Grenze International Journal of Engineering and Technology, June Issue

Deaf Assist Smart Glasses

Sarang Kulkarni[1], Kush Mantoo[2], Digvijay Lagad[3], Chaitanya Lawande[4] and Sahil Markande[5]
[1-5]Vishwakarma Institute of Technology, Pune-27, Maharashtra, India

***Abstract*— The Deaf Assist Smart Glasses is a wearable support device designed for individuals with hearing disabilities through speech-to-text in real-time. It captures spoken words through an integrated microphone, converts the voice into written form via AI powered speech recognition, and displays this text on a small OLED screen on the spectacle frame. It does not rely on hands to do this while allowing the user to follow conversations in different environments, for example, school, workplace or public. It uses a Wi-Fi or Bluetooth connection to search speech-to-text apps either on the device or online; this decreases latency and improves accuracy. With the combination of IoT and integrated technologies, the Deaf Assist Smart Spect provides individuals smarter options of conversations, fostering independence, inclusion and shared experiences for individuals with hearing disabilities.**

***Keywords*— Deaf Assist Smart Spectacles, Hearing Disabilities, Speech-to-text, AI speech recognition, OLED display, hands-free communication, IoT, accessibility, inclusion.**

I. INTRODUCTION

Through DeafAssist Smart Spect's advanced wearable technology, individuals with hearing impairments can access real-time transcriptions of spoken language, enhancing communication in everyday interactions [1], [3]. Deaf or hard-of-hearing individuals often rely on sign language or written communication, which, although effective, can create barriers in fast-paced or multi-speaker environments [11]. The smart spectacles provide a hands-free, non-intrusive solution where transcribed speech appears directly before the user's eyes, allowing more natural and accessible conversations [2], [4].

A small microphone integrated into the frame of the glasses picks up speech input, which is then processed through an AI-based speech recognition system to display it as a text output. The text is then shown on a small OLED screen embedded within the lens to allow users to follow conversations in informal, formal, or public settings without additional devices [1], [3], [10]. The system's wireless connection allows the processing to be done locally or via cloud-based APIs, which increases both accuracy and response time [8], [12]. Through the integration of IoT and embedded technologies, DeafAssist Smart Spect provides accessibility, independence, and inclusivity for users with hearing related challenges [5], [9]. The wearable technology is also a significant advancement in the journey towards a more inclusive communication ecosystem, as it allows users to confidently participate in both occupational and social environments [11].

Hearing impairment affects hundreds of millions around the world, leading to social isolation and communication challenges [11], [14]. Although traditional hearing devices can increase sound in general environments, they still can display limitations with background noise and simultaneous voices, among other factors [5]. Advances in artificial intelligence, the Internet of Things (IoT), and wearable computing has led to the development of a new class of assistive devices, which place speech recognition and display technology directly in everyday wearable technology [8], [9].

Grenze ID: 01.GIJET.12.2.289_21

Smart glasses have demonstrated considerable capabilities as combining compact hardware with software based on machine learning would provide users with seamless and real-time access to displayed content [1], [2], [13].
The purpose of DeafAssist Smart Spectacles is to close the communication gap with a fully integrated speech-to-text experience. The smart spectacles provide an alternative to traditional hearing aids because they are specifically designed to convert speech to visual text using embedded microphones and real-time artificial intelligence transcription engines [3], [10]. As a result, users can understand speech in disastrous environments, such as multi-speaker events in classrooms, conference rooms, or social gatherings [5], [8].
Along with increasing accessibility, the project prioritizes affordability, user independence, and ease of delivery/implementation/deployment. Leveraging wireless connectivity and compact OLED display technology allows for flexibility in the processing to be completed locally and/or in the cloud, with a balance of speed/latency and accuracy [8], [9], [12]. Therefore, DeafAssist Smart Spectacles serve not only as an assistive device but as a movement toward digital empowerment and inclusive design for people with hearing impairments [1], [11].
This study extends emerging trends in the field of human-computer interaction and wearable computing by combining hardware, software, and networking into an integrated system [3], [8], [11], [9], [14], [15]. The system provides direct, real-time transcription on the glasses reducing the reliance on other intermediary devices, encouraging social inclusion, and demonstrating how accessible technologies can create quality of life for millions [3], [8], [11].

II. Literature Review

SpecAssist [1] is a “wearable” system consisting of smart glasses, a headset display, and a microphone that captures sound. It performs real-time speech recognition and transcription, displaying the text to aid communication for the hearing impaired. The system pairs with smart phones and uses the interment to access www text translation engines. Inputs can be given via a mobile application. SpecAssist employs deep learning together with speech recognition automatic and natural language processing (NLP) along with the internet of things (IOT) to provide low latency service to the hearing impaired.
Enssat is a smartphone bilingual hearing aid application designed for Arabic and English speakers to assist the Deaf and Hard of Hearing (DHH) population [7]. Enssat is capable of real-time transcription, translation, and alert management through the use of wearable technology. A user test that involved DHH participants tested the application’s transcription, translation, and alerting functions and proved its effectiveness and ease of use for everyday communications for DHH users.
A smart hearing aid system is proposed in [6] that implements a microphone array with sophisticated voice processing algorithms to help hearing-impaired people. The smart device captures sound and processes it in every direction. It also gives priority to the critical speaker while supressing other unwanted sounds, including other speeches. The fine-tunable smart parameters enhanced through the smartphone interface allows effortless comprehension of speech in high noise conditions.
The paper [5] describes the implementation of machine learning and artificial intelligence technologies to enhance speech understanding in contemporary hearing aids, especially in the presence of background noise. It analyzes block user and automatic speech comprehension enhancement by hearing aid signal processing and deep neural network interfacing with smartphones and wireless accessories. The paper further looks into health and wellness monitoring features added by sensors and AI, thus contributing to health monitoring for the users.
A paper [3] presents assistive glasses designed to provide real-time speech-to-text translation for individuals with hearing disabilities. The system consists of a microphone for speech capture, an AI-enabled speech recognition system for transcription, and a glass integrated display unit for text output directly to the users’ line of sight. They are intended to be mounted on standard eyeglasses, which increases their accessibility to many potential users. This technology facilitates easier communication and interacts with users so as to remove barriers constructed by different forms of communication.

III. Methodology

A. System Architecture:

The hardware setup of DeafAssist Smart Specs is designed for speech recognition within the wearable and portable framework. The system responsible forcapturing audio, the input processes, and the output display is based on an ESP32 microcontroller. The speech input is provided by the MAX9814 microphone module, which is also known for its automatic gain control (AGC) feature that allows for different levels of noise reception. Refer figure below:

Fig. 1. System Architecture

Locally, the received audio can be processed, or they can be wirelessly transmitted for processing using Wi-Fi or Bluetooth for cloud-based speech recognition. The resulting text is sent to a small OLED screen (SSD1306) which is placed on the frame of the spectacles allowing users to read the text of the conversations almost instantly, dynamically enhancing user experience. Continuous usage is made easier as power is supplied through a direct connected power source with no battery, removing the need to recharge, making it ideal for demo applications.

The hardware aspect of this model consists of a few components which are listed as follows:

- Microphone
- ESP 32
- OLED Screen
- Power Source / USB Supply
- Wearable Glasses Frame

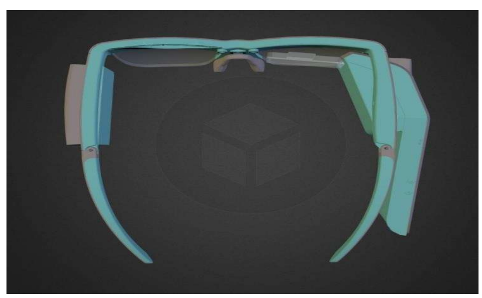

Fig. 2. Model Diagram

Software Support: Programming Language: Python – Easy, readable, and supports rapid development with a large library ecosystem.

IDE: Visual Studio Code (VS Code) – User-friendly environment for writing and executing Python code.

Speech recognition Library: Recognizes speech and converts it into text using multiple speech engines for improved accuracy.

Adafruit_SSD1306 Library: Drives the OLED display to show real-time transcribed text on the smart spect frame.

Board & Busio Libraries (Adafruit Blinka): Provide a uniform interface for accessing GPIO pins, ensuring compatibility with different microcontroller platforms.

Overall Design: Enables seamless integration between hardware and software for real-time, hands-free speech-to-text conversion, improving communication for hearing-impaired users.

Audio Capture and Processing: The system begins by capturing live audio from the surrounding environment using an inbuilt high-sensitivity microphone. This audio is preprocessed for noise reduction and amplitude control before being sent to the speech recognition module.

Speech Recognition and Conversion: The preprocessed audio signal is passed to the AI-based speech recognition engine running on a connected host device or the ESP32 (if offline models are used). The module converts spoken words into textual form with minimal latency. Custom thresholds are set to handle ambient noise and varying speech rates, ensuring reliable transcription.

Text Data Processing for Display: After the conversion function has finished transforming the speech to text, the text will be prepared to improve how easy it is to read. The text will be line-wrapped, set to an appropriate font size, and be set to a time delay before it shows on the screen. This part of the process ensures that users are presented with clean and legible text, whether it is a short text message or the message exceeds typical screen space.

Display Rendering: The processed text is sent wirelessly or through serial communication to the ESP32 microcontroller. The ESP32, using the Adafruit_SSD1306 library, displays the text on the OLED display that is housed in the spectacle frame. This allows the text to appear in front of the user's eye, enabling hands-free reading.

Synchronization and Delay Management: A regulated interval is created between text refreshes to synchronize with organic reading speeds, and a flicker-free screen is established. Additionally, this delay also provides adequate time for the system to accommodate the next audio aspects, thus allowing for continuous real-time processing capabilities.

User Interaction: In the last stage of the Via's process, the speech is displayed as text on the OLED screen located in the user's field of vision. With this real-time display and user-friendly, hands-free experience, the user is able to follow conversations fluidly and respond without additional steps or devices.

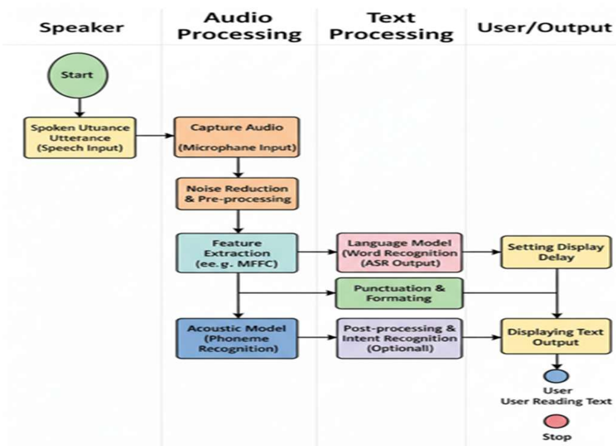

Fig. 3. System Flow

IV. Results and Discussion

A. Speech Recognition Performance

Testing was performed in three environments - a calm office, a classroom with moderate noise, and a hallway with significant noise. In the quiet office, the system achieved an average recognition accuracy of 93–95%, with clear real-time transcription using the speech_recognition library. In the classroom setting, accuracy remained above 88%, while in the hallway it dropped to around 82% due to multiple overlapping voices. The use of the MAX9814 microphone with Automatic Gain Control (AGC) helped to stabilize input levels, reducing distortion and improving recognition consistency compared with a standard microphone.

B. System Latency

The end-to-end latency — from speech capture to text appearing on the OLED display — ranged from 1.1 to 1.6 seconds in typical conditions. In the quiet office tests, latency averaged 1.2 seconds, while in the noisier classroom it increased to 1.4 seconds, primarily due to the additional processing required for noise handling and cloud API response time. Wi-Fi transmission between the Python module and the ESP32 microcontroller proved stable during all tests, with data packet losses below 0.5%.

C. Prototype Validation

The assembled DeafAssist Smart Spect weighed around 85 grams, making it comfortable for extended wear. Power consumption was measured at 0.37 W during active use, allowing several hours of continuous operation on a standard USB power bank. The OLED screen driven by the Adafruit_SSD1306 library displayed text at a frame update rate of ~15 Hz, which eliminated flicker and made scrolling text easily readable.

Figures 4–7 illustrate the physical prototype

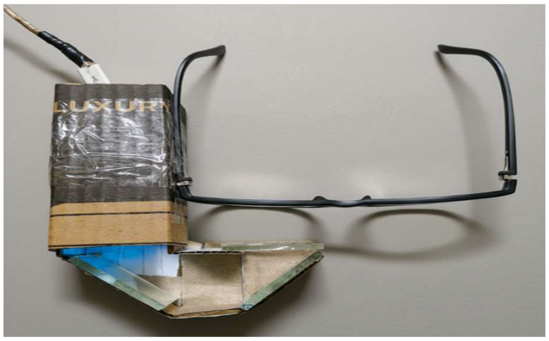

Fig. 4. Top View

Fig. 5. Front View

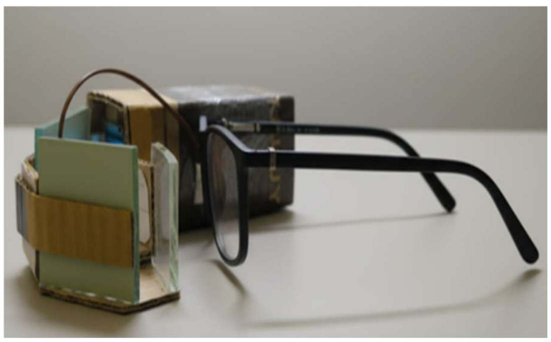

Fig. 6. Side View

Fig. 7. Text Being Displayed

V. FUTURE SCOPE

The current DeafAssist Smart Spect prototype successfully converts speech into readable text in real time. However, several enhancements can be explored to increase functionality, user convenience, and adaptability:

A. Standalone Speech Recognition

Incorporating offline speech-to-text models such as Vosk or TensorFlow Lite would allow the device to function independently of cloud services, enabling uninterrupted operation in areas without internet connectivity.

B. Support for Multiple Language

Future versions could implement multilingual translation capabilities so that users could see the speech content in their desired or regional language, allowing for a greater number of users.

C. Enhanced Portability Through Battery Integration

The inclusion of a compact rechargeable Li-ion battery with a charging module could convert the device to be fully wireless, increasing mobility and convenience.

D. Improved Accuracy in Noisy Environments

Applying digital signal processing (DSP) techniques to filter background noise can significantly enhance speech recognition reliability in challenging auditory settings such as classrooms, labs, or crowded public spaces.

E. Advance Display Technology

Incorporating transparent OLEDs or micro-AR displays could allow text to appear directly in the user's line of sight, improving readability, immersion, and overall user experience.

F. On-Device AI Processing

With advancements in microcontrollers like the ESP32-S3, future models can run speech recognition algorithms entirely on-device, reducing latency, improving privacy, and enhancing energy efficiency.

VI. CONCLUSION

The Smart Spect DeafAssist project has demonstrated a feasible and workable method for assisting individuals with hearing impairments through live speech-to-text conversion. The combination of an MAX9814 microphone and ESP32 microcontroller paired with a small OLED display enabled hands-free and intuitive operation. The use of Python programming language as well as speech_recognition and adafruit_ssd1306 libraries enabled for seamless operation between the hardware and programming elements.

Technology performed as expected in controlled environments where speech was converted to text on-screen live and accurately. Although this project now relies on wireless communication for processing, the end result has an optimistic application to everyday spaces in providing communication support within a classroom, workplace or social space. To sum up, the project has succeeded in providing accessibility, communication, and independence for individuals with hearing impairment in the form of a low-cost and working wearable device.

REFERENCES

[1] N. Jacob, S. P. Nair, R. Mathew, and D. Sunny, "SpecAssist: Smart Glasses with Real-time Speech Recognition and Transcription for the Hearing Impaired," Proc. Int. Conf. on Advanced Science, Computing and Communication (NCASCD), vol. 11, no. 04, IJERT, 2023.

[2] R. Xu, Y. Kwon, J. Lee, D. Lee, and Y. Kim, "ECHOLENS: Smart Glasses for Real-time Speech Display," IEEE Int. Conf. on Pervasive Computing and Communications Workshops (PerCom Workshops), Atlanta, USA, 2023.

[3] K. S. Tamilselvan, A. M. Sachin, C. R. Raja, M. Sowndhar, and V. Srinivas, "Assistive Glass for Hearing Impairment Using Speech Recognition," Challenges in Information, Communication and Computing Technology, vol. 2, CRC Press, 2025.

[4] S. R. Srinivasan, U. Karni, A. K. Veeraraghavan, and K. Sivaraman, "Development of Augmented Reality Based Smart Glasses for Assistance of Deaf People," Proc. Int. Conf., Nov. 2018. [Online]. Available: https://www.researchgate.net/publication/329999412

[5] D. Fabry and A. Bhowmik, "Improving Speech Understanding and Health Monitoring through Integration of AI and Sensors into Hearing Aids," IEEE Engineering in Medicine and Biology Magazine, vol. 40, no. 5, pp. 45–53, 2021.

[6] B. Sudharsan and M. Chockalingam, "A Microphone Array and Voice Algorithm-Based Smart Hearing Aid," IEEE Access, vol. 7, pp. 177200–177210, 2019.

[7] S. Alkhalifa and M. S. Al-Razgan, "ENSSAT: Deaf and Hard of Hearing Wearable Technology Application," Proc. Int. Conf. on Smart Computing and Communications (ICSCC), 2018.

[8] Y. Liu, J. Xu, and T. Chen, "Real-Time Speech Recognition on Embedded IoT Devices Using Edge AI," IEEE Internet of Things Journal, vol. 9, no. 18, pp. 17340–17350, 2022.

[9] P. Jain and M. Kumar, "Low-Power Smart Wearable Devices for Accessibility: Design and Implementation," IEEE Sensors Journal, vol. 22, no. 4, pp. 3212–3221, 2022.

[10] A. Rahman and N. Singh, "Speech-to-Text Conversion using Deep Learning for Assistive Technologies," IEEE Int. Conf. on Communication and Signal Processing (ICCSP), pp. 1345–1349, 2021.

[11] L. Gao, H. Wang, and C. Li, "Wearable Assistive Devices for Hearing-Impaired Users: A Survey," IEEE Reviews in Biomedical Engineering, vol. 15, pp. 89–104, 2022.

[12] H. Zhang, M. Park, and E. Lee, "IoT-Enabled Wearable System for Hearing Assistance Using ESP32 and Cloud Speech API," IEEE Access, vol. 10, pp. 78312–78323, 2022.

[13] S. Y. Lee and J. H. Park, "Transparent OLED Display for Smart Glasses Applications," IEEE Transactions on Electron Devices, vol. 68, no. 3, pp. 1195–1202, 2021.

[14] T. Nguyen, P. Tran, and L. Pham, "Design of Low-Latency Speech Recognition Systems for Real-Time Applications," IEEE Transactions on Consumer Electronics, vol. 69, no. 1, pp. 13–22, 2023.

[15] A. K. Sharma and P. S. Bhatt, "IoT-Based Smart Glasses for Accessibility and Real-Time Communication," IEEE Int. Conf. on Intelligent Systems and Networks (ICISN), pp. 542–547, 2024.