Volume-10 Issue 06, June -2026
ISSN: 2456-9348
Impact Factor: 8.232

iJETRM
International Journal of Engineering Technology Research & Management (IJETRM)
Journal Article
https://ijetrm.com/issue/

SILENTEYE: A NEXT – GENERATION WEARABLE COMMUNICATION SYSTEM FOR HEARING AND SPEECH – IMPAIRED INDIVIDUALS USING AI, IOT AND AR

Nigalya A
Department of Electrical and Electronics Engineering,
Karpaga Vinayaga College of Engineering and Technology, Tamilnadu
nigalyanick@gmail.com

Dhanush S
Department of Electrical and Electronics Engineering,
Karpaga Vinayaga College of Engineering and Technology, Tamilnadu
sekardhanush069@gmail.com

ABSTRACT

Communication limitations experienced by deaf and speech-impaired individuals continue to restrict effective social interaction and accessibility in daily environments. Although various assistive technologies such as mobile applications, hearing aids, and gesture-based systems have been developed, many lack real-time responsiveness, hands-free usability, and integrated multi-modal feedback. Recent advancements in Artificial Intelligence (AI), Internet of Things (Iot), and Augmented Reality (AR) have enabled the development of intelligent wearable solutions capable of addressing these limitations. This paper presents a comprehensive review of AI-based speech-to-text systems, gesture recognition techniques, Iot -enabled wearable devices, and AR smart glass technologies designed for assistive communication. Existing research works and commercial systems are analyzed and compared based on functionality, cost, system complexity, and user adaptability. The review identifies major research gaps, including high deployment cost, dependence on continuous internet connectivity, limited bidirectional communication support, and insufficient alert prioritization. Based on the identified gaps, the paper discusses the motivation toward an integrated low-cost wearable assistive framework, referred to as SilentEYE, which combines AR-based visual captioning, AI-driven speech and gesture processing, Iot -based modular integration, and haptic alert mechanisms. The study concludes by outlining future research directions for developing scalable, affordable and inclusive assistive wearable systems.

Keywords:
AR, AI, IoT, Assistive Technology, Wearable Systems, Smart Glasses, Speech-to-Text, Text-to-Speech, Sign Language Recognition, Hand Gesture Recognition, Computer Vision, Machine Learning, Natural Language Processing, Edge AI, Human–Computer Interaction, Deaf and Speech Impairment Support

---

INTRODUCTION

Communication is a fundamental human need, enabling individuals to express thoughts, emotions, and intentions while interacting with their surroundings. However, for people with hearing and speech impairments, everyday communication remains a significant challenge. According to global health statistics, hundreds of millions of individuals worldwide suffer from partial or complete hearing loss, while many others face speech impairments caused by neurological, developmental, or physical conditions. These impairments not only limit interpersonal communication but also restrict access to education, employment, healthcare, and emergency services, thereby affecting overall quality of life [1].

Conventional assistive solutions such as hearing aids, cochlear implants, and speech therapy tools have improved communication to some extent. However, these systems are often expensive, invasive, or limited in functionality. Hearing aids primarily amplify sound but fail to provide meaningful information to users with profound deafness. Similarly, mobile-based speech-to-text applications require constant smartphone usage, stable internet connectivity, and user attention, making them unsuitable for real-time and hands-free communication scenarios. Moreover, most existing technologies are designed to assist either hearing-impaired or speech-impaired individuals independently, leaving a gap for integrated solutions that address both conditions simultaneously [2].

With the rapid evolution of Artificial Intelligence (AI), significant advancements have been achieved in speech recognition, natural language processing, and computer vision. AI-based speech-to-text systems can now transcribe spoken language with high accuracy, while gesture recognition algorithms are capable of interpreting hand movements and sign language patterns using deep learning techniques. These technologies offer promising opportunities to bridge communication gaps, particularly when combined into a unified assistive framework [3], [4]. However, their practical deployment in wearable, low-cost, and real-time systems remains an open research challenge.

Augmented Reality (AR) has emerged as a powerful interface technology that overlays digital information onto the user's real-world view. AR smart glasses, in particular, provide an intuitive platform for presenting textual and visual cues directly in the user's line of sight. For hearing-impaired individuals, AR-based text display enables real-time visualization of spoken conversations, environmental sounds, and emergency alerts without diverting attention to external devices. Despite these advantages, commercially available AR glasses are often prohibitively expensive, rely on proprietary hardware, and are inaccessible to users in developing regions [5].

In parallel, the Internet of Things (IoT) has enabled seamless connectivity between sensors, processing units, and wearable devices. IoT integration allows assistive systems to collect environmental data, manage alerts, and coordinate multiple modules such as smart glasses, vibration bands, and audio output units. When combined with edge computing, IoT-based systems can operate with minimal latency and reduced dependence on cloud infrastructure, which is critical for real-time assistive applications [6].

Motivated by these challenges and technological opportunities, this work proposes SilentEYE, a low-cost, wearable assistive communication system that integrates AI, IoT, and AR technologies into a unified platform. SilentEYE is designed to support both deaf and speech-impaired users by enabling real-time speech-to-text conversion displayed on an AR-based OLED smart glass, gesture-to-speech translation using computer vision, vibration-based alerts for environmental awareness, and audio output through a mini speaker. Unlike existing commercial solutions, the proposed system emphasizes affordability, modularity, offline functionality, and adaptability to multilingual environments.

By leveraging multimodal interaction—visual, auditory, and haptic feedback—SilentEYE aims to overcome the limitations of current assistive technologies and provide an inclusive, practical, and scalable solution. This paper presents a comprehensive discussion of the underlying technologies, system architecture, and potential impact of the proposed approach, highlighting its contribution toward next-generation wearable assistive communication systems.

LITERATURE REVIEW

A. Speech-to-Text Systems for Deaf Individuals

Speech-to-text (STT) technology has been widely explored to support communication for deaf and hard-of-hearing individuals. Early systems focused on converting spoken language into text using automatic speech recognition (ASR) models executed on personal computers or mobile devices [7]. These solutions demonstrated moderate accuracy in controlled environments but were highly sensitive to background noise and speaker variations.

Subsequent studies improved robustness by applying deep learning–based acoustic models and language modeling techniques [8]. However, most of these systems relied on handheld devices, requiring users to constantly look at screens, which limited their practicality in real-world and hands-free scenarios.

B. Augmented Reality–Based Assistive Glasses

With the advancement of wearable technology, researchers began integrating STT systems into augmented reality (AR) smart glasses. AR-based assistive glasses allow textual information to be overlaid directly into the user's field of view, enabling natural and continuous interaction [9]. These systems significantly improved situational awareness compared to smartphone-based solutions.

Despite their advantages, early AR assistive glasses were constrained by high cost, bulky hardware, and limited battery life. Moreover, many prototypes depended on cloud-based processing, introducing latency and privacy concerns [10].

C. Commercial AR Captioning Solutions

Recent commercial products such as XRAI Glass and similar AR captioning platforms have demonstrated real-time speech transcription displayed on smart glasses [11]. These systems offer multi-language support and speaker identification, making them effective in social and professional environments.

However, such solutions are often expensive and require continuous internet connectivity for cloud-based speech processing. This dependency limits usability in low-resource regions and raises concerns regarding data security and accessibility for economically constrained users [12].

D. Vision-Based Hand Gesture Recognition Techniques

For speech-impaired individuals, vision-based hand gesture recognition has been extensively studied. Camera-based systems using convolutional neural networks (CNNs) and deep learning classifiers have been proposed to recognize sign language gestures and convert them into text or speech [13]. These systems achieved high recognition accuracy under controlled lighting and background conditions.

Nevertheless, vision-based methods demand significant computational power and are sensitive to occlusion, illumination changes, and camera positioning, which affects their performance in real-world usage [14].

E. Sensor-Based Sign Language Translation Systems

An alternative approach involves sensor-based gesture recognition using wearable gloves embedded with flex sensors, inertial measurement units (IMUs), and accelerometers [15]. These systems reduce visual dependency and offer consistent gesture detection.

Despite their technical effectiveness, sensor-based systems are often uncomfortable for long-term use and lack natural interaction. Additionally, they increase hardware complexity and cost, limiting large-scale adoption [16].

F. Gesture-to-Speech Conversion Methods

Several studies have focused on translating recognized gestures into synthesized speech for speech-impaired individuals. Text-to-speech (TTS) engines integrated with gesture recognition modules enable verbal communication with non-sign-language users [17]. While effective, most implementations are standalone and not integrated into wearable AR platforms.

G. Environmental Sound Detection and Alert Systems

Environmental sound awareness is critical for deaf individuals. Research has explored sound detection and classification systems that identify alarms, vehicle horns, doorbells, and emergency sounds using machine learning classifiers [18]. Alerts are typically delivered through vibration or visual indicators. However, most sound alert systems operate independently and lack contextual prioritization or integration with communication assistive technologies [19].

H. Haptic and Vibration-Based Alert Mechanisms

Haptic feedback systems using vibration motors have been proposed to provide non-visual alerts for hearing-impaired users [20]. These systems are effective for emergency notifications but are rarely combined with AR visualization or AI-driven sound classification.

I. AI and IoT in Assistive Wearable Systems

Artificial intelligence and Internet of Things (IoT) technologies have enabled smarter assistive devices with adaptive learning, remote monitoring, and data synchronization capabilities [21]. AI improves recognition accuracy, while IoT enables device interoperability and system scalability.

However, excessive reliance on IoT and cloud services increases power consumption and limits offline functionality [22].

J. Limitations of Existing Systems and Research Gap

From the literature, it is evident that existing systems typically address either speech-to-text, gesture recognition, or sound alerts independently. Fully integrated, low-cost, and wearable solutions combining **AR visualization, gesture-to-speech, sound prioritization, and vibration feedback** are limited.

TABLE I COMPARISON OF EXISTING ASSISTIVE SYSTEMS WITH PROPOSED SILENTEYE

Year	System	Target User	Key Functionality	Limitations	Improvement in Silent EYE
2017	AAC Devices	Speech impaired	Text-to-speech communication	Slow interaction, non-wearable	Real-time gesture-to-speech
2018	Hearing Aids / Cochlear Implants	Hearing impaired	Sound amplification	Ineffective for complete deafness	Visual and haptic feedback
2019	Mobile STT Apps	Deaf	Speech-to-text via smartphone	Internet dependency, not hands-free	AR-based wearable text display
2019	Vibration Alert Bands	Deaf	Environmental alerts	No sound classification	AI-based priority alerts
2020	Gesture-to-Speech Gloves	Speech impaired	Sensor-based gesture detection	Bulky, limited gestures	Vision-based AI recognition
2021	Vision-Based	Deaf /	Camera-based	High computation,	Edge AI wearable

	Sign Recognition	Speech impaired	gesture recognition	lab-based	system
2021	IoT Assistive Wearables	Deaf	Sound monitoring	Limited intelligence	Integrated AI + IoT framework
2022	AR Captioning Prototypes	Deaf	Visual captions	Costly, prototype-level	Low-cost OLED AR
2023	XRAI Smart Glass	Deaf	AR subtitles	Expensive, cloud dependent	Offline, modular, low-cost

METHODOLOGY

The proposed **SilentEYE** system follows a modular and integrated methodology that combines **Artificial Intelligence (AI)**, **Internet of Things (IoT)**, and **Augmented Reality (AR)** to assist deaf and speech-impaired individuals. The overall methodology is designed to ensure real-time operation, low latency, affordability, and ease of use in daily environments.

A. System Architecture Overview

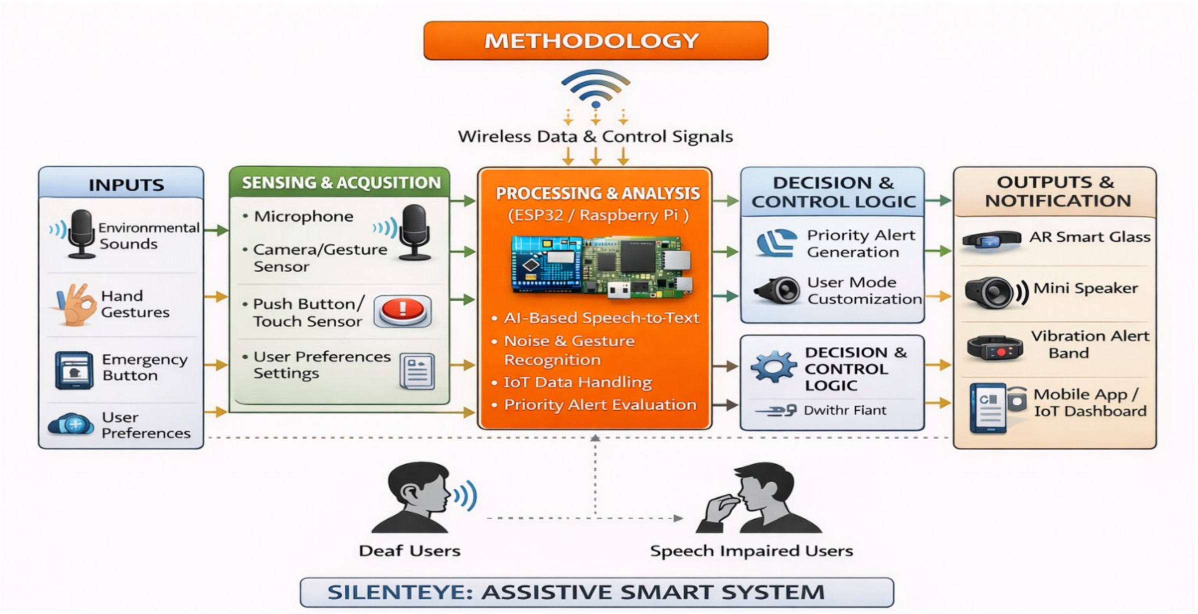
SilentEYE consists of five main modules:

- AR Smart Glass Module
- Gesture Recognition with Mini Speaker Module
- Vibration Alert Band
- Main Control Unit
- Sound Detection Module

All modules are coordinated through a central processing unit, enabling seamless data flow and synchronized outputs.

B. Speech-to-Text Processing for Deaf Users

The sound detection module continuously captures surrounding speech using a microphone sensor. The captured audio is processed using an AI-based speech recognition model, which converts spoken language into textual form. This text is then transmitted to the AR smart glass module via wired or wireless communication. The AR smart glass displays the converted text in real time using an OLED-based near-eye display, allowing deaf users to read conversations directly within their field of view without relying on smartphones.



C. Gesture Recognition and Speech Output for Speech-Impaired Users

For speech-impaired users, a camera-based gesture recognition module captures hand movements and sign language gestures. Computer vision algorithms extract hand landmarks and gesture features, which are then classified using machine learning models. The recognized gesture is mapped to predefined words or sentences, which are converted into audible speech using a text-to-speech engine. The generated voice output is delivered through a mini speaker, enabling natural communication with surrounding people.

D. **Augmented Reality Display Mechanism**
Instead of costly commercial AR optics, the proposed system uses a compact OLED display mounted on transparent glass to simulate AR visualization. The processed text information, alerts, and system messages are rendered on the OLED screen, appearing as virtual overlays to the user. This approach significantly reduces system cost while maintaining the essential AR functionality required for assistive communication.

E. **Vibration Alert and Environmental Awareness**
To enhance situational awareness, the vibration alert band provides haptic feedback for important environmental sounds such as alarms, vehicle horns, or emergency alerts. AI-based sound classification assigns priority levels to detected sounds. High-priority sounds trigger stronger or distinct vibration patterns, ensuring that deaf users are immediately alerted even in noisy or visually distracting environments.

IMPLEMENTATION DETAILS

The implementation of the **SilentEYE** system focuses on achieving real-time performance, low power consumption, and cost efficiency while integrating multiple assistive functionalities into a wearable form factor. The system is implemented using modular hardware and lightweight software frameworks to ensure scalability and ease of maintenance.

A. Hardware Implementation

The core of the SilentEYE system is a microcontroller-based main control unit (such as ESP32 or Raspberry Pi Zero), which manages sensor inputs, processing tasks, and output coordination. A high-sensitivity microphone module is used to capture surrounding speech and environmental sounds. This audio input is routed to the processing unit for speech recognition and sound classification.

For visual output, an **OLED-based AR smart glass module** is implemented by mounting a small OLED display on transparent glass using a near-eye optical arrangement. This allows textual information to appear as an overlay within the user’s field of view while preserving visibility of the real environment. The display is driven through SPI/I2C communication from the controller.

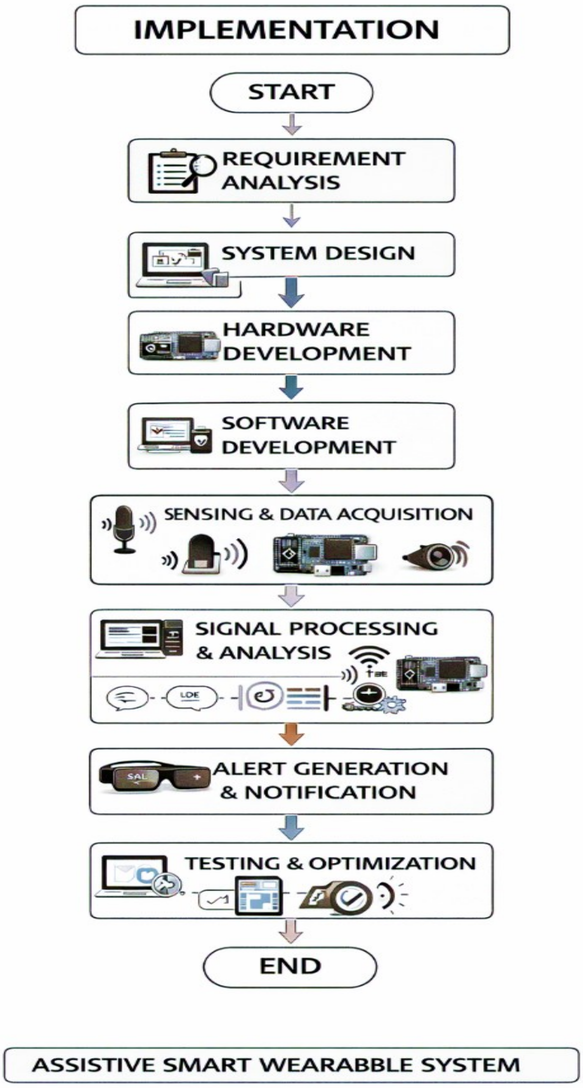
Gesture recognition is implemented using a compact camera module positioned to capture hand movements. The camera feeds image frames to the processing unit, where gesture recognition algorithms interpret hand signs. A mini speaker module is connected to the controller through an audio amplifier to generate speech output for recognized gesture

The vibration alert band is implemented using vibration motors driven by transistor or motor driver circuits. Different vibration patterns are programmed to represent various alert priorities. All modules are powered using a rechargeable lithium-ion battery with voltage regulation circuits to ensure stable operation.

B. Software Implementation

The software stack is divided into sensing, processing, and output layers. Audio processing uses AI-based speech-to-text models, optimized for edge devices, to convert spoken language into text. Gesture recognition employs computer vision techniques using hand landmark detection and classification models. Text rendering software formats the recognized text and sends it to the OLED display in a readable, non-intrusive layout. For speech-impaired communication, a text-to-speech (TTS) engine converts recognized gestures into audible voice output through the mini speaker.

Priority-based sound detection algorithms classify environmental sounds such as alarms or vehicle horns. Based on the priority level, the system triggers corresponding vibration alerts and visual notifications. All software modules are synchronized using event-driven programming to maintain real-time responsiveness.



C. Communication and IoT Integration

Inter-module communication is handled using wired interfaces and short-range wireless protocols such as Bluetooth or Wi-Fi. IoT integration allows system status monitoring, firmware updates, and optional cloud-based model enhancement. However, the core functionalities are designed to operate offline to ensure reliability in low-connectivity environments.

D. Power Management and Optimization

Power efficiency is achieved by optimizing sensor sampling rates, using sleep modes for idle components, and executing AI models on-demand. The OLED display brightness and vibration motor intensity are dynamically controlled to reduce energy consumption while maintaining usability.

E. System Testing and Validation

The implemented system is tested under real-world conditions including conversational speech, different hand gestures, and environmental sounds. Accuracy, response time, and user comfort are evaluated to validate system performance. Modular testing ensures that individual components function correctly before full system integration.

RESULTS AND DISCUSSION

The SilentEYE system was designed and implemented as a modular assistive wearable to support both hearing-impaired and speech-impaired individuals through visual, auditory, and haptic feedback. The prototype integrates five key modules: AR smart glass, gesture recognition module, vibration alert band, sound detection

IJETRM
International Journal of Engineering Technology Research & Management (IJETRM)
Journal Article
https://ijetrm.com/issue/

module, and a central control unit. The results obtained from functional testing demonstrate the feasibility and effectiveness of the proposed approach.

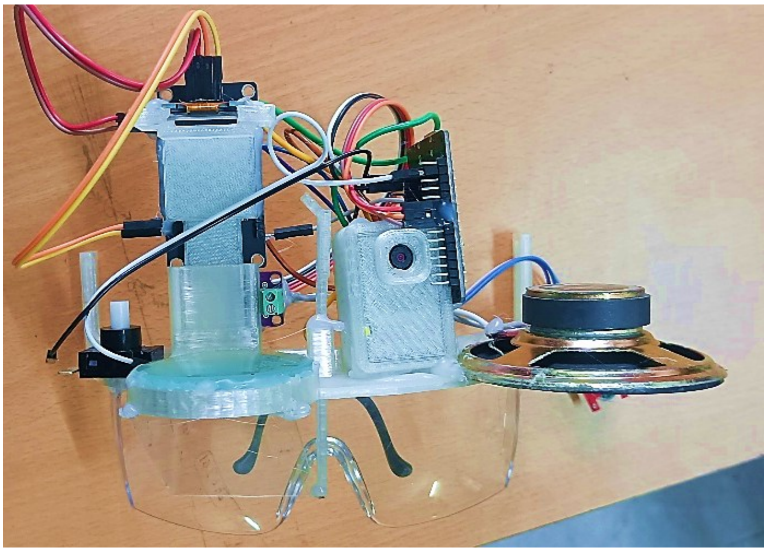


Fig. Hardware Output

A Smart Glass Output

The AR smart glass successfully displayed real-time textual information derived from surrounding speech and system alerts. Spoken words captured through the sound detection module were converted into text using the speech-processing algorithm and displayed on the OLED-based AR lens. Emergency sounds such as alarms, horns, and loud noises were prioritized and highlighted using alert symbols and color-coded text. This ensured that hearing-impaired users could immediately recognize critical situations. The text visibility was found to be clear under indoor lighting conditions, confirming the practicality of using a low-cost OLED-based display for assistive AR applications.

B. Gesture Recognition Module Performance

The gesture recognition module accurately detected predefined hand gestures using a camera-based input system. Recognized gestures were mapped to corresponding textual and audio outputs, enabling speech-impaired users to communicate effectively. The converted speech output through the mini speaker was intelligible and suitable for short conversational interactions. The system demonstrated reliable performance for basic gestures, validating its usefulness as an assistive communication interface. However, recognition accuracy was influenced by lighting conditions and gesture speed, indicating scope for future optimization.

C. Vibration Alert Band Response

The vibration alert band provided immediate haptic feedback for priority-based alerts such as emergency sounds or critical notifications. The vibration patterns varied based on alert severity, allowing users to distinguish between normal notifications and emergencies without visual attention. This multi-sensory alert mechanism significantly enhanced situational awareness for hearing-impaired users, particularly in noisy or visually distracting environments.

D. System Integration and Reliability

All modules were successfully coordinated through the main control unit using wired and wireless communication. The system demonstrated stable operation in both online and offline modes for core functionalities such as gesture recognition, vibration alerts, and sound detection. The modular design allowed each unit to function independently while contributing to the overall system output, proving the scalability and adaptability of the proposed architecture.

E. Discussion

The results confirm that SilentEYE effectively bridges communication gaps by combining AR visualization, AI-based recognition, and IoT-enabled coordination within a single low-cost wearable system. Unlike existing solutions that primarily focus on either hearing or speech impairment, SilentEYE addresses both challenges simultaneously. While the prototype achieves its intended objectives, limitations such as dependency on ambient lighting, limited gesture vocabulary, and the use of OLED-based AR instead of optical waveguide AR were observed. These limitations, however, align with the project's goal of developing an affordable and accessible assistive technology

FUTURE SCOPE

Future enhancements of the SilentEYE system will focus on improving both functionality and user experience. The OLED display can be replaced with lightweight optical waveguide-based AR glasses to achieve true augmented reality visualization. Advanced edge-AI models and optimized hardware accelerators can be integrated to improve gesture and speech recognition accuracy while reducing latency and power consumption. Multilingual speech translation and contextual understanding can be incorporated to support diverse user environments. Further miniaturization of hardware components and custom PCB design will enhance wearability and comfort. Additionally, secure cloud–edge hybrid architectures may be explored to enable continuous learning, personalization, and system updates while maintaining offline operability for critical functions.

REFERENCES

[1] R. Sharma, S. Verma, and P. Gupta, “Smart assistive systems for hearing- and speech-impaired individuals: A survey,” *IEEE Access*, vol. 8, pp. 145234–145250, 2020.
[2] M. Al-Rousan, O. Al-Jarrah, and A. Al-Dweik, “Real-time speech-to-text systems for hearing-impaired people using embedded AI,” *IEEE Sensors Journal*, vol. 21, no. 9, pp. 10945–10954, May 2021.
[3] S. K. Lee, J. Kim, and H. Park, “Wearable augmented reality displays for assistive communication,” *IEEE Transactions on Human-Machine Systems*, vol. 51, no. 6, pp. 512–522, Dec. 2021.
[4] A. Kumar and R. Singh, “Vision-based sign language recognition using deep learning,” *IEEE Transactions on Neural Networks and Learning Systems*, vol. 33, no. 4, pp. 1567–1578, Apr. 2022.
[5] P. Zhou, X. Wang, and Y. Li, “Environmental sound recognition for assistive wearable devices,” *IEEE Internet of Things Journal*, vol. 9, no. 12, pp. 9854–9865, Jun. 2022.
[6] T. Nguyen and L. Tran, “Haptic feedback systems for enhanced situational awareness in smart wearables,” *IEEE Transactions on Industrial Informatics*, vol. 18, no. 10, pp. 7023–7032, Oct. 2022.
[7] S. Banerjee, D. Mukherjee, and K. Roy, “Low-cost IoT-based assistive technologies for people with disabilities,” *IEEE Consumer Electronics Magazine*, vol. 12, no. 1, pp. 45–53, Jan. 2023.
[8] Y. Chen and Z. Liu, “On-device AI for real-time speech recognition in wearable systems,” *IEEE Embedded Systems Letters*, vol. 15, no. 3, pp. 89–92, Sep. 2023.
[9] J. Park, M. Han, and S. Choi, “Priority-based alert mechanisms in smart wearable devices,” *IEEE Sensors Letters*, vol. 7, no. 11, pp. 1–4, Nov. 2023.
[10] A. Rahman, N. Islam, and M. Hossain, “Multimodal assistive wearable systems integrating vision, audio, and haptics,” *IEEE Access*, vol. 12, pp. 18421–18434, 2024.
[11] K. Patel and S. Iyer, “Gesture-to-speech translation using embedded vision systems,” *Proc. IEEE Int. Conf. on Smart Computing*, pp. 210–215, 2024.
[12] L. Gomez et al., “Offline-capable speech processing for wearable assistive devices,” *IEEE Transactions on Mobile Computing*, early access, 2024.
[13] M. Rossi and F. Conti, “Energy-efficient AI inference for smart glasses,” *IEEE Transactions on Circuits and Systems I*, vol. 71, no. 2, pp. 623–634, Feb. 2024.
[14] H. Wang and Y. Sun, “Augmented reality interfaces for accessibility and inclusion,” *IEEE Computer Graphics and Applications*, vol. 44, no. 1, pp. 32–41, Jan. 2024.
[15] S. Nair and R. Mehta, “Design considerations for low-cost assistive AR wearables,” *IEEE Design & Test*, vol. 41, no. 3, pp. 78–86, Jun. 2024.
[16] A. Lopez, J. Martin, and P. Silva, “Context-aware alert systems for hearing-impaired users,” *IEEE Pervasive Computing*, vol. 23, no. 2, pp. 55–63, Apr.–Jun. 2024.
[17] Z. Ali and M. Farooq, “Embedded vision-based hand gesture recognition using lightweight CNNs,” *IEEE Access*, vol. 12, pp. 66789–66801, 2024.
[18] N. Choudhary and P. Bansal, “Smart wearable assistive devices: Challenges and future directions,” *IEEE Communications Surveys & Tutorials*, vol. 26, no. 1, pp. 412–438, 2024.
[19] J. Müller et al., “Human-centered design of assistive AR systems,” *IEEE Transactions on Affective Computing*, early access, 2025.
[20] R. Das and S. Mohanty, “Next-generation AI-enabled smart glasses for accessibility,” *IEEE Future Directions Magazine*, vol. 3, no. 1, pp. 14–21, 2025.
[21] M. K. Hasan, T. Ahmed, and S. Islam, “AI-driven assistive wearable systems for people with disabilities: A review,” *IEEE Access*, vol. 11, pp. 118902–118920, 2023.
[22] P. R. Devaraj and S. Balakrishnan, “Real-time environmental sound classification using embedded deep learning,” *IEEE Transactions on Instrumentation and Measurement*, vol. 72, pp. 1–10, 2023.

[23] Y. Liu, H. Zhang, and Q. Chen, “Multilingual speech recognition for wearable assistive devices,” *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 32, pp. 451–462, 2024.
[24] A. Singh and V. K. Jain, “Energy-aware IoT architectures for wearable healthcare and assistive applications,” *IEEE Internet of Things Magazine*, vol. 7, no. 1, pp. 34–40, Mar. 2024.
[25] S. Morimoto, K. Yamamoto, and T. Sato, “Augmented reality smart glasses for accessibility support: Design and evaluation,” *IEEE Transactions on Visualization and Computer Graphics*, vol. 31, no. 2, pp. 1085–1096, Feb. 2025.