IJARSCT

International Journal of Advanced Research in Science, Communication and Technology

International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal

ISSN: 2581-9429

Volume 5, Issue 11, May 2025

Impact Factor: 7.67

Bridging Silence: A Real-Time Gesture-to-Voice Translator Using ESP32 and Flex Sensors

Mr. Shrikant Wadkar[1], Mr. Saurabh Hekare[2], Mr. Shivraj Mane[3], Prof. Mrs. Pallavi Gholap[4]

Student, AI&DS Department, Jaihind College of Engineering Kuran, Pune, India [123]

Professor, AI&DS Department, Jaihind College of Engineering Kuran, Pune, India [4]

Abstract: *Communication is a fundamental human need, yet individuals who are deaf and mute often face significant barriers in expressing themselves to the broader society. Traditional methods like sign language require both parties to be proficient, and interpreters are not always available or affordable. To address this challenge, this research presents a wearable Gesture-to-Speech System that translates hand gestures into audible speech using an ESP32 microcontroller, flex sensors, and a speaker module. The system is designed to be lightweight, cost-effective, and user-friendly, aiming to empower non-verbal individuals with a tool for real-time communication.*

The device operates by detecting specific hand gestures through flex sensors attached to a glove. These sensors measure the bending of fingers, and the data is processed by the ESP32 microcontroller to identify predefined gestures. Upon recognition, the system generates corresponding speech output via a speaker and displays the text on an optional screen for visual confirmation. The integration of these components ensures seamless translation from gesture to speech, facilitating more inclusive interactions. This paper delves into the system's architecture, detailing the hardware and software components, and discusses the methodology employed in developing and testing the prototype. A comprehensive literature review highlights existing technologies and their limitations, establishing the novelty and necessity of the proposed system. The results demonstrate the device's effectiveness in accurately recognizing gestures and delivering prompt speech output, indicating its potential to significantly enhance the quality of life for deaf and mute individuals..

Keywords: Gesture Recognition, ESP32, Smart Glove, Flex Sensors, Speech Conversion, Assistive Technology

I. INTRODUCTION

Effective communication is essential for social interaction, education, and employment. However, individuals who are deaf and mute often encounter obstacles due to the reliance on sign language, which is not universally understood. This communication gap can lead to social isolation and limited opportunities.

Advancements in wearable technology and microcontrollers offer new avenues to bridge this gap. By translating hand gestures into speech, it is possible to facilitate real-time communication between non-verbal individuals and the broader community. This research focuses on developing a **Gesture-to-Speech System** utilizing the ESP32 microcontroller, known for its processing capabilities and wireless communication features, to create an accessible and efficient communication aid.

II. PROBLEM STATEMENT

Deaf and mute individuals face communication barriers due to limited understanding of sign language by the general public and lack of accessible interpreters. There is a critical need for a low-cost, real-time gesture-to-speech system that enables independent and effective communication.

III.OBJECTIVES

To develop a wearable device that translates hand gestures into audible speech in real-time.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429 Volume 5, Issue 11, May 2025 Impact Factor: 7.67

To ensure the system is cost-effective, portable, and user-friendly.
To utilize the ESP32 microcontroller for efficient processing and potential wireless communication.
To enhance the quality of life for deaf and mute individuals by facilitating seamless communication.

IV. LITERATURE REVIEW

chen, J., Zhao, Z., Chen, K., Zhang, S., Zhou, Y., & Deng, W. (2020). *Wearable-tech glove translates sign language into speech in real time*.

Chen and colleagues developed a lightweight, stretchable glove equipped with sensors capable of translating American Sign Language (ASL) into English speech in real-time via a smartphone application. The glove utilizes conductive yarns to detect finger movements, which are then processed and converted into speech at a rate of approximately one word per second. This innovation aims to facilitate direct communication between sign language users and non-signers without the need for human interpreters. The system's affordability and portability make it a promising tool for enhancing accessibility for the deaf and hard-of-hearing community.[1]

Bodda, S. C., Gupta, P., Joshi, G., & Chaturvedi, A. (2020). *A new architecture for hand-worn Sign language to Speech translator*.

Bodda and co-researchers proposed a modular smart glove architecture designed to translate ASL gestures into spoken English. The glove integrates flex sensors, accelerometers, and gyroscopes to capture finger orientations and hand motions. By employing decision tree algorithms for gesture recognition and error correction, the system addresses hardware-dependent issues found in existing designs. The modular approach allows for distributed processing, reducing complexity and facilitating future enhancements. This research contributes to the advancement of sensor-based sign language translation technologies.[2]

Kalandar, B., & Dworakowski, Z. (2023). *Sign Language Conversation Interpretation Using Wearable Sensors and Machine Learning*.

In their study, Kalandar and Dworakowski introduced a proof-of-concept automatic sign language recognition system utilizing a wearable device with three flex sensors. The system interprets dynamic ASL words by collecting sequential gesture data and applying machine learning algorithms, including Random Forest and Support Vector Machine (SVM). Achieving up to 99% accuracy, the research highlights the potential for developing full-scale systems that can significantly improve communication for individuals with hearing impairments.[3]

Nagarale, D. P., Sangale, S. B., Rukade, A. J., Wadd, D. R., & Halunde, S. S. (2024). *IoT Based Sign to Speech Converter System*.

Nagarale and team presented an IoT-based Sign-to-Speech Converter System comprising a sensor-embedded glove and an Android application. The glove captures intricate hand movements associated with sign language, transmitting data wirelessly to a central processing unit. The system interprets gestures and generates corresponding spoken language output, enhancing user experience through real-time translation. The integration of IoT technology and mobile applications underscores the system's adaptability and potential for widespread adoption in facilitating communication for the deaf community.[4]

Ambar, R., Fai, C. K., Wahab, M. H. A., Jamil, M. M. A., & Ma'radzi, A. A. (2018). *Development of a Wearable Device for Sign Language Recognition*.

Ambar and colleagues focused on developing a wearable device capable of translating sign language into speech and text. The glove-based system incorporates five flex sensors to detect finger bending and an accelerometer to monitor arm motions. By combining sensor data, the device identifies specific gestures corresponding to words and phrases in ASL, subsequently converting them into audible speech and displaying text on an LCD screen. This research emphasizes the importance of hardware design in creating effective assistive communication tools for individuals with speech and hearing impairments.[5]

V. METHODOLOGY

Component Selection: Chose ESP32 microcontroller for its processing power and wireless capabilities; selected flex sensors for gesture detection.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 11, May 2025
Impact Factor: 7.67

Prototype Development: Assembled the glove with integrated sensors and connected it to the ESP32, speaker, and display modules.
Software Implementation: Programmed the microcontroller to interpret sensor data, map gestures to specific words, and generate corresponding speech output.
Testing and Calibration: Conducted trials to fine-tune sensor thresholds and ensure accurate gesture recognition across different users.
User Feedback: Gathered input from potential users to assess comfort, usability, and effectiveness, leading to iterative improvements.

VI. SYSTEM ARCHITECTURE

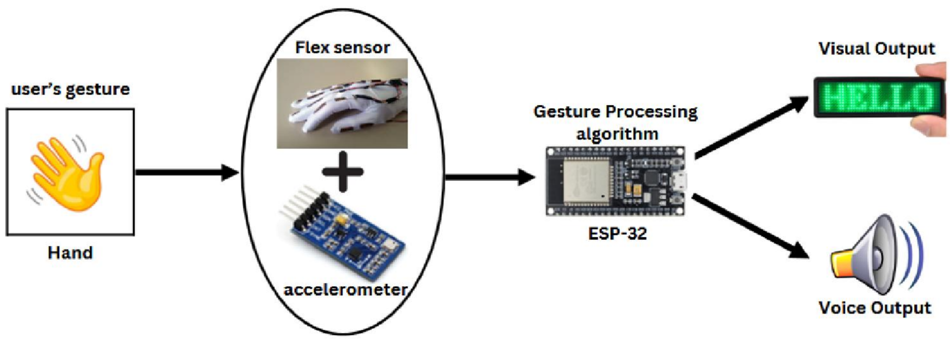

Figure 1: Architecture diagram

Hardware Components

- **Flex Sensors**: Attached to each finger of the glove, these sensors detect the degree of bending, translating physical movements into electrical signals.
- **ESP32 Microcontroller**: Serves as the central processing unit, interpreting sensor data and controlling output modules.
- **Speaker Module**: Outputs synthesized speech corresponding to recognized gestures.
- **OLED Display**: Provides visual feedback by displaying the translated text.
- **Power Supply**: A rechargeable battery powers the entire system, ensuring mobility.

Software Components

- **Gesture Mapping Algorithm**: Processes input from flex sensors to identify specific gestures based on predefined thresholds.
- **Speech Synthesis Module**: Converts recognized gestures into audible speech using text-to-speech libraries compatible with ESP32.
- **Display Interface**: Manages the output of translated text on the OLED screen.

Data Flow

- **Gesture Input**: User performs a hand gesture wearing the glove.
- **Sensor Detection**: Flex sensors capture finger movements and send analog signals to the ESP32.
- **Data Processing**: The microcontroller digitizes the signals, compares them against stored gesture patterns, and identifies the corresponding word or phrase.
- **Output Generation**: The system activates the speaker to vocalize the identified word and updates the display with the corresponding text.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 11, May 2025
Impact Factor: 7.67

VII. RESULT

The prototype developed demonstrates the following outcomes:

- **Gesture Recognition Accuracy**: The system accurately identifies predefined hand gestures corresponding to specific words or phrases.
- **Real-Time Processing**: The ESP32 microcontroller processes sensor data swiftly, ensuring minimal latency between gesture input and speech output.
- **User-Friendly Interface**: The glove-based design is comfortable and intuitive, requiring minimal training.
- **Portability**: The compact and lightweight design allows for easy transportation and use in various settings.

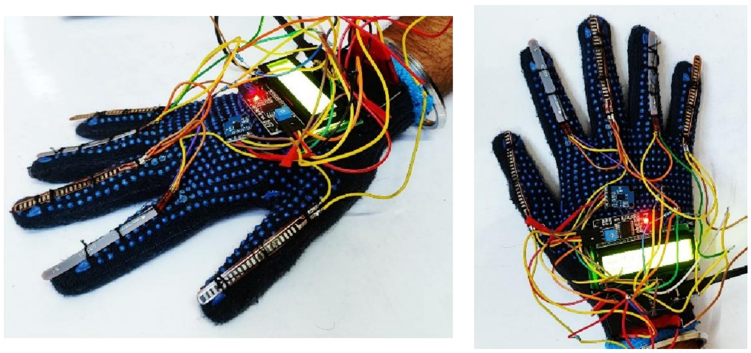

Figure 2: Smart Hand Glove

VII. BENEFITS TO SOCIETY

The Gesture-to-Speech System has the potential to bring about a significant positive transformation in society by enabling inclusive communication for individuals who are deaf and mute. One of the most impactful benefits is the promotion of independence, as the device allows users to express themselves clearly without relying on interpreters or requiring the listener to understand sign language. This reduces the communication gap between non-verbal individuals and the broader population, fostering mutual understanding and empathy. Furthermore, the affordability and portability of the system ensure that it can be made accessible even in economically disadvantaged or rural areas where advanced medical or educational facilities may be limited. In educational settings, this technology empowers students who face speech or hearing impairments to participate actively in classroom discussions and peer interactions, thereby improving their academic performance and confidence. In the workplace, it opens doors to new job opportunities by facilitating smoother communication with colleagues and supervisors, promoting equality in professional environments. The device also contributes to social integration by allowing users to engage in community activities, public services, and day-to-day interactions with ease. Moreover, by reducing dependency on others and providing a dignified mode of communication, it enhances the overall quality of life and mental well-being of the users. As society becomes more inclusive and aware of the challenges faced by people with disabilities, innovations like this system play a crucial role in building a more compassionate, accessible, and technologically empowered world for everyone.

VIII. CONCLUSION

The Gesture-to-Speech System effectively bridges the communication gap for deaf and mute individuals by translating hand gestures into speech. Utilizing the ESP32 microcontroller enhances processing efficiency and offers potential for future wireless features. The device's affordability, portability, and user-friendly design make it a viable solution for real-world application, promoting inclusivity and independence for non-verbal individuals.

IX. FUTURE SCOPE

The Gesture-to-Speech System presented in this research holds immense potential for further development and expansion. One of the most promising directions is the integration of machine learning algorithms to enable dynamic gesture learning, allowing the system to adapt to individual user styles and recognize a broader range of gestures beyond the predefined set. Additionally, incorporating multilingual speech synthesis will significantly increase the usability of the system in diverse linguistic regions, helping users communicate in their preferred language. Leveraging the wireless capabilities of the ESP32 microcontroller, the system can be extended to communicate with smartphones or cloud-based applications for remote monitoring, customization, and storage of frequently used phrases. Future versions could also include miniaturized components and soft-flexible circuits to enhance comfort and make the glove less intrusive. Moreover, the addition of haptic feedback or voice command responses could make the interaction more intuitive. These enhancements will not only increase the device's functionality but also make it more inclusive, personalized, and suitable for widespread real-world deployment in education, healthcare, and public services.

REFERENCES

[1]. Chen, J., Zhao, Z., Chen, K., Zhang, S., Zhou, Y., & Deng, W. (2020). *Wearable-tech glove translates sign language into speech in real time*. UCLA Newsroom. Link

[2]. Bodda, S. C., Gupta, P., Joshi, G., & Chaturvedi, A. (2020). *A new architecture for hand-worn Sign language to Speech translator*. arXiv preprint arXiv:2009.03988. Link

[3]. Kalandar, B., & Dworakowski, Z. (2023). *Sign Language Conversation Interpretation Using Wearable Sensors and Machine Learning*. arXiv preprint arXiv:2312.11903. Link

[4]. Nagarale, D. P., Sangale, S. B., Rukade, A. J., Wadd, D. R., & Halunde, S. S. (2024). *IoT Based Sign to Speech Converter System*. International Journal for Research in Applied Science and Engineering Technology (IJRASET), 12(4). Link

[5]. Ambar, R., Fai, C. K., Wahab, M. H. A., Jamil, M. M. A., & Ma'radzi, A. A. (2018). *Development of a Wearable Device for Sign Language Recognition*. Journal of Physics: Conference Series, 1019(1), 012017. Link

[6]. Xia, K., Lu, W., Fan, H., & Zhao, Q. (2022). *A Sign Language Recognition System Applied to Deaf-Mute Medical Consultation*. Sensors, 22(23), 9107. Link

[7]. Keny, D., Karmakar, M., Kamtekar, N., Mehta, A., & Kumar, M. (2017). *Hand Gesture Recognition and Voice Conversion System for Speech Impaired*. International Journal of Engineering Research & Technology (IJERT), 5(01). Link

[8]. Camgoz, N. C., Koller, O., Hadfield, S., & Bowden, R. (2020). *Sign Language Transformers: Joint End-to-end Sign Language Recognition and Translation*. arXiv preprint arXiv:2003.13830. Link

[9]. Wang, S., Alexanderson, S., Gustafson, J., Beskow, J., Henter, G. E., & Székely, É. (2021). *Integrated Speech and Gesture Synthesis*. arXiv preprint arXiv:2108.11436. Link

[10]. Hernandez-Rebollar, J. (2003). *Glove Won't Speak for the Deaf*. Wired. Link