advanced Object Recognition: Enhance the system's OCR capabilities to go beyond text recognition and incorporate object recognition. This would enable visually impaired users to identify and interact with physical objects, such as identifying household items, reading labels on products, or recognizing faces.
- Navigation Assistance: Integrate the system with navigation services to provide real-time directions and guidance for visually impaired users. This could involve incorporating GPS technology, indoor mapping systems, or augmented reality to provide audio-based navigation instructions, landmark descriptions, and obstacle detection.
- Smart Home Integration: Extend the system's connectivity to various smart home devices, allowing visually impaired users to control their home environment effortlessly. By integrating with smart speakers, thermostats, lighting systems, and other connected devices, users could use voice commands to adjust settings, turn on/off appliances, and receive audio feedback about the status of their home.
- Social Integration and Communication: Develop features that facilitate social interaction and communication for visually impaired users. This could include voice-based social media integration, audio description of images or videos shared by friends, and real-time transcription of conversations during phone calls or video chats.

The future scope of such a system is not limited to these. It is crucial to continually engage with visually impaired users and accessibility experts to identify areas of improvement and new opportunities for enhancing the system's functionality and usability.