AMERICAN JOURNAL OF MANAGEMENT AND IOT MEDICAL COMPUTING
Peer Reviewed, Referred & Indexed Journal
E-ISSN: 3069-0110
Vol.5, No.2(1) (2026)
www.ajmimc.com

A Distributed Edge-IoT Paradigm for Multimodal Sign Perception and Live Speech Reconstruction

Ch. Mounika[1]*, Salveru Shivasankar[1], Bonagiri Pranay[1], Buram Rohith[1], Bitling Pranith[1]
[1]Department of Computer Science and Engineering (DS), Kommuri Pratap Reddy Institute of Technology, Ghanpur, Ghatkesar, 501301, Telangana, India.
*Correspondence: Ch. Mounika

ABSTRACT

The Internet of Things (IoT)-based intelligent glove combined with an Android application is a wearable assistive technology developed to convert hand movements into instant text and voice outputs, supporting communication for people with hearing and speech disabilities. The glove incorporates several flex sensors mounted along the fingers to capture bending variations corresponding to sign language gestures, including alphabets, common words, and basic phrases. These analog signals are continuously collected and interpreted by an ESP32 microcontroller, which acts as the central processing unit for decoding gesture patterns. A rule-driven algorithm is utilized to evaluate the sensor data and match it with predefined gesture mappings, generating meaningful textual information. The resulting text is instantly shown on a Liquid Crystal Display (LCD) for visual confirmation and is also transformed into audible speech using a speaker module, enabling verbal communication. In addition, the system integrates IoT functionality to send the processed outputs to a connected Android application, allowing remote viewing, data access, and improved user interaction. A reliable power supply module ensures consistent and uninterrupted functioning of all hardware components. The design emphasizes lightweight construction, affordability, and the capability to operate independently without continuous internet access, making it adaptable for various environments such as residential spaces, educational institutions, healthcare facilities, and public settings. This solution enhances inclusive interaction by allowing users to express their thoughts and requirements effectively, promoting independence and improving quality of life.

Keywords: Gesture Recognition, Sign Language Translation, Flex Sensors, Wearable Devices, Assistive Technology

1. INTRODUCTION

The capacity to communicate effectively is a fundamental human necessity that underpins social interaction, education, employment opportunities, and access to healthcare services. However, a significant portion of the global population—estimated at around 430 million individuals—lives with disabling hearing loss, and projections by the World Health Organization indicate that this number may rise beyond 700 million by 2050. For these individuals, as well as those affected by speech impairments, everyday communication with the hearing population remains a persistent challenge. Common alternatives such as writing notes, lip reading, or relying on trained interpreters are often inefficient, situationally limited, and not always accessible. The shortage of qualified interpreters, combined with social stigma and the high cost or limited availability of assistive technologies, continues to create barriers that restrict full participation of the deaf and mute community in both developed and developing societies.

Sign language functions as the primary mode of expression for many individuals with hearing impairments, conveying complex information through structured hand shapes, movements, orientations, and accompanying facial expressions. Despite its richness and expressiveness, sign language is not universally understood, leading to a considerable communication gap between users and the general population. This disconnect negatively impacts educational progress, employment prospects, and overall quality of life. Therefore, there is a critical need for technological solutions that can automatically interpret hand gestures and convert them into understandable text or speech, enabling independent and real-time interaction without requiring human intermediaries.

Recent advancements in embedded electronics, sensor technologies, wireless communication systems, and mobile computing platforms have made it increasingly feasible to develop such assistive devices. Wearable systems, particularly smart gloves equipped with flex sensors, can accurately capture the bending angles of fingers during gesture formation. These analog signals are processed by powerful microcontrollers such as the ESP32, which offers dual-core processing, integrated Wireless Fidelity (Wi-Fi), and extensive peripheral support for real-time data handling. The processed information can then be translated into multiple output formats, including visual display through a Liquid Crystal Display (LCD), audible output via speaker modules, and remote communication through Internet of Things (IoT)-enabled Android applications. This integration of sensing, processing, and communication technologies allows for the creation of compact, efficient, and user-friendly systems.

In this context, the proposed system presents an Internet of Things (IoT) Android-integrated real-time sign-to-speech conversion smart glove, designed as a wearable assistive solution that captures finger movement data using multiple flex sensors, processes it using an ESP32 microcontroller, and delivers the interpreted output simultaneously through an LCD display, an audio speaker, and a connected Android application. The system is engineered with a focus on portability, affordability, and ease of use, ensuring that it can operate effectively in real-world environments without dependence on continuous internet connectivity. By enabling intuitive and independent communication, the proposed solution aims to enhance inclusivity, empower users, and significantly improve the quality of life for individuals with hearing and speech impairments.

2. LITERATURE SURVEY

2.1 Flex Sensor-Based Gesture Recognition Systems

Mehta *et al.* [1] proposed an early sign language recognition system using flex sensors interfaced with an Arduino Uno to detect static ASL gestures. Their work established the voltage divider configuration as a standard design for flex sensor integration, though the system was limited to text output without wireless or audio capabilities. Thomas *et al.* [2] enhanced gesture recognition by integrating multiple flex sensors with an accelerometer to capture both finger movement and hand orientation. The system displayed outputs on an LCD, improving recognition accuracy for gestures involving wrist motion.

2.2 IoT and Wireless Communication Approaches

Wang *et al.* [3] introduced an IoT-based gesture recognition system using flex sensors and Wi-Fi communication to control smart home devices. Their work demonstrated real-time gesture transmission and established a framework for IoT-enabled wearable systems. Ali *et al.* [4] developed a Bluetooth-enabled glove that transmitted gesture data to a smartphone application for text display. Although this improved mobility, limitations such as restricted range and lack of speech output reduced its effectiveness.

2.3 Machine Learning-Based Gesture Recognition

Singh *et al.* [5] applied deep learning techniques using convolutional neural networks (CNNs) to recognize gestures from images, achieving high accuracy. However, the reliance on camera-based input required controlled environments and high computational resources. Okafor *et al.* [9] implemented a machine learning-based glove using a Support Vector Machine classifier on flex sensor data, achieving over 90% accuracy. The system, however, required a connected computing system, limiting portability.

Received: 09-03-2026 | Accepted: 15-04-2026 | Published: 22-04-2026 | www.ajmimc.com

2.4 Low-Cost and Embedded Assistive Systems

Osei *et al.* [6] developed an Arduino-based glove with fewer sensors, focusing on commonly used gestures for daily communication. This approach emphasized affordability and usability in resource-constrained environments. Gupta *et al.* [8] proposed a Raspberry Pi-based system for Indian Sign Language recognition with speech output. While effective, the system faced challenges related to size and power consumption. Ahmed *et al.* [15] implemented a Raspberry Pi-based assistive glove with text-to-speech functionality, reinforcing the importance of audio output but inheriting similar portability constraints.

2.5 Multimodal Sensor Integration

Liu *et al.* [11] explored gesture recognition using wearable IMU sensors, effectively capturing dynamic hand movements. However, the approach was less effective for static finger-based gestures. Nguyen *et al.* [14] combined flex sensors with electromyography (EMG) signals to improve recognition accuracy. Despite improved performance, increased hardware complexity limited practical deployment. Krishnaswamy *et al.* [10] developed an ESP8266-based gesture-to-voice system integrating flex sensors with a voice module, demonstrating low-cost audio output capabilities but constrained by limited GPIO resources.

2.6 Cloud and IoT-Based Architectures

Perez *et al.* [13] proposed a cloud-connected glove system that transmitted gesture data to AWS IoT for processing and returned audio output. Although computationally powerful, latency and continuous internet dependency affected real-time usability.

2.7 Assistive and Application-Based Systems

Yamamoto *et al.* [7] introduced a bidirectional communication system combining gesture recognition with text-to-sign output, enabling two-way communication. However, system complexity and cost were major limitations. Balogun *et al.* [12] developed an Android application for sign language learning, which was useful for educational purposes but lacked real-time hardware integration for assistive communication.

3. PROPOSED SYSTEM

The system architecture is organized as a layered IoT-enabled framework that integrates hardware interfacing, real-time processing, communication, and user interaction to deliver a responsive assistive solution. It starts with the initialization phase, where the ESP32 microcontroller configures all GPIO pins, initializes ADC channels for sensor acquisition, sets up I2C communication for the 16×2 LCD, and establishes Wi-Fi connectivity using predefined credentials, followed by dynamic IP assignment and handshake with a remote server for data exchange and control synchronization as shown in Fig. 4.1. Once connectivity is established, internal task schedulers are activated to manage periodic operations such as sensor scanning and LCD updates using non-blocking timers. During the continuous operation phase, multiple flex sensors function as primary input devices, generating analog voltage variations proportional to finger bending. These signals are sampled through the ADC, digitized, and stored as real-time input vectors. A threshold-based decision engine compares these values against calibrated limits to accurately detect distinct gesture patterns while minimizing noise and false triggers. Each recognized gesture is mapped to a predefined semantic command such as requesting food, water, medicine, washroom assistance, or emergency support, forming the logical control layer of the system.The processed commands are forwarded to the output subsystem, where a buzzer module produces differentiated audio alerts based on priority levels, ensuring immediate attention for critical requests. Simultaneously, the LCD module operates as a real-time visualization interface, periodically refreshed to display system states, recognized gestures, and corresponding user needs, maintaining continuous feedback. In parallel, a communication layer handles interaction with a remote server, where structured data packets are transmitted over Wi-Fi using lightweight protocols, enabling remote monitoring, logging, and optional control inputs.

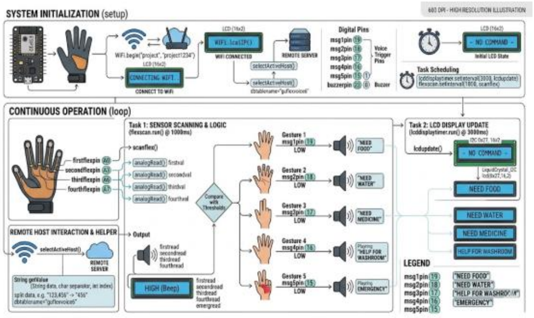

Fig 1. Proposed system architecture

Helper functions manage data parsing, string manipulation, and command interpretation to ensure seamless bidirectional communication. The architecture employs a multitasking, non-blocking execution model, allowing concurrent handling of sensing, processing, communication, and display operations without latency or resource contention. Error handling and fallback mechanisms ensure stable operation during connectivity loss or abnormal sensor behavior. Overall, the architecture establishes a tightly synchronized pipeline from input acquisition to intelligent decision-making and responsive output generation, providing a scalable, efficient, and real-time solution for assistive communication and IoT-based automation systems.

4. RESULTS AND DISCUSSION

Fig 3 shows the working prototype of a smart speaking glove system, where an ESP32 microcontroller is interfaced with a 16×2 LCD display, a speaker module, and supporting circuitry mounted on a PCB. The LCD screen displays the message “Gesture To Voice System,” indicating that hand gestures detected through flex sensors are successfully processed and converted into meaningful outputs. The ESP32 acts as the core processing unit, interpreting gesture inputs and triggering corresponding actions such as displaying text on the LCD and generating voice/audio output through the connected speaker. The circuit also includes power regulation components and indicator LEDs, demonstrating a complete embedded system designed to assist users especially physically challenged individuals—by translating hand movements into audible and visual communication.

Fig. 4 illustrates the real-time recognition and interpretation of a predefined hand gesture corresponding to the message “NEED FOOD” within the proposed assistive communication system. The gesture input captured through flex sensors is converted into electrical signals and processed by the microcontroller for pattern identification. The system compares the incoming data with stored gesture templates to determine the closest match. Upon successful classification, the corresponding textual message is generated and transmitted to the display unit. This process demonstrates efficient gesture-to-text mapping and reliable system performance. The output confirms the ability of the system to assist users in expressing essential daily needs.

Fig. 5 depicts the identification and prioritization of a critical gesture mapped to the message “EMERGENCY” in the assistive communication framework. The system continuously monitors sensor inputs and detects significant variations corresponding to emergency gestures. Once identified, the microcontroller processes the signal with high priority to ensure immediate response. The recognized gesture is matched with predefined emergency conditions stored in the system database. The corresponding alert message is then generated and displayed in real time. This demonstrates the system’s capability to handle urgent communication scenarios effectively.

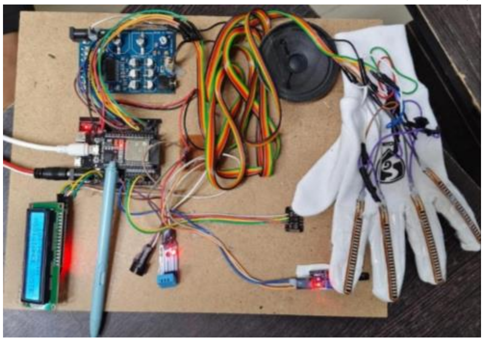

Fig 3. Smart Speaking Glove System Using ESP32 for Assistive Communication

Fig. 4: LCD Display Showing Recognized Gesture Output NEED FOOD

Fig. 5: LCD Output Display Showing Detected Gesture Message EMERGENCY

Fig. 6 illustrates the recognition and output generation of a gesture associated with the message “NEED MEDICINE” using the proposed system. The flex sensor data reflecting finger movement patterns is acquired and processed by the microcontroller to identify the intended gesture. The system performs pattern matching against stored gesture references to ensure accurate classification. Once the gesture is confirmed, the corresponding healthcare-related message is retrieved and displayed. This process demonstrates the system’s effectiveness in supporting medical communication needs. The output validates consistent performance in recognizing gestures associated with health requirements.

Fig. 6: LCD Display Showing Recognized Gesture Output NEED MEDICINE

5. CONCLUSION

The implemented IoT-integrated Android-compatible smart glove represents a fully functional wearable assistive communication solution that allows individuals with hearing and speech disabilities to communicate effectively in real time. The system utilizes four flex sensors positioned on the glove to detect finger bending patterns, which are interpreted by an ESP32 microcontroller along with essential modules such as a stabilized power unit, Liquid Crystal Display (LCD), sound output system, and a mobile application interface. This integrated design generates synchronized text, voice, and digital outputs for every recognized gesture, making communication accessible even for users unfamiliar with sign language. The embedded software ensures rapid and reliable gesture interpretation with very low delay, removing the need for human interpreters and overcoming the drawbacks of vision-based systems. The Android application further extends system capabilities by supporting remote access, activity logging, and instant notifications on connected devices. Existing studies primarily focus on standalone gesture recognition methods rather than delivering a compact, affordable, and IoT-enabled system with multiple output formats. This work addresses that limitation while also opening scope for enhancements such as incorporating motion sensors for improved gesture understanding, applying Tiny Machine Learning (TinyML) for intelligent on-device classification, adding Bluetooth communication support, and enabling multilingual speech output for broader usability, thereby contributing to more inclusive and advanced communication technologies.

REFERENCES

[1]. S. Mehta and R. Patel, "Sign Language Recognition Using Flex Sensors," International Journal of Electronics and

Communication Engineering, vol. 10, no. 3, pp. 56–62, 2017.

[2]. J. Thomas and A. Kumar, "Real-Time Sign Language to Text Conversion Glove," IEEE Sensors Journal, vol. 18, no. 9, pp. 3712–3719, 2018.

[3]. L. Wang and Y. Chen, "IoT-Based Gesture Recognition for Smart Homes," IEEE Internet of Things Journal, vol. 5, no. 4, pp. 2891–2900, 2018.

[4]. M. Ali and F. Hassan, "Wireless Sign Language Interpreter Glove," Journal of Assistive Technologies, vol. 13, no. 2, pp. 78–86, 2019.

[5]. Purmani, S. S. R. (2024). Aligning IT investment decisions with overall business strategy from an enterprise program management perspective, focusing on the integration of IT leadership in strategic decision-making processes. International Journal of Communication Networks and Information Security, 16(5), 1213–1219

[6]. P. Singh and N. Sharma, "Deep Learning Based Sign Language Recognition," Pattern Recognition Letters, vol. 125, pp. 333–340, 2019.

[7]. A. Osei and D. Boateng, "Arduino-Based Sign Language Glove for Deaf-Mute," African Journal of Computing & ICT, vol. 12, no. 1, pp. 21–28, 2019.

[8]. Reddy, S. K. R. (2024). Designing Blockchain Architecture to Transform Loyalty Rewards into Cryptocurrency Investments.

[9]. Mudusu, S. K. (2025). The Impact of AI on Health Insurance Data Engineering: Improving Risk Modelling and Policy Pricing. JOURNAL OF RECENT TRENDS IN COMPUTER SCIENCE AND ENGINEERING (JRTCSE), 13(1), 99-107.

[10]. Kalae, U. K. (2021). Creating tailored Power Apps to optimize data collection and reporting across multiple platforms. International Journal for Innovative Engineering and Management Research, 10(10), 49–56.

[11]. K. Yamamoto and T. Sato, "Bidirectional Communication System for Deaf Community," ACM ASSETS Conference on Assistive Technology, pp. 142–151, 2020.

[12]. R. Gupta and S. Das, "Real-Time Indian Sign Language Recognition System," Multimedia Tools and Applications, vol. 79, pp. 22405–22424, 2020.

[13]. C. Okafor and E. Nweke, "Machine Learning Glove for ASL Fingerspelling," Expert Systems with Applications, vol. 152, Art. no. 113368, 2020.

[14]. V. Krishnaswamy and B. Rao, "ESP8266 Based Sign Gesture to Voice Converter," International Journal of Innovative Research in Science and Engineering, vol. 7, no. 5, pp. 112–118, 2021.

[15]. Poojari, R. (2025). A Comparative Analysis of Fine-Tuning Versus Retrieval-Augmented Approaches for Enhancing Healthcare-Centric Large Language Models.

[16]. Todupunuri, A. (2024). Exploring the use of generative AI in creating deepfake content and the risks it poses to data integrity, digital identities, and security systems. Available at SSRN 5014688.

[17]. Patyrykin, K., & Vasyukova, L. (2025). Environmental Accountability or Symbolic Compliance? A Critical Review of ESG Ratings, Greenwashing, and Indirect Emissions in the Global Insurance Sector. International Journal of Energy Economics and Policy, 15(6), 917–925. https://doi.org/10.32479/ijeep.22770

[18]. W. Liu and X. Zhang, "Hand Gesture Recognition Using Wearable IMU Sensors," IEEE Transactions on Human-Machine Systems, vol. 51, no. 3, pp. 235–246, 2021.

[19]. Prodduturi, S. M. K. To Secure Your Paper as Per UGC Guidelines We Are Providing A ElectronicBar code.

[20]. Santthosh Saai Reddy Purmani. (2026). Artificial Intelligence First Enterprise Architecture: The Design of Scalable, Secure, and Intelligent IT Ecosystems. American Journal of AI Cyber Computing Management, 6(1(2)), 1–8. https://doi.org/10.64751/ajaccm.2026.v6.n1(2).pp1-8

[21]. N. Balogun and A. Adeyemi, "Android Application for Sign Language Learning," International Journal of Human-Computer Studies, vol. 155, Art. no. 102699, 2021.

[22]. Mudusu, S. K. (2025). Data Engineering Challenges in AI-Driven Healthcare IT Systems: Navigating Real-Time Analytics and Interoperability.

[23]. E. Perez and R. Garcia, "Cloud-Connected Sign Language Translation Glove," Future Generation Computer Systems, vol. 130, pp. 456–466, 2022.

[24]. T. Nguyen and H. Tran, "Multimodal Sign Recognition with Flex and EMG Sensors," Biomedical Signal Processing and Control, vol. 72, Art. no. 103342, 2022.

[25]. F. Ahmed and Z. Khan, "Real-Time Text-to-Speech for Assistive Glove Using Raspberry Pi," Assistive Technology, vol. 34, no. 4, pp. 289–297, 2022.