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 6, Issue 3, February 2026
Impact Factor: 8.2

Gesture Sense: Translating Sign Language into Text and Speech

Dr. Vilas S. Ubale[1], Miss. Shraddha Pulate[2], Miss. Shivani Shinde[3], Miss. Sai Suryavanshi[4]
Professor, Electronics & Computer Engineering Department[1]
Students, Electronics & Computer Engineering Department[2,3,4]
Amrutvahini College of Engineering, Sangamner, India

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 Gloves, Speech Conversation, Assistive Technology, Sensors

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.

In recent years, there has been growing interest in the development of wearable devices that can translate sign language into spoken or written language, enabling individuals who are deaf and mute to communicate more effectively with others. Sign gloves are a type of wearable device that use sensors and machine learning algorithms to detect hand gestures and translate them into spoken or written language. These devices have the potential to revolutionize communication for individuals who are deaf and mute, enabling them to communicate more effectively with people who do not know sign language. The declining proportion of literate and employed deaf and mute individuals is due to the hearing disability of deaf individuals and the inability to communicate verbally for mute individuals. This results in a lack of communication between normal individuals and deaf and mute individuals. It becomes a problem similar to two people, who know two different languages, and none of them knows any common language, so it becomes challenging to converse with each other. Therefore, they require a physical interpreter, which may not always be practical to arrange. The same kind of problem occurs between normal individuals and deaf or mute individuals.

To overcome this issue, a unique application is presented. The application is an attractive interpreter that translates regular English sentences as text input by a normal person for a deaf person and sign language in the form of gestures by a mute person into synthesized English words, which have a corresponding meaning in sign language. This translates a particular item as an audio output for normal people. This will help normal, deaf, and mute communities by removing the communication gap between them. Gesture-based communication is an essential and only method of communication for deaf and mute individuals. It is a formal language using a set of hand gestures for communication (by the deaf). In this project, a replacement of the flex sensor is used, which plays a significant role, and it is set on fingers. As fingers bend, it changes resistance depending on the amount of bend on the sensor [2].

II. LITERATURE SURVEY

For our project we are surveying some reports and references which are helping us to make it easy and simplest and they are as follows

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. 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]

2. 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]

3. 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]

4. 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]

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. 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]

III. SYSTEM ARCHITECTURE

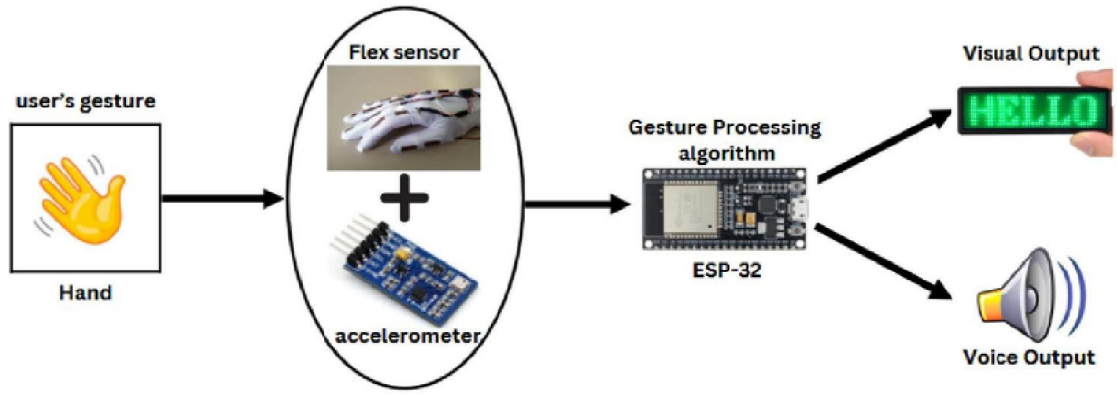

Fig. 1. Architecture Diagram

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

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 across different users.

User Feedback: Gathered input from potential users to assess comfort, usability, and effectiveness leading to iterative improvements.

IV. METHODOLOGY

The system comprises of a method to convert the hand gestures to text and audio messages. The prototype built has modules like PIC microcontroller, bend sensitive flex sensor, analog to digital converter and accelerometer. The architecture diagram is shown in fig. 1.

The working of the GestureSense: Translating Sign Language to Text & Speech system is based on detecting and interpreting hand gestures through sensors and converting them into understandable text and voice outputs. The system uses flex sensors attached to the fingers of a glove to measure the degree of finger bending. When a user performs a specific sign, the flex sensors produce varying resistance values corresponding to each gesture. Along with these, a MEMS (Micro Electro Mechanical System) sensor detects the orientation, tilt, and movement of the hand, providing additional data for accurate gesture recognition. All these sensor signals are sent to the ESP8266 microcontroller, which processes the data by comparing it with pre-stored gesture patterns in its program memory. Once a gesture is identified, the ESP8266 generates a text output that is displayed on the LCD screen, showing the interpreted message. Simultaneously, the same text is sent to a Text-to-Speech (TTS) module, which converts it into an audible voice signal. The audio amplifier strengthens this signal so it can be heard clearly through a speaker. A keypad is included for manual input or control functions, such as resetting the system or entering predefined commands.

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 6, Issue 3, February 2026
Impact Factor: 8.2

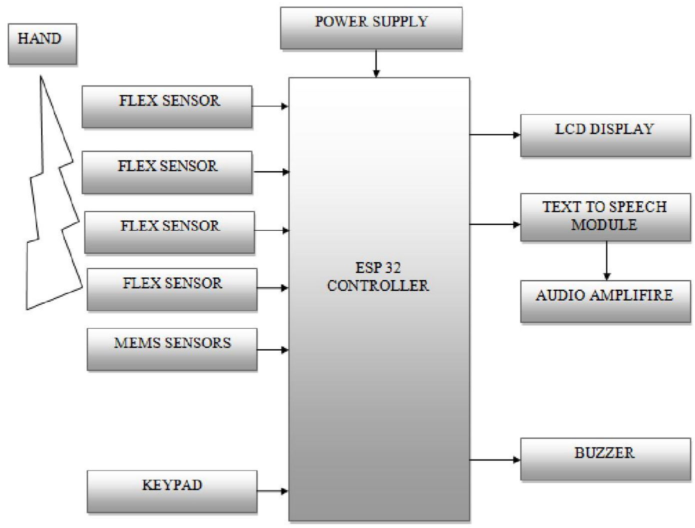

Fig. 2. Block Diagram

A. ESP32:

ESP32 IoT Development Board Module for Arduino.ESP32 is far superior to Arduino UNO. That means it has all the features you need to create your IoT & non-IoT projects, ESP32 can perform as a complete standalone system or as a slave device to a host MCU, reducing communication stack overhead on the main application processor. ESP32 can interface with other systems to provide Wi-Fi and Bluetooth functionality through its SPI / SDIO or I2C / UART interfaces ESP32 is different from ESP32-WROOM-32U in that ESP32-WROOM-32D have on board Antenna you do not need to connect the antenna externally to the ESP32-WROOM-32D IoT Development board

Fig. 3. ESP32 Dev Module

B. MPU-6050 3 AXIS ACCELEROMETER & GYRO SENSOR:

The MPU6050 devices combine a 3-axis gyroscope and a 3-axis accelerometer on the same silicon together with an onboard Digital Motion Processor (DMP) capable of processing complex 9-axis MotionFusion algorithms.

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 6, Issue 3, February 2026
Impact Factor: 8.2

Features :

1. Chip built-in 16bit AD converter, 16-bit data output.
2. I2C Digital-output of 6 or 9-axis Motion Fusion data in rotation matrix, quaternion, Euler Angle, or raw data format.
3. Selectable Solder Jumpers on CLK, FSYNC, and AD0.
4. Digital Motion Processing™ (DMP™) engine offloads complex Motion Fusion, sensor timing synchronization, and gesture detection.
5. Embedded algorithms for run-time bias and compass calibration. No user intervention is required.
6. Digital-output temperature sensor.

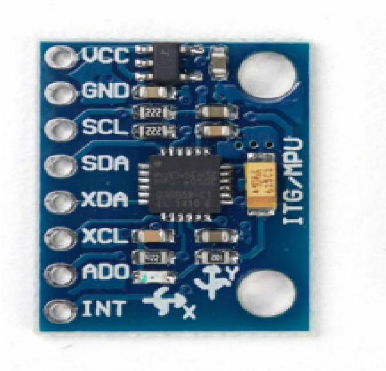

Fig. 4. MPU-6050

C. Flex Sensor:

A simple Flex Sensor 2.2″ – Bend Sensor with a length of 2.2″ which bends and flexes with a physical device. As the sensor is flex, the resistance across the sensor increases. A connector is 0.1″ spaced and breadboard friendly. Applications in – Robotics, Gaming (Virtual Motion), Medical Devices, Computer Peripherals, Musical Instruments, Physical Therapy The flex sensor is a bend detecting sensor that has got numerous applications in robotics, medical and haptic technology. The resistance of these sensors changes in accordance with the bend, which can be measured using any microcontroller.

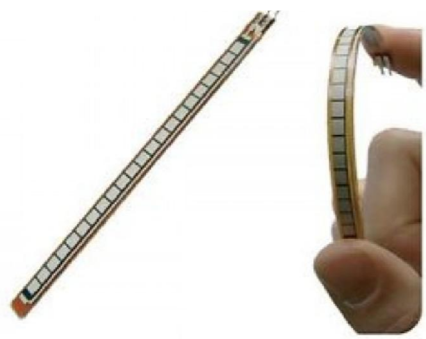

Fig. 5. Flex Sensor

D. Mp3 player:

The Serial MP3 Player v1.2 (often based on the YX5300 chip) is a compact UART-TTL module designed for Arduino/microcontrollers to play MP3/WAV files from a Micro SD card (FAT16/FAT32). It operates at 9600 baud, supports 30-level volume control, and uses 3.3V/5V logic, commonly used with the MD_YX5300 library

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 6, Issue 3, February 2026
Impact Factor: 8.2

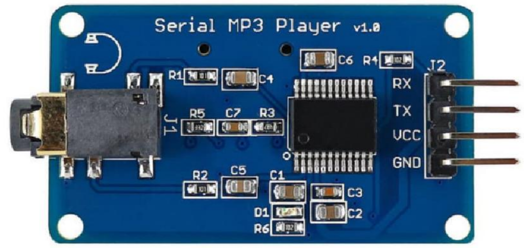

Fig.6. Serial MP3 Player

V. 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.

Smart gloves for dumb and deaf people, also known as assistive technology gloves, are devices that can help those with speech and hearing impairments communicate and interact with the world around them. They typically incorporate sensors and other technology that can translate hand gestures or finger movements into speech or text, allowing users to communicate with others or control devices such as smart phones or computers. Overall, these gloves can be a useful tool for people with speech and hearing impairments to improve their communication and independence. Smart gloves for people who are deaf or hard of hearing can be a useful tool for communication and navigation. They can convert sign language or speech into text or audio, allowing for more effective communication with hearing individuals. They can also supply navigation help through haptic feedback, helping users navigate unfamiliar spaces. However, it is important to note that these devices are not a replacement for sign language or other forms of communication used by the deaf and hard of hearing community. Additionally, the development and accessibility of such technology is still ongoing, and more research is needed to improve the usability of these devices.

VI. 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.

ACKNOWLEDGMENT

It gives us great pleasure in presenting the paper on “Gesture Sense: Translating Sign Language into Text and Speech”. We would like to take this opportunity to thank our guide, Prof. Vilas S. Ubale, Professor, Department of Electronics & Computer Engineering Department, Amrutvahini College of Engineering, Sangamner, for giving us all the help and guidance we needed. We are grateful to him for his kind support, and valuable suggestions were very helpful.

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 6, Issue 3, February 2026
Impact Factor: 8.2

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.

[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.

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

[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).

[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.

[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.

[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).

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

[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.

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

DOI: 10.48175/568