Transforming Text to Audio: Enchancing Accessibility and Visually Imparied

A.Narayana Murthy[1], S.Swarna[2], K.Santhi Priya[3], Vyshtnavi[4], R.RaviChandra[5]

[1]Associate Professor, Department of Electronics and Communication Engineering, VSM College of Engineering, Ramachandrapuram, Andhra Pradesh, India
[2,3,4,5]UG Students, Department of Electronics and Communication Engineering, VSM College of Engineering, Ramachandrapuram, Andhra Pradesh, India

ABSTRACT

Engineering solutions have involved in everyone's life, most importantly are those aiming to help people with disabilities, however, the modern assistance devices with their current prices are not meeting the requirements of the market. This project is mainly focusing on people with visual impairments and more specifically their education life. It is presenting a concept of smart glasses to provide assistance in multiple tasks represented as modes to be chosen by the user. To prove the concept, this project implements only one mode which is reading using text detection techniques.

Taking into consideration the cost, this project is using the single board computer raspberry pi 4 B as the heart of the processing and the camera for image capturing and video recording. The video taken is processed using open CV, and the description of the live scenes for the text recognition mode or any other future implemented modes will be provided to the user in an audio format. Finally the results of this project show the design of the smart glasses prototype. The bigger the size the better results are obtained. This project is using the advantage of the wearable devices and the capabilities of the new raspberry pi therefore it can be taken further to provide assistance in more tasks.

Keyword: Speech Synthesis, Natural Language Processing (NLP),Voice Synthesis, Phonetic Transcription, Neural Networks ,Deep Learning, WaveNet, Prosody, Pitch and Intonation, Text Normalization, Speech Corpus, Voice Cloning, Speech Generation.

1. INTRODUCTION

Visual impairment can severely hinder access to information and independence. In recent years, wearable assistive technology has emerged as a promising solution to help individuals with such disabilities lead more independent lives. The primary goal of this project is to create a cost-effective, compact, and user-friendly device that assists visually impaired users by reading printed or digital text aloud. mounted on the glasses frame. The images are processed using OpenCV to extract and convert text into audio using text-to-speech libraries. This device, powered by a Raspberry Pi 4 Model B, provides a practical and scalable platform that can be expanded to support future features like navigation or object recognition. This project demonstrates how accessible technology can significantly improve the daily experiences of visually impaired individuals.

The smart glasses developed in this project capture real-time images using a camera

2. LITERATURE REVIEW

2.1 Assistive Technologies for the Visually Impaired

Various research efforts have explored wearable systems designed to aid visually impaired individuals through tactile, audio, or visual feedback. Previous works highlight the importance of integrating text recognition and speech synthesis to offer real-time assistance.

2.2 OCR and Text Detection in Wearables

OCR (Optical Character Recognition) has been successfully integrated into mobile applications and embedded systems for recognizing printed text. Research shows that combining OCR with Raspberry Pi and OpenCV provides a low-cost solution for image processing in real-time applications.

2.3 Audio-Based Feedback Systems

Text-to-speech (TTS) systems, when paired with OCR, can provide instant feedback for users. Several projects have leveraged Python- based TTS libraries on microcontrollers and single-board computers for real-world use cases.

3. EXISTING METHOD

3.1 OrCam MyEye

A commercially available wearable camera that reads text and recognizes faces and products. However, it is highly expensive and limited in customization.

3.2 Seeing AI by Microsoft

A mobile application that uses smartphone cameras to detect scenes, text, people, and currency. The drawback lies in the need for a smartphone, internet, and user interaction.

3.3 Google Lookout

An Android-based app offering real-time visual assistance. It is limited to smartphone users and is not wearable in the true sense.

3.4 Be My Eyes

A community-based solution that connects blind users with volunteers for real-time help, relying on human intervention and availability.

3.5 Aira Smart Glasses

Provides guided assistance using human agents and AR smart glasses, requiring subscriptions and live communication.

4. PROPOSED SOLUTION

The proposed solution is a cost- effective pair of smart glasses powered by Raspberry Pi 4 Model B that assists visually impaired individuals by converting text from images to speech. A webcam mounted on the glasses captures the surroundings in real-time. The system processes the captured video frames using OpenCV and performs OCR (Optical Character Recognition) to detect and extract text.

The detected text is then processed using a TTS engine, and the resulting audio is delivered to the user via headphones. An ultrasonic sensor is also integrated to detect nearby obstacles and enhance spatial awareness. The device is compact, lightweight, and can be powered using a rechargeable battery or power bank.

Fig.4.1.1: Block Diagram (*Insert block diagram illustration here showing Raspberry Pi, camera, ultrasonic sensor, speaker, and power supply interconnected*)

5. HARDWARE COMPONENTS

5.1 Raspberry Pi 4 Model B

A powerful single-board computer serving as the processing core. It handles image capture, processing, and audio output generation.

5.2 USB Webcam

Used for capturing real-time images and videos that serve as input for text detection.

5.3 Ultrasonic Sensor

HC-SR04 sensor used for detecting obstacles or objects in proximity to alert the user and prevent collisions.

5.4 Headphones/Speaker

Audio output device used to communicate the converted text to the user in real-time.

5.5 Power Bank/Battery

Used as a power source to make the system portable and usable on-the-go.

5.6 SD Card

Holds the Raspberry Pi OS, Python programs, and supporting libraries.

5.7 Jumper Wires

Used for connections between the Raspberry Pi and various sensors or components.

5.8 Glass Frame

A wearable frame to mount the camera, sensors, and Raspberry Pi for user convenience.

6. SOFTWARE

6.1 Raspberry Pi OS Installation

- Download Raspberry Pi Imager from official website.
- Flash the OS to SD card and boot the Pi.
- Set up environment with Python, OpenCV, Tesseract OCR, and pyttsx3 TTS libraries.

6.2 Program Execution Steps

- Open terminal on Raspberry Pi.
- Execute Python script for OCR-based text reading**.**
- Ensure webcam and audio output are working.
- Text is captured, extracted using OCR, and played via speaker.

7. ADVANTAGES

7.1 Real-Time Text Reading

Captures and reads out printed or handwritten text from surroundings instantly.

7.2 Portability and Wearability

Compact design mounted on a regular glasses frame enables mobility and ease of use.

7.3 Low Cost

Affordable components make it accessible compared to commercial alternatives.

7.4 Expandability

The modular design supports future enhancements like navigation, object detection, etc.

8. APPLICATIONS

8.1 Educational Assistance

Helps visually impaired students read books, documents, and signs.

8.2 Navigation Aid
With future development, can assist in obstacle detection and route guidance.
8.3 Public Accessibility
Helps users identify and read signboards, price tags, menus, etc.
8.4 Daily Living Support
Improves independence in daily tasks like reading prescriptions or currency notes.

9. EXPERIMENTAL RESULTS
The prototype was tested in different lighting and font conditions. It successfully detected printed text from books, posters, and labels. The OCR algorithm efficiently extracted and read text aloud with minimal delay. Larger image areas produced better accuracy, especially with clear and high-contrast fonts. The system was effective for indoor and moderately lit environments.

10. FUTURE SCOPE
The project can be expanded to include multiple modes such as:
- Scene Description: Describing objects or people in the camera frame.
- Real-Time Navigation: Integration with GPS and path detection.
- AI-Based Object Recognition: Identifying daily objects or obstacles.
- Voice Command Interface: Allowing the user to switch between modes using voice.
- Cloud Connectivity: For storing user data or improving OCR accuracy via online engines.

11. CONCLUSION
The Smart Glasses system provides an effective assistive solution for visually impaired individuals by integrating text detection and audio feedback in a wearable format. Using components like Raspberry Pi 4, webcam, and OpenCV, the system demonstrates efficient performance for reading printed text in real-time. Its affordability, scalability, and user-friendliness make it a valuable tool for education and independent living. Future enhancements will allow the device to assist with a broader range of tasks, making it a vital companion for visually impaired users.

12. REFERENCES
1. Patel, A., & Sharma, R. (2022). Assistive Technologies for the Visually Impaired. IJERT, 11(3), 65-72.
2. Gupta, S., & Verma, P. (2021). OCR and Text-to-Speech Integration Using Raspberry Pi. IEEE Access, 9, 11290-11297.
3. Singh, A., & Kumar, R. (2020). Wearable Technologies for Accessibility. International Journal of Smart Devices, 6(1), 34-41.
4. OpenCV Documentation: https://opencv.org/
5. Tesseract OCR Documentation: https://github.com/tesseract-ocr/
6. Raspberry Pi Documentation: https://www.raspberrypi.com/do cumentation/