Sign2Speak-A User-Configurable Glove For Mute People

Vidula Meshram
Department of Computer Engineering
Vishwakarma Institute of Technology, Pune
Pune, India
vidula.meshram@vit.edu

Laksh Mundada
Department of Computer Engineering
Vishwakarma Institute of Technology, Pune
Pune, India
laksh.mundada241@vit.edu

Naman Doshi
Department of Computer Engineering
Vishwakarma Institute of Technology, Pune
Pune, India
naman.doshi24@vit.edu

Tanvi Nanaware
Department of Computer Engineering
Vishwakarma Institute of Technology, Pune
Pune, India
tanvi.nanaware24@vit.edu

Vaishnavi Narkhede
Department of Computer Engineering
Vishwakarma Institute of Technology, Pune
Pune, India
vaishnavi.narkhede24@vit.edu

***Abstract*— People with speech and hearing disabilities often face communication difficulties because many people do not understand sign language. To address this problem, an improved Smart Glove system was developed using capacitive touch sensors, MPU6050 and DF Player Mini that converts hand gestures into speech. Unlike the existing project, which supported only fixed gestures, this system allows users to create and save their own personalized gestures through mobile applications. The glove transmits sensor data via Bluetooth to the smartphone, where the app processes and recognizes gestures and then converts them to speech. This project aims to enhance accessibility, personalization, and ease of communication for Indian Sign Language (ISL) users. It recognized gestures successfully with good level of confidence in testing.**

Keywords—Bluetooth, ESP32, Indian Sign Language (ISL), mobile application, personalized gestures.

I. Introduction

Communication is one of the most important parts of human interaction. People who are unable to speak often face difficulties while expressing their needs, emotions, and thoughts in daily life. Although sign language is widely used among speech-impaired individuals, many people are not familiar with it, which creates a communication gap between speech-impaired users and society. Because of this limitation, there is a growing need for affordable and easy-to-use assistive communication systems [6], [14].

Recent developments in embedded systems, wearable devices, and mobile technologies have created new possibilities for smart communication solutions. Gesture recognition systems are becoming increasingly popular because they provide a natural and simple method of interaction [3], [4], [14]. However, many existing systems are expensive, complex, or limited to fixed predefined gestures [1]-[5]. Most available solutions also depend completely on internet connectivity or external processing systems, reducing portability and practical usability.

To address these challenges, this paper presents a Smart Glove for speech-impaired communication using ESP32 and Bluetooth-enabled mobile integration. The proposed system uses five touch sensors placed on the fingertips to detect hand gesture patterns. The ESP32 microcontroller processes the sensor inputs and identifies the performed gesture. The system supports two modes of operation: offline predefined gesture recognition and personalized gesture customization through a mobile application.

Gestures which are commonly used are saved on the memory card connected to the DF Player Mini Module. If these gestures are detected, then the corresponding audio message would be played through the speaker without the need for the internet/mobile. In customized mode, users can connect their gloves to the Android application via Bluetooth and personalize them according to their needs. Users will make the gestures, calibrate them on the app and provide the corresponding word or phrase associated with that particular movement of hand gestures. The corresponding gesture is then translated into spoken form through the mobile's text-to-speech function.

Some of the benefits associated with this project include portability, cost-effectiveness, offline functioning, real-time speech synthesis and customizations among others. The combination of wearable hardware along with a mobile application makes this project flexible enough and more user friendly than traditional devices used for communication purposes. Moreover, this project supports (ISL) gestures making it more unique from already existing models [4], [9].

The primary goal of this project is to design an efficient, and effective assistance communication system that enables those who have difficulty speaking to communicate independently. The designed Smart Glove is intended to minimize communication gaps and facilitate social communication via intelligent and wearable technology.

II. Literature Review

[1] The research paper titled, “Hand Gesture Recognition for User-Defined Textual Inputs and Gestures” by Wang el al. (2025) presents a camera-based gesture recognition system that focuses on individual user style variation. Unlike previously existing systems with predefined sets of gestures and phrases, this system uses a customizable user interface, where an individual can define their own gestures. The system uses the MediaPipe library to extract 21 hand keypoints per frame from a single RGB camera. The system supports four modes: hand gesture verification, new gesture addition, gesture management, and model analysis. In the user study (N=12), which was based on User Experience Questionnaire, the overall satisfaction score was 6.96/10,

where the scores for Novelty and Attractiveness were 7.63 and 7.24 respectively. One of the disadvantages of this research is that it does not cater for any dynamic or continuous gestures but only for static gestures.

[2] In the paper "Vision Based Approach to Sign Language Recognition" by Gupta, Ramjiwal, and Jose (2018), the authors suggest an entirely camera-based algorithm for the recognition of predetermined number of hand movements that help deaf and mute people in their day-to-day interactions. The system processes the input video frame-by-frame and applies YCbCr color space in order to distinguish skin from the background. After this, the computation of center of gravity (COG) is done and the farthest point from it is considered to make a circle around the gesture to provide a consistent reference for analyzing the hand's geometry. The processed gesture is compared against a pre-stored set of gesture templates. There is the use of 2D cross-correlation for matching. This method can handle hand movements and the gesture is recognized. One of the limitations is that there are no dynamic updates and the background should remain clear for detecting a hand gesture.

[3] The paper "Hand Gesture Recognition using Machine Learning" by Na Rang, Jeronima, Mora, and Jardim (2025) from the journal Procedia Computer Science gives a comprehensive method of identifying static gestures which are associated with all 26 letters of the Latin alphabet using Random Forest machine learning algorithm. A custom dataset which consists of 54000 RGB images was collected under specific lighting conditions using a webcam, and the gestures were performed by 5 individuals. The feature extraction was done using the MediaPipe Hands Landmarks framework which detected 21 knuckle points for each image. These were then used as input features to the random forest model, which was trained on 67% of the dataset and tested on the remaining 33%. According to the test results, the accuracy of the model was 92.3%. The algorithm had some problems in recognizing small differences between finger positions. Additionally, an application in Python language was built for validation of the classifier in real time using a camera. Though this research provides the users with mobile application functionality, one of the drawbacks of this system is that it works for predefined Latin letters and doesn't input new gestures.

[4] In the paper "A Robust Sign Language and Hand Gesture Recognition System using Convolutional Neural Networks" by Saurav Kumar, Pratiyush Kumar, Priyesh Mishra, and Pragya Tewari (2023), the authors present a vision-based technique which is based on deep learning that recognizes Indian Sign Language (ISL). The technique uses a custom dataset to fine-tune a CNN model. The system recorded an accuracy of 98.6%.

Based on the reviewed papers, it can be seen that most gesture recognition systems use cameras, machine learning algorithms, or gestures themselves for interaction [1]–[4]. They has high accuracy in most cases, but the majority of such solutions requires a specialized environment, increased computing capabilities, or cannot allow offline communication and personalization. Additionally, many gesture detection devices based on smart gloves uses flex sensors for tracking finger movements [5]. Though flex sensors have been successfully used for years, their analog output should be calibrated and is not so accurate [11]. The suggested Sign2Speak solution solves those problems since capacitive sensors and the MPU6050 sensor are used for gesture detection [8], providing accurate binary signals by touch sensors and minimizing additional processing required. Moreover, the suggested device allows for offline and more customized gestures can be added using mobile application using Bluetooth communication. Thus, Sign2Speak provides an efficient, affordable, and reliable communication method for people with speech impairments [5]–[7].

III. Methodology

A. System Overview

The extended Smart Glove is a wearable device which translates gesture to speech for hearing impaired individuals [9], [10], [13]. The glove has an ESP32 as a microcontroller with each finger consisting of a touch sensor, an MPU6050 motion sensor to detect hand movements, DF Player where pre-defined gestures are stored and a speaker which plays the audio when a pre-defined gesture is recognized [8]-[10].

For a user who wants to customize their own gestures, it can be done via mobile application [1]. The user can click on the calibrate button and record the gestures. It is then saved in ESP32. When that personalized gesture is recognized the audio is played through the mobile device itself as it has a Text-to Speech module. This allows the user to update the vocabulary and help in communicating well.

B. Flowchart

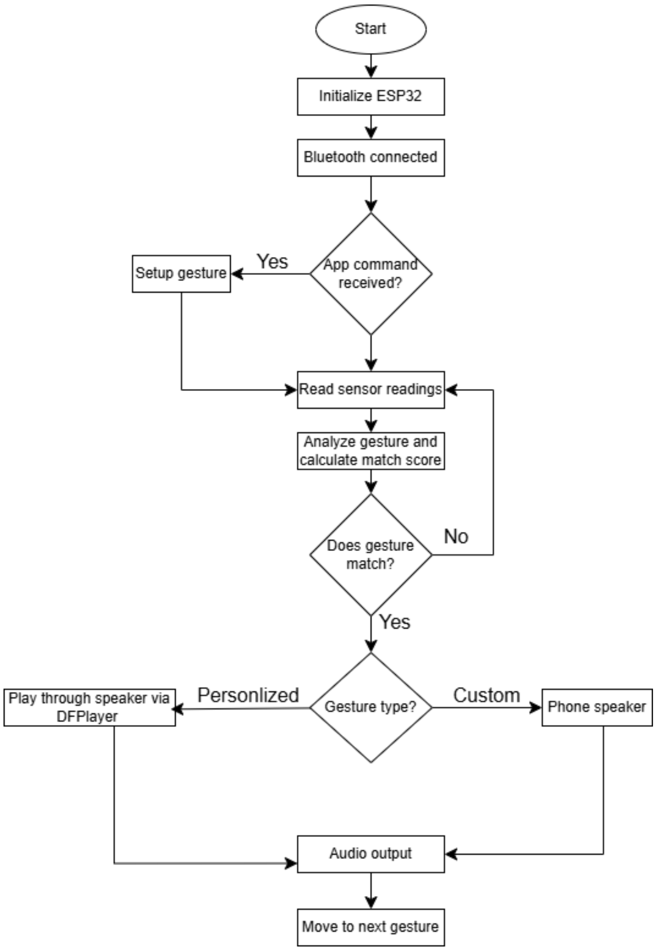

Fig. 1. Flowchart of the system

C. *System Architecture and Process Flow*

Fig. 1. Shows the end-to-end process flow of the system from device initialization to audio output. The system operates continuously in two parallel paths managed by the dual-core architecture.

Core 0- Bluetooth and Configuration Layer: There is a dedicated background task that handles the bluetooth connection. It allows users to add a new gesture, calibrate the current one, delete a saved gesture. All the cross-core memory access to the shared gesture array is kept in check by the semaphore that prevents both processors from changing the gesture data at the same time, which avoids system crashes.

Core 1- Sensor Reading and classification loop- The main loop continuously reads data from the motion sensor and the capacitive touch sensors. To filter the accidental vibrations and readings, the system converts the raw accelerometer values to a filtered scale of 0-100 using an 8-sample filter. This is then sent directly to the gesture matching/classification engine.

D. *Gesture Recognition Process*

The system checks how a hand movement matches the saved gesture by looking at the two things:

- Finger touch
- Wrist position

Every movement gets an overall match score out of 100%. It is calculated like:

$$M = 0.50 * TouchScore + 0.25 * OrientationScore(X) + 0.25 * OrientationScore(Z)$$

To avoid false triggers, the gesture is recognized if its total score is above 60% and it beats the second closest gesture by at least 12%.

To eliminate accidental triggers, the system double checks using a 4-out-of 5-rule. A gesture is confirmed or matched only if it passes the high-score test in 4 out of 5 consecutive scans.

Pre-defined gestures are stored in the SD card present in the DF Player mini. For custom gestures, they are saved to the same internal memory from where the profiles are received over bluetooth from the mobile application i.e, the non-volatile memory.

E. *Mobile Application Implementation*

The mobile application is developed for Android and is connected via standard Bluetooth. The app is split into two sections.

I. Adding and Setting up Gestures: Users can create custom gestures without needing to reprogram the glove. Just by clicking on “CALIBRATE” the glove takes a snapshot of the hand and finger position. Then typing the name and clicking on “SAVE”, the custom gesture is added and this data is sent to glove’s memory.

II. Live Recognition and Voice Mode: When the application receives a continuous stream of sensor data from the glove, each packet extracts the sensor states and X-axis and Z-axis orientation values. After confirmed recognition, the phrase is retrieved from the memory and passed to the Android Text-To-Speech. The app instantly reads it out using the phone’s Text-To-Speech (TTS) voice.

IV. Results And Discussion

A. *Device Implementation*

The successful prototyping of the Smart Glove involved the use of ESP32, touch sensors, MPU6050, DFPlayer Mini, and an Android application. This enabled it to detect touch combination gestures and provide speech output accordingly.

For capturing the finger movements, five touch sensors were fitted in all the fingertips. For tracking hand position and movement direction, the MPU6050 sensor was used. These two sensors kept on detecting the value of gesture movement and comparing with the existing ones stored in the ESP32 microcontroller memory.

If any match is found, the corresponding voice message would be created. In case of pre-programmed gesture recognition, the audio will play using the DFPlayer Mini hardware and speaker, while in customized gesture recognition, the text message will be transmitted using Bluetooth in the mobile app for generating speech output from the phone speaker.

As shown in Fig. 2 the mobile app has been designed for gesture calibration and communication via Bluetooth. From here users can control their connections, add/delete any gesture and check connectivity status.

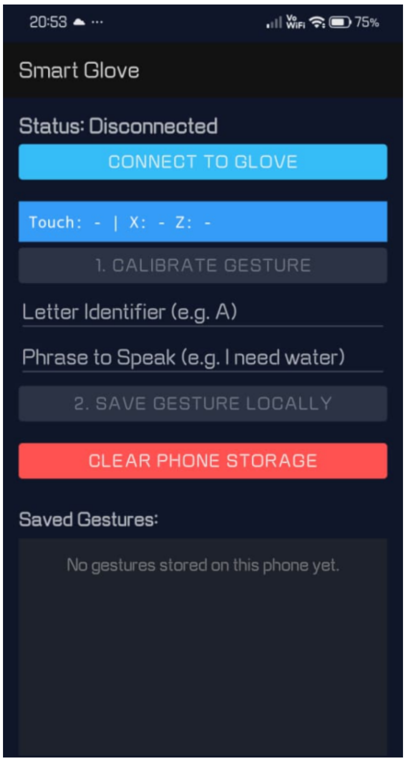

Fig. 2 Smart Glove Application

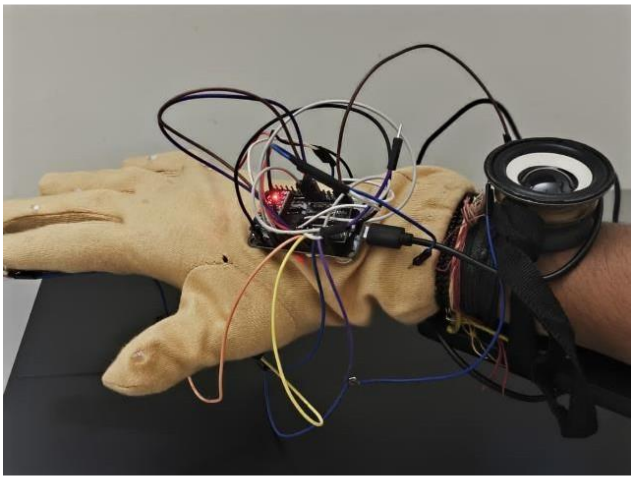

Fig. 3 Glove

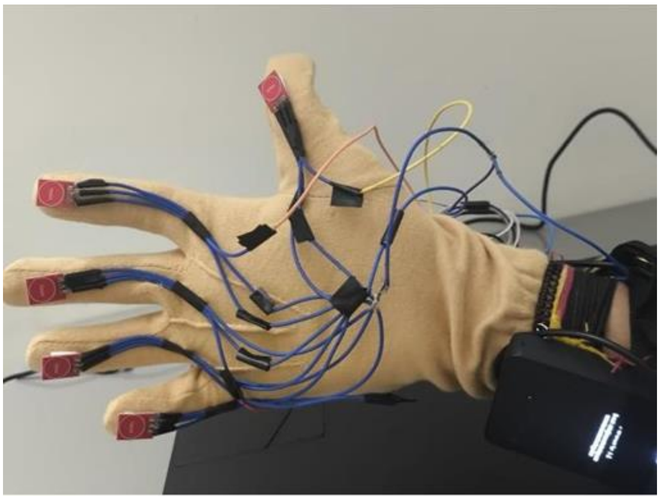

Fig. 4 Placement of touch sensors

B. Gesture Recognition Results

Various combinations of gestures were repeatedly tested in order to analyze the system's performance. Both pre-defined gestures and trained gestures were analyzed independently.

Pre-defined gestures demonstrated better results due to their being calibrated and saved within controlled conditions.

TABLE I. OBSERVATION

Operation	**Observed Performance**
Sensor Detection	Fast (~10 ms)
Gesture Processing	Stable
Bluetooth Communication	Smooth
Voice Output	Immediate
Overall System Response	Suitable for real-time use

Custom gestures demonstrated somewhat lower accuracy due to the possibility of performing gestures differently each time.Utilization of touch sensors together with motion sensors increased accuracy by providing more detailed information about gestures.

TABLE II. PER-GESTURE RECOGNITION RESULTS (SINGLE CALIBRATED USER)

Gesture	**Touch pattern**	**BendX**	**BendZ**	**Tolerance**	**Status**
A	[1,0,0,0,0]	22.0	100.0	18.0°	Recognized
B	[1,0,1,0,0]	72.0	63.0	22.0°	Recognized
C	[0,0,0,0,0]	39.0	84.0	25.0°	Recognized
D	[1,1,0,0,0]	48.0	86.0	18.0°	Recognized

E	[1,1,0,0,0]	73.5	68.9	22.0°	Recognized
F	[0, 1, 1, 0, 0]	31.2	94.2	18.0	Recognized
G	[1, 0, 1, 0, 0]	38.0	49.0	28.0	Recognized
H	[0, 0, 0, 1, 1]	41.0	90.0	18.0	Recognized
I	[0, 1, 0, 0, 0]	60.0	81.2	22.0	Recognized
J	[0, 1, 0, 0, 0]	19.5	95.3	25.0	Recognized

The system successfully recognized all tested ISL gestures.

1. Serial Monitor Output Analysis

For a successful recognition of alphabet 'A' the serial monitor showed a real-time information:

T: [1,0,0,0,0] B: [X:23.5 Z:98.7] -> A (93% | margin: 38%)

T: [1,0,0,0,0] B: [X:22.8 Z:99.2] -> A (95% |margin: 41%)

T: [1,0,0,0,0] B: [X:22.1 Z:100.5] -> A (96% |margin:43%)

T: [1,0,0,0,0] B: [X:23.0 Z:99.8] -> A (94% | margin:40%) –

>RECOGNIZED

The alphabet 'A' got recognized on the 4th reading.

The system rejects gestures below 60% threshold or with an insufficient margin below 12%.

C. Overall Discussion

The implemented system has shown satisfactory performance throughout the testing process, allowing the system to perform satisfactorily both in offline communication and in personalized gesture-based communication.

One of the strengths of this project is the ability of the user to create their own gestures rather than rely on the predefined gestures in the system. The offline voice capability also helps make the use of this application better when there is no access to Bluetooth or internet.

There are some challenges faced by the application. Accuracy was found to be somewhat lessened whenever the hand position of the user changed, and the inconsistency in gestures made. There is also some issue with external movement and improper sensor contact. Despite all these weaknesses, this system proves to be useful as an assistive device for communication for those with difficulty in speaking.

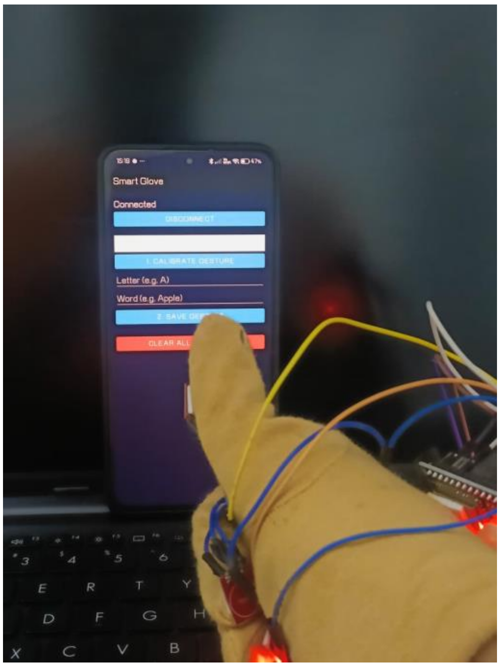

Fig.5 Sign2Speak

V. Future Scope

The project can be made better by making use of a few technologies.

1. Machine Learning-Based Gesture Recognition

The current system is rule-based. Future enhancement can be made by integrating machine learning techniques such as k-NN, SVM, or small neutral networks can be used to improve accuracy [3], [4], [7]. Using AI directly on the device using Tensor Flow Lite will also keep user data private.

2. Dual-hand and Full-sentence support

The current system uses only one glove for one hand, so it can recognize only one-handed signs. Adding a second glove would increase the vocabulary and make the communication more natural.

3. Cloud Synchronization and Collaborative vocabulary

Cloud allows users to back up and restore their gesture profiles on different users [12]. A common library of frequently used ISL phrases could help new users start quickly without creating every gesture from scratch.

4. Emergency and context-aware phrase templates

The system could include a special emergency mode with simple yet quick gestures for urgent phrases such as "I need help" or "call ambulance".

5. Multi-language

Currently, the system mainly supports English. Future versions could support Indian languages such as Hindi, Marathi, etc.

6. Improve wearability

The current prototype uses large electronics modules attached to the glove, which can feel messy and bulky. Future designs could use a custom compact PCB and a small rechargeable battery built into a wristband. Decreasing the number of wirings could improve durability.

VI. Conclusion

This paper presented a communication system for speech-impaired people that is, Smart Glove. The system combines wearable, embedded systems, and mobile application support that is used to change hand movements into personalized voice output. The glove used has five touch sensors that detect and recognize movements from the user by linking with the ESP32 microcontroller.

The system has the ability to recognize both offline predefined gestures and customized user-defined gestures. Predefined gestures are recorded in the SD card present in DFPlayer Mini module where the system does not require the use of internet or Bluetooth connection to produce voice. Additionally, a user is able to customize gestures that he/she wishes to use using the mobile application through the Bluetooth connection. Customized gestures can then be changed to voice on the mobile phone speaker.

The suggested solution is portable, affordable, simple to use, and applicable to real-world communications. Different from other solutions that work only with predefined gestures, the Smart Glove that is being developed enables users to establish personalized communication sequences that are appropriate for their needs. Adding the possibility of operating in an offline mode makes the system more useful for real-world use.

This project is a good example of how wearable assistive devices can be used to eliminate communication obstacles. In the future, this system may be improved through using machine learning techniques to make gesture recognition better, expanding support of more sign languages, and adding the functionality of working in cloud networks for remote communications.

This solution in the field of assistive technologies in smart healthcare is affordable, portable, and able to perform all tasks quickly and efficiently.

Acknowledgement

We would like to appreciate and thank the team members and colleagues who offered valuable suggestions and support during this phase.

We also like to thank Vishwakarma Institute of Technology, Pune for providing the resources and an environment for learning. The faculties were also helpful and guided us to make this project a success.

References

[1] Y. Wang et al., "Hand Gesture Recognition for User-Defined Textual Inputs and Gestures," 2025 International Conference on Human-Computer Interaction and Intelligent Systems, 2025.

[2] A. Gupta, S. Ramjiwal, and B. Jose, "Vision Based Approach to Sign Language Recognition," International Journal of Computer Applications, vol. 179, no. 18, pp. 1–5, 2018.

[3] N. Rang, J. Mora, and J. Jardim, "Hand Gesture Recognition using Machine Learning," Procedia Computer Science, vol. 245, pp. 1123–1130, 2025.

[4] S. Kumar, P. Kumar, P. Mishra, and P. Tewari, "A Robust Sign Language and Hand Gesture Recognition System using Convolutional Neural Networks," 2023 International Conference on Smart Computing and Communication Systems, 2023.

[5] S. Maharajan, S. Fernando, S. Shrestha, D. Mariyanayagam, and B. Virdee, "Sign Language Detection and Translation Using Smart Glove 2.0 Using Hybrid CNN-Transformer Model," Proceedings of Data Analytics and Management (ICDAM 2025), Lecture Notes in Networks and Systems, vol. 1597.

[6] S. N. Sawant, "Sign language recognition system to aid deaf-dumb people using PCA," Int. J. Comput. Sci. Eng. Technol. (IJCSET), vol. 5, no. 5, 2014.

[7] H. Ahuja, K. Varadarajan and M. Jammal, "Towards Smart Interaction: Hand Gesture Recognition Using Machine Learning in IoT Scenarios," IEEE Global Conf. Artif Intell. Internet of Things (GCAIoT), Dubai,2023, pp. 48-52.

[8] K. S. Abhishek, L. C. F. Qubeley, and D.Ho, "Glove-based hand gesture recognition sign language translator using capacitive touch sensor," in 2016 IEEE Int. Conf. Electron Devices Solid-State Circuits (EDSSC), 2016.

[9] S. Y. Heera, M. K. Murthy, S. V. S, and S. Salvi, "Talking hands- An Indian sign language to speech translating gloves," in 2017 Int. Conf. on Innov. Mech. for Ind. Appl. (ICIMIA), 2017, pp. 746-751

[10] A. Snehaa, S. Suryaprakash, A. Sandeep, S. Monikapriya, M. Mathankumar and P. Thirumoorthi, "Smart audio glove for deaf and dumb impaired," 2021 Int Conf on Adv Electr, Electron., Commun, Comput. Autom (ICAECA), 2021, pp. 1- 6.

[11] U. Farooq, A. Asmat, M. S. B. M. Rahim, N. S. Khan and A. Abid, "A Comparison of Hardware Based Approaches for Sign Language Gesture Recognition Systems," Int. Conf. on Innov. Comput. (ICIC), 2019, pp. 1-6.

[12] D. Balazh, V. Mrak, A. Sydor, V. Andrushchak, B. Rusyn and T. Maksymyuk, "Intelligent IoT Control System based on Hand Gesture Recognition," IEEE 5th Int. Conf. Adv. Inf. Commun. Technol (AICT), Ukraine, 2023, pp. 102-104.

[13] A. M. H. Khandoker, M. A. Hossain, and M. R. Islam, "Smart Glove Based Sign Language Recognition System for Deaf and Dumb People," *International Conference on Electrical, Computer and Communication Engineering (ECCE)*, 2019.

[14] A. Osman Hashi, S. Zaiton Mohd Hashim and A. Bte Asamah, "A Systematic Review of Hand Gesture Recognition: An Update From 2018 to 2024," in IEEE Access, vol. 12, pp. 143599-143626, 2024.