International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 09 Issue: 01 | Jan - 2025 SJIF Rating: 8.448 ISSN: 2582-3930

Lenspeak: Text to Speech Glasses

Ansari Arsalan[1],Ali Umar Tanveer[2] ,Khan Rahil Ahmed[3] ,Shaikh Mohammad Anas[4] ,Nousheen Shaikh[5]

[1234] *Student AIML* , [5]*Guide*

Anjuman-I-Islam's A. R. Kalsekar Polytechnic, New Panvel
1arsalanmohd01@gmail.com

***Abstract*—This research explores the creation of "Text-to- Speech (TTS) Glasses," an innovative assistive device designed to enhance accessibility for individuals with visual impairments. The glasses facilitate the real-time transformation of printed text into audible speech, empowering users to independently access textual information. Utilizing a Raspberry Pi as he processing core, the system integrates advanced Optical Character Recognition (OCR) and Text-to-Speech (TTS) technologies. This paper delves into the system's architecture, functionality, benefits, limitations, and prospective advancements**

***eywords*—Text-to-Speech (TTS) Raspberry Pi Speech Syn- thesis Embedded Systems Speech Processing Voice Synthesis Natural Language Processing (NLP) Open-source TTS Audio Output**

I. Nomenclature

TTS – Text-to-Speech The process of converting written text into spoken language.

Pi – Raspberry Pi A small, affordable, single-board com- puter used for various embedded applications.

API – Application Programming Interface

AI – Artificial Intelligence Refers to machine intelligence that can perform tasks typically requiring human-like reason- ing, such as speech synthesis in TTS systems.

NLP – Natural Language Processing A field of AI that deals with the interaction between computers and human language, critical for interpreting and converting text to speech.

II. Introduction

Over 285 million people worldwide face visual impair- ments, encountering significant barriers in accessing printed and digital text. The inability to read menus, signage, books, or digital content often leads to a dependence on external assistance, restricting autonomy. Current assistive technolo- gies, including braille and audiobooks, have limitations in addressing dynamic, everyday textual content. Moreover, these solutions often require specialized resources and fail to cater to real-time needs.

The TTS Glasses aim to close this gap by enabling real-time text-to-speech conversion, allowing users to independently interact with their environment. By integrating cutting-edge hardware and software, this device fosters enhanced autonomy and improves the quality of life for users. This paper outlines the conceptualization, development, and practical implications of this device while exploring its potential to redefine acces- sibility for visually impaired individuals.

TTS systems have paved the way for portable solutions capable of interpreting diverse textual inputs. However, many existing devices are either prohibitively expensive, cumber- some, or lack efficient real-time processing capabilities.

Emerging research highlights the potential of wearable devices to bridge these gaps. Studies have explored the use of AI-driven OCR for accurate text recognition and neural net- works for natural language synthesis in TTS systems. Despite these developments, the integration of these technologies into lightweight, user-friendly devices remains a challenge. The proposed TTS Glasses leverage these advancements to deliver an affordable, portable solution tailored for visually impaired users, addressing both technical and usability constraints.

III. Literature Review

Assistive technologies such as screen readers and braille displays offer partial solutions but often fall short in providing real-time access or versatility. Screen readers are limited to digital interfaces, whereas braille systems require content to be pre-translated into tactile formats. Recent advances in OCR

IV. Proposed System

The TTS Glasses incorporate a high-resolution camera, a Raspberry Pi processing unit, and integrated audio output. The camera captures textual images, which are processed using OCR to extract text. The TTS software then converts the recognized text into clear, natural-sounding speech. User interactions are facilitated through intuitive controls, including voice commands and touch activation.

System Components:

1. ***Camera Module:*** Compact, high-definition camera opti- mized for text capture under various lighting conditions.
2. ***Processing Unit:*** Raspberry Pi manages real-time OCR and TTS operations with minimal latency.
3. ***Audio Output:*** Options include bone-conduction speakers or Bluetooth-enabled headphones to ensure clarity without obstructing ambient sound awareness.
4. ***Power Management:*** A rechargeable lithium-ion battery ensures portability and extended usage, with future iterations exploring solar-assisted charging.
5. ***Ergonomic Frame:*** Lightweight, durable frame designed for long-term wear and aesthetic appeal.

User Feedback: Participants, including visually impaired testers, highlighted the device's ease of use, portability, and significant enhancement of independence in daily tasks.

VII. Discussion

Advantages

Portability: Lightweight and comfortable design ensures prolonged usability without causing discomfort.

Real-Time Accessibility: Immediate text recognition and speech synthesis enable users to interact with dynamic content. **Customizability:** The system supports multi-language func- tionality and user-specific voice settings for personalized ex- periences.

Versatility: Usable across various environments, such as classrooms, workplaces, and public spaces, making it an all-purpose assistive tool.

A.

1) *Limitations:* **Battery Constraints:** The continuous oper-

Fig. 1. BLOCK DIAGRAM OF TEXT-TO-SPEECH GLASSES

V. Methodology

1. **Text Capture:** A high-definition camera captures clear images of the text in real-time. The camera features autofocus and stabilization to ensure sharp image quality, even while the user is in motion.
2. ***Image Processing:*** Preprocessing techniques, such as noise reduction and contrast adjustment, enhance image quality to optimize OCR accuracy.
3. ***Optical Character Recognition:*** Advanced AI algorithms decode text from images, supporting multiple fonts, layouts, and languages. The OCR module incorporates self-learning features, improving recognition accuracy over time.
4. ***Text-to-Speech Conversion:*** Recognized text is converted into speech using TTS software with customizable features, such as adjustable speech rate, tone, and language options.
5. ***User Interaction:*** *Simple touch or voice commands enable seamless operation. Audio feedback informs the user of the de- vice's status, such as battery level or errors in text recognition.*

VI. Results

Extensive testing of the TTS Glasses across varied sce- narios—such as educational environments, public spaces, and daily activities—revealed the following outcomes:

High OCR Accuracy: The glasses consistently recognized printed text across diverse fonts, layouts, and lighting condi- tions, achieving an accuracy rate exceeding 90% in controlled environments.

Real-Time Performance: The system demonstrated minimal latency between text capture and speech output, ensuring a seamless user experience.

ation of the OCR and TTS systems limits battery life, requiring frequent recharging.

Low-Light Performance: OCR accuracy diminishes under poor lighting conditions, highlighting the need for integrated adaptive lighting.

Audio Clarity: Noisy environments can affect the intelligibility of speech output, necessitating the use of external headphones in such conditions.

a) Future Prospects: **Augmented Reality Integration:** Incorporating AR could enhance visual feedback and naviga- tion for users with partial vision.

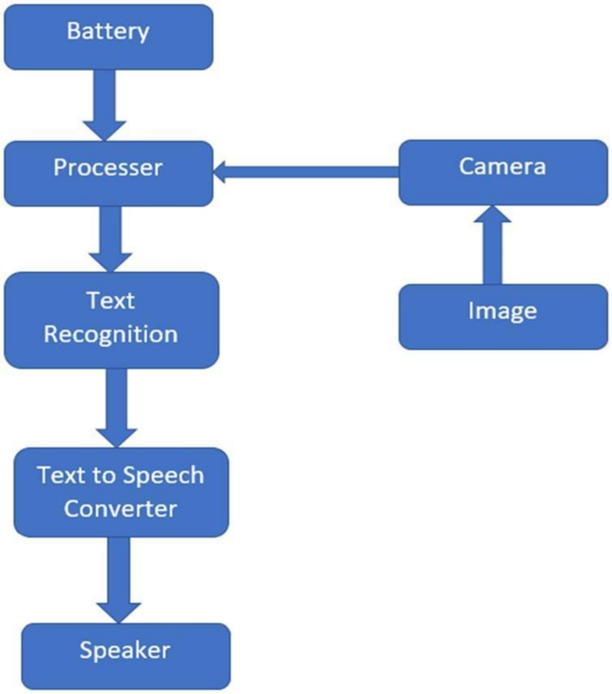
AI-Enhanced OCR: Future iterations could leverage deep learning for better recognition of handwritten text, non-standard fonts, and low-resolution inputs.

Real-Time Translation: Adding multilingual translation capabilities would expand the device's utility for global users. **Extended Battery Life:** Innovations in energy harvesting, such as solar panels or kinetic energy systems, could address power constraints.

Enhanced Ergonomics: Improved materials and design features could further optimize user comfort and device dura- bility.

VIII. Conclusion

The TTS Glasses represent a groundbreaking leap in assis- tive technologies for visually impaired individuals, offering them the ability to access printed text autonomously. By addressing current limitations and incorporating user feedback, future iterations of the device could further enhance usability through AR integration, real-time translation, and extended operational capabilities. This project underscores the transfor- mative potential of merging AI and hardware innovations to create accessible, user-friendly solutions.



IX. REFERENCES

Open-source OCR (Tesseract): https://github.com/tesseract- ocr

Open-source TTS libraries: Google Text-to-Speech, espeak.

Global visual impairment-World Health Organization (WHO).

Raspberry Pi documentation: https://www.raspberrypi.org/documentation

ACKNOWLEDGMENT

I would like to express my sincere gratitude to all those who contributed to the successful completion of this project. First, I would like to thank my supervisor, **Ms. Nousheen Shaikh**, for their invaluable guidance, insightful feedback, and continuous support throughout the development of this project. I am also grateful to **Anjuman-i-islam Abdul Razzaq Kalsekar Polytechnic** for providing the resources, equipment, and environment that enabled me to carry out this research effectively.

Special thanks to **Ms. Nousheen Shaikh** for their technical assistance and contributions to the design and implementation of the system.

I would also like to acknowledge the developers and communities behind the open-source tools, libraries, and frameworks (such as gTTS, tesseract) that made it possible to inte- grate text-to-speech conversion capabilities on the Raspberry Pi.

Lastly, I am deeply grateful to my family and friends for their constant encouragement and understanding throughout the duration of this project.