

## Edge-Enabled IoT Multimodal Wearable System for Real-Time Sign Language to Speech Conversion with Mobile Integration

Rithesh Kumar<sup>1\*</sup>, Mohammad Abrar Ul haq<sup>2</sup>, Jayya Akash<sup>2</sup>, M. Charan Theja Reddy<sup>2</sup>

<sup>1</sup>Assistant Professor, <sup>1,2</sup>UG Student, <sup>1,2</sup>Department of Computer Science and Engineering.

<sup>1,2</sup>Kommuri Pratap Reddy Institute of Technology, Ghanpur, Ghatkesar, 501301, Telangana, India.

\*Correspondence: Rithesh Kumar ([riteshkumar237@gmail.com](mailto:riteshkumar237@gmail.com))

### ABSTRACT

This research is a wearable assistive system designed to translate hand gestures into real-time text and speech, enabling effective communication for individuals with hearing and speech impairments. The glove is embedded with multiple flex sensors placed along the fingers to detect variations in bending patterns associated with sign language gestures, including letters, words, and simple expressions. These sensor readings are continuously acquired and processed by an ESP32 microcontroller, which serves as the core unit for signal interpretation. A rule-based approach is employed to analyze the incoming data and map it to predefined gesture patterns, which are then converted into meaningful textual outputs. The interpreted text is immediately displayed on an LCD for visual feedback and simultaneously converted into audio through a speaker module to facilitate verbal interaction. Furthermore, the system incorporates IoT capabilities to transmit the generated outputs to a connected Android application, allowing remote access, monitoring, and enhanced usability. A stable power supply unit supports uninterrupted operation of the system components. The design prioritizes portability, cost-effectiveness, and the ability to function independently without requiring constant internet connectivity, making it suitable for diverse environments such as homes, schools, hospitals, and public areas. Overall, the system promotes accessible and inclusive communication by enabling users to convey their needs efficiently and independently.

**Key words:** Gesture Recognition, Sign Language Translation, Flex Sensors, Wearable Devices, Assistive Technology

### 1. INTRODUCTION

Communication is a fundamental human right and a cornerstone of social participation, education, employment, and healthcare access. However, for the estimated 430 million people worldwide who experience disabling hearing loss — a figure projected to rise to over 700 million by 2050 according to the World Health Organization — meaningful communication with the hearing majority remains a daily challenge. Speech-impaired individuals face similar barriers, often relying on written notes, lip reading, or the presence of a trained sign language interpreter to interact in everyday situations. The scarcity of trained interpreters, the social stigma associated with disability, and the lack of affordable assistive technology continue to marginalize the deaf and mute community across both developed and developing countries.

Sign language is the natural and primary language of the deaf community, conveying meaning through hand shapes, orientations, movements, and facial expressions. While sign language is rich and expressive, it is not universally understood. The vast majority of the hearing population has no knowledge of sign language, creating a profound communication barrier that affects the quality of life, educational outcomes, and economic opportunities of those who rely on it. Assistive technologies that can automatically interpret sign language gestures and convert them into spoken or written output

have the potential to dramatically reduce this barrier and empower individuals to communicate independently without requiring a human interpreter.

Recent advances in microelectronics, sensor technology, wireless communication, and mobile computing have created unprecedented opportunities to develop such assistive devices. Wearable gloves equipped with flex sensors can accurately capture the angular bending of each finger during sign formation. Microcontrollers like the ESP32, with their dual-core processing capability, integrated Wi-Fi, and extensive peripheral support, can process this sensor data in real time and drive multiple output modalities including audio speakers, LCD displays, and IoT-connected mobile applications. The convergence of these technologies makes it possible to design a compact, affordable, and highly functional sign-to-speech conversion device that can be worn by the user and operated autonomously in real-world environments.

This research proposes and implements the IoT Android Integrated Real Time Sign to Speech Conversion Smart Glove, a wearable assistive device that captures finger-bend data from four flex sensors, processes it on an ESP32 microcontroller, and delivers the recognized sign language output simultaneously through an LCD display, an audio speaker, and an Android mobile application connected via IoT. The system is designed for portability, affordability, and ease of use, targeting the needs of the deaf and mute community across diverse socioeconomic and geographic contexts.

## 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.

## 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 proposed IoT Android Integrated Real Time Sign to Speech Conversion Smart Glove directly addresses the limitations of existing systems by delivering a fully integrated, multi-modal, and IoT-connected assistive communication device built around the ESP32 microcontroller. Four flex sensors mounted on the glove continuously monitor the bending of the index, middle, ring, and little fingers, generating analog voltage signals that are sampled by the ESP32's ADC channels at regular intervals. The ESP32 firmware processes the four-channel sensor data using a classification algorithm that maps specific combinations of bending angles to predefined sign language gestures, covering alphabets, numbers, and high-frequency words and phrases. Upon successful gesture recognition, the system simultaneously triggers three output modalities: the recognized text is displayed on the LCD for visual confirmation; the audio speaker plays the corresponding pre-recorded spoken word for acoustic output accessible to hearing individuals; and the gesture label is transmitted via Wi-Fi to the paired Android APK, which displays it as a live text feed with notification alerts and maintains a session log for review. The regulated power supply ensures stable, noise-free operation across all components. The entire system is self-contained, operates without internet or cloud dependency, and can be powered by a rechargeable battery pack worn on the wrist.

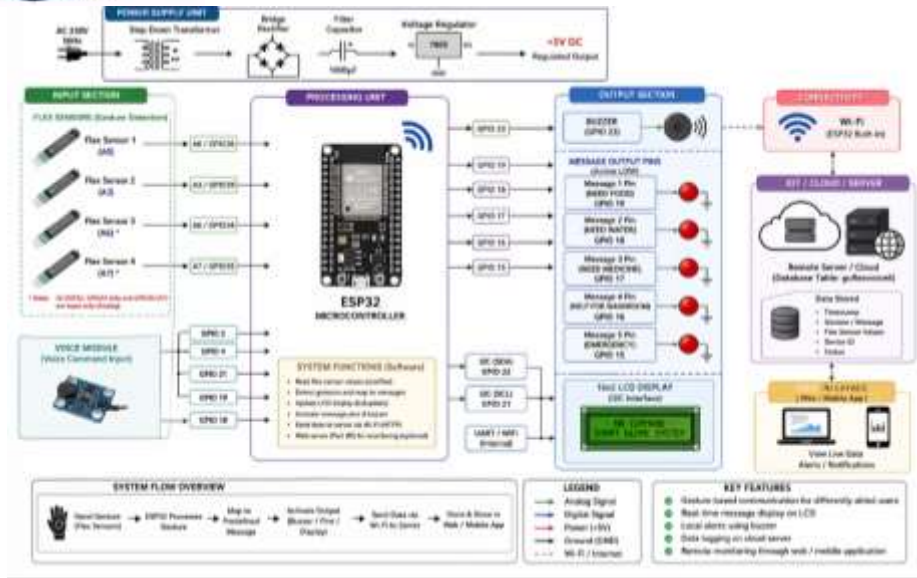


Fig 1. Proposed system architecture of smart speaking gloves

The use of ESP32 as the core platform, combined with low-cost flex sensors and an open-source Android application, makes this solution significantly more affordable and accessible than commercial assistive communication devices, positioning it as a viable and impactful tool for the deaf and mute community in both urban and rural settings

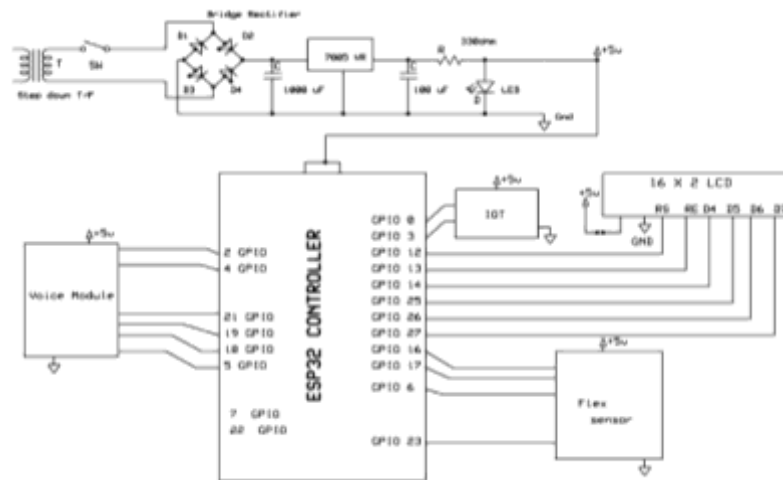


Fig 2. ESP 32 based circuit diagram

As shown in fig 2 the circuit diagram represents an ESP32-based embedded system powered through a regulated supply, where an AC input is stepped down using a transformer, rectified by a bridge rectifier (D1–D4), filtered with capacitors, and regulated to a stable +5V using a 7805 voltage regulator, with an LED indicator for power status. The ESP32 microcontroller acts as the central unit, interfacing with multiple components: a voice module connected through GPIO pins for voice command input, an IoT module for wireless communication, a 16×2 LCD display (connected via control and data pins such as RS, E, D4–D7) for output visualization, and a flex sensor for sensing bending or motion input. All components are powered by the regulated +5V supply, and the ESP32 processes inputs from the sensors and voice module, communicates data via IoT, and displays relevant information on the LCD, forming a complete smart embedded system.

#### 4. RESULTS

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.

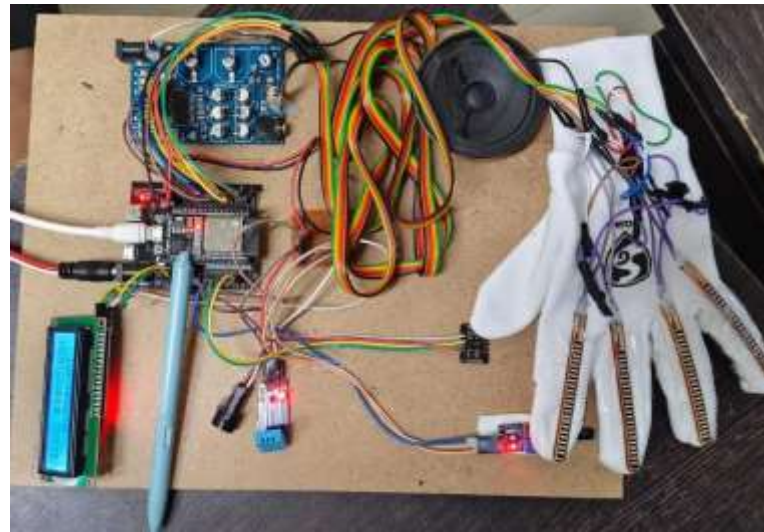


Fig 3. Smart Speaking Glove System Using ESP32 for Assistive Communication

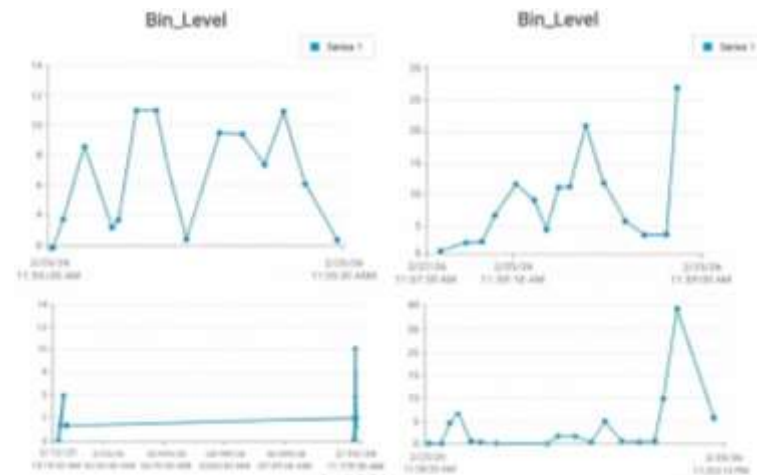


Fig 4. Real-Time Sensor Data Analysis for Smart Speaking Glove System

Fig 4 presents multiple time-series graphs labeled “Bin\_Level,” representing real-time sensor data variations captured from the smart speaking glove system, likely corresponding to flex sensor readings associated with different finger movements or gesture patterns. Each graph illustrates how sensor values change over time, reflecting dynamic hand gestures performed by the user. Peaks in the graphs indicate significant finger bending or gesture activation, which are interpreted by the ESP32 microcontroller to trigger predefined messages such as “NEED FOOD” or “EMERGENCY.” The

variation across different plots suggests multiple sensor channels or gesture scenarios being monitored simultaneously. This data visualization plays a crucial role in analyzing gesture recognition accuracy, system responsiveness, and threshold calibration, ultimately helping improve the reliability and efficiency of the assistive communication system.

## 5. CONCLUSION

The developed IoT-connected Android-supported smart glove demonstrates a complete wearable assistive communication system that enables individuals with hearing and speech impairments to interact with others in real time. The system integrates four flex sensors embedded within a glove to capture finger movements, which are processed by an ESP32 microcontroller along with supporting components such as a regulated power supply, LCD display, audio output module, and a mobile application interface. This coordinated architecture produces simultaneous text, speech, and digital outputs for each detected gesture, ensuring that communication is accessible to users regardless of their familiarity with sign language. The embedded firmware provides fast and consistent gesture recognition with minimal latency, eliminating dependence on interpreters and avoiding the limitations of camera-based systems. The connected Android application enhances functionality by enabling remote monitoring, session tracking, and real-time notifications on paired devices. A review of existing research indicates that most approaches focus on isolated gesture recognition techniques rather than delivering a unified, portable, and cost-effective IoT-enabled system with multi-modal outputs. This work bridges that gap while also establishing a foundation for future improvements, including the integration of motion sensing through accelerometers for better gesture context, adoption of TinyML (Tiny Machine Learning) models for on-device intelligent classification and expanded vocabulary, incorporation of Bluetooth as an alternative communication protocol, and support for multilingual speech generation to accommodate diverse user communities, thereby advancing inclusive and technology-driven communication solutions.

## 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]. P. Singh and N. Sharma, "Deep Learning Based Sign Language Recognition," *Pattern Recognition Letters*, vol. 125, pp. 333–340, 2019.
- [6]. 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.
- [7]. K. Yamamoto and T. Sato, "Bidirectional Communication System for Deaf Community," *ACM ASSETS Conference on Assistive Technology*, pp. 142–151, 2020.
- [8]. R. Gupta and S. Das, "Real-Time Indian Sign Language Recognition System," *Multimedia Tools and Applications*, vol. 79, pp. 22405–22424, 2020.
- [9]. C. Okafor and E. Nweke, "Machine Learning Glove for ASL Fingerspelling," *Expert Systems with Applications*, vol. 152, Art. no. 113368, 2020.

- [10]. 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.
- [11]. 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.
- [12]. N. Balogun and A. Adeyemi, "Android Application for Sign Language Learning," *International Journal of Human-Computer Studies*, vol. 155, Art. no. 102699, 2021.
- [13]. E. Perez and R. Garcia, "Cloud-Connected Sign Language Translation Glove," *Future Generation Computer Systems*, vol. 130, pp. 456–466, 2022.
- [14]. T. Nguyen and H. Tran, "Multimodal Sign Recognition with Flex and EMG Sensors," *Biomedical Signal Processing and Control*, vol. 72, Art. no. 103342, 2022.
- [15]. 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.