ISSN No:-2456-2165 https://doi.org/10.5281/zenodo.14591355

Development of Gesture-to-Voice Glove for Individuals with Speech and Hearing Impairments

Shwet Kamal[1]
UG Scholar – ETE
Bangalore Institute of Technology
Bangalore, Karnataka

Nithyashree M[2]
UG Scholar – ETE
Bangalore Institute of Technology
Bangalore, Karnataka

B Sudha[3]
Assistant Professor – ETE
Bangalore Institute of Technology
Bangalore, Karnataka

Abstract:- This project presents the development of a cost-effective assistive device aimed at bridging the communication gap between deaf-mute individuals and the hearing community. The device is built around Smart Gloves embedded with flex and motion sensors, designed to detect hand movements corresponding to sign language gestures. These gestures are processed by a microcontroller that interprets the signals and converts them into voice output, allowing real-time verbal communication. The system integrates with a server that translates sign language into digital speech, providing an accessible and efficient communication solution. The prototype is capable of recognizing all 26 letters of the English alphabet, numbers, and commonly used words, covering a wide range of everyday communication needs. With its affordability and ease of use, the device offers a practical and economical solution for individuals with speech disabilities, making it accessible across diverse economic backgrounds. The gloves demonstrate high accuracy in recognizing gestures, ensuring reliable communication without requiring extensive training or technical skills. Although the prototype shows promising results, further testing with a more diverse set of users is recommended to assess its real-world effectiveness and optimize its performance. This solution has the potential to significantly enhance the communication experience for individuals with speech impairments.

***Keywords*:-** *Gesture; Accessibility; Communication; Sensors; Speech; IoT; Recognition.*

I. INTRODUCTION

Effective communication is essential for fostering connections and opportunities, but individuals with speech and hearing impairments often encounter significant challenges in their interactions. These barriers can lead to social isolation and reduced access to educational, professional, and social opportunities. While various assistive technologies exist, many are prohibitively expensive, overly complex, or limited in functionality, leaving many individuals without adequate solutions.

This project proposes the development of a cost-effective and user-friendly assistive device designed to address these challenges. The device incorporates Smart Gloves equipped with advanced flex and motion sensors to accurately interpret sign language gestures. These gestures are then processed and converted into audible speech in real-time, enabling seamless verbal communication. By eliminating the need for additional tools or interpreters, the system provides a straightforward and efficient way to bridge the communication gap. [1]

The primary goal of this innovation is to offer an affordable and practical solution that enhances communication for individuals with speech disabilities. The device is designed to be intuitive and accessible, ensuring that users can operate it with minimal training. By facilitating real-time communication, it empowers users to participate more freely in daily activities, social interactions, and professional engagements. This technology represents a significant step forward in creating a more inclusive and equitable society, offering individuals with speech impairments the tools they need to engage meaningfully with the world around them.

II. LITERATURE SURVEY

As shown in TABLE 1., from papers 1, 2 and 3 flex and motion sensors are referred, from paper 1 Machine Learning algorithms are referred, from papers 1, 2 and 3 gesture recognition are referred and from papers 1, 2 and 3 IoT techniques are also referred. From various papers, it is observed that the limitations of the papers are slow processing speed, high initial costs and accuracy challenges. The future scope is to develop enhanced gesture recognition and improve accessibility and usability with more language integration apart from English and Kannada.

ISSN No:-2456-2165 https://doi.org/10.5281/zenodo.14591355

Table 1 Comparative Study of Research Papers

Parameter	Paper 1: Smart Gloves for Deaf and Dumb People (IJARSCT) (2024)	Paper 2: Smart Hand Gloves for Deaf and Dumb (IRJMETS) (2023)	Paper 3: Smart Hand Gloves for Deaf and Dumb People (IJRASET) (2023)
Technology Used	IoT, machine learning, AI algorithms	Gesture recognition, text-to-speech conversion, LCD display	Gesture-based recognition using Arduino UNO and Bluetooth communication
Sensors	Flex sensors, accelerometers	Flex sensors	Accelerometer, Flex sensors
Microcontroller Used	Arduino UNO	PIC16F877A	Arduino UNO
Data Processing	Captures sensor data, and translates gestures into text/speech using ML algorithms	Converts sensor signals to text (LCD) and speech (voice module)	Analog signals from sensors are converted to digital using Arduino
Output Format	Text and speech; include haptic feedback and vibration.	Text displayed on LCD; synthesized speech output	Text, audio (via Bluetooth-connected app)
Advantages	Enhances communication, promotes independence, is portable, and supports multiple languages	Real-time communication, portable, aids speech-impaired users	Portable, real-time communication, low cost
Limitations	Accuracy challenges, limited battery life, high initial costs, difficulty with complex gestures	Slower processing speed	Slower processing

➢ *Problems Identified*

From the Literature Survey carried out, several problems were identified. They are:

- Many systems are prohibitively expensive, making them inaccessible to individuals from low-income backgrounds.
- Existing solutions often focus only on basic gestures, with limited support for complex or customizable sign language.
- Some devices require extensive training or technical expertise, which can discourage widespread adoption.
- Systems dependent on continuous internet connectivity or external tools are impractical in certain environments.
- Most systems lack multi-language capabilities, especially for regional languages like Kannada, which restricts their usability among non-English-speaking users. [2][3][4]

➢ *Proposed Solution*

Following the problems identified in the existing system, the solutions proposed are:

- To build a smart glove that supports multiple regional languages.
- To reduce the cost and complexity of the model by using alternate IoT tools and equipment.
- To increase the gesture recognition speed of the glove.
- To use the advanced and reliable Machine Learning algorithm for recognizing the gestures.

By following the above methods, it can be made sure to make a user-friendly smart glove that can be used to bridge the communication gap and improve the quality of life for individuals with speech and hearing impairments.

➢ *Software and Hardware Requirements*

Hardware and software requirements are shown in the component diagram in Fig.1.

- *The Hardware Requirements for the Project are:*

✓ *Flex Sensor*
✓ *ESP32 Microcontroller*
✓ *MPU-6050*
✓ *Hand Gloves*
✓ *Speaker with Amplifier Module*
✓ *Power Supply*
✓ *Jumper Wires*

- *The Software Requirements for the Project are:*

✓ *CPP programming language*
✓ *Python for Machine Learning*
✓ *Arduino IDE*

ISSN No:-2456-2165 https://doi.org/10.5281/zenodo.14591355

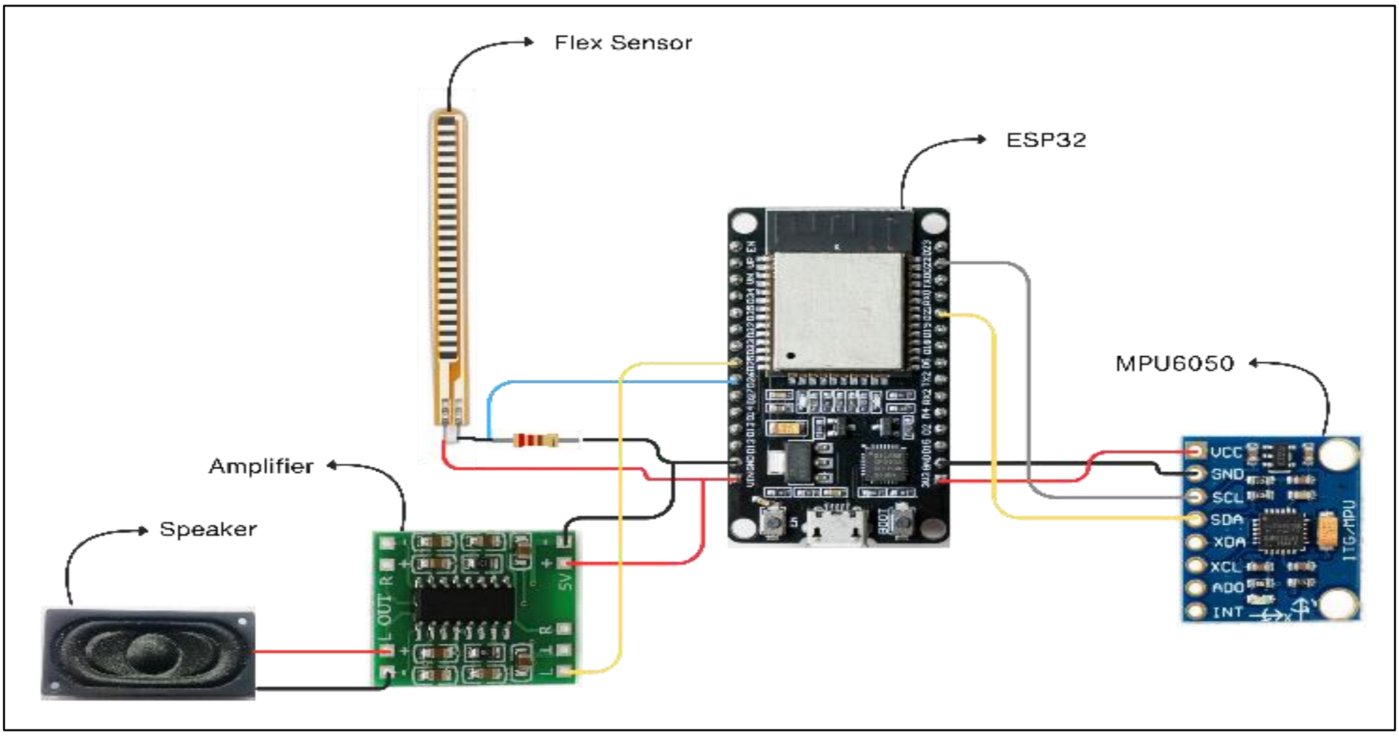


Fig 1 Component Diagram

III. METHODOLOGY

A. *System Overview*

The **Gesture-to-Voice Glove** is designed as a wearable device that captures hand gestures and converts them into audible speech in real-time. The system integrates flex and motion sensors embedded in gloves, a microcontroller for processing sensor data, and a server- based machine learning model for gesture recognition. The final speech output is delivered through a speaker module. The entire system aims to provide a seamless communication experience for individuals with speech and hearing impairments.

➢ *Hardware Requirements*

The hardware requirements for the project are:

- Flex Sensors: Detect bending of fingers to capture static gestures. Measures resistance changes corresponding to finger movements.
- MPU-6050 Motion Sensor: Combines a 3-axis gyroscope and 3-axis accelerometer. Tracks hand orientation and dynamic movements.
- ESP32 Microcontroller: Processes sensor inputs and communicates with the server for gesture recognition and voice synthesis. Features built-in Wi-Fi and Bluetooth for connectivity.
- Speaker with Amplifier Module: Converts text data into audible speech for real- time communication.
- Hand Gloves: Serves as the base for integrating sensors and the circuit. Designed to ensure comfort and durability.
- Regulated Power Supply: Provides stable voltage for powering the components.
- Jumper Wires: Connects sensors, microcontrollers, and other modules efficiently. [2][3]

➢ *Software Requirements*

The software tools and frameworks employed include:

- Embedded C: Used for programming the microcontroller to handle sensor data and communication.
- Python: Facilitates machine learning algorithms and gesture recognition on the server. Arduino IDE: Provides a platform for writing and uploading code to the ESP32 microcontroller.
- Text-to-Speech (TTS) Library: Converts recognized gestures into natural-sounding speech.

The Fig 2 & 3 explains the working of a smart glove system by showcasing two interconnected processes: the hardware flow and the software/data processing flow.

On the hardware side, the flex sensors measure the bending or movement of the fingers, while the MPU6050 sensor captures the hand's orientation, such as tilt or rotation. Both these sensors send their respective data to the ESP32 microcontroller, which acts as the central processing unit. The ESP32 processes the input data from the sensors and generates an audio output, which is then sent to a connected speaker for playback. This setup allows the glove to convert physical hand movements into audio signals.

On the software side, the data collected from the flex sensors and MPU6050 is used as input sensor data. This combined data is passed to a machine learning model, which interprets the hand gestures or movements and generates a corresponding text string. The text string is then converted into speech using a text-to-speech (TTS) system, which creates an audio file. The audio file is returned to the hardware system, allowing the speaker to deliver verbal feedback or communication.

Together, the hardware and software flows enable the smart glove to translate hand gestures into meaningful audio feedback, making it useful for applications such as gesture-based communication or assistive devices for individuals with speech impairments.

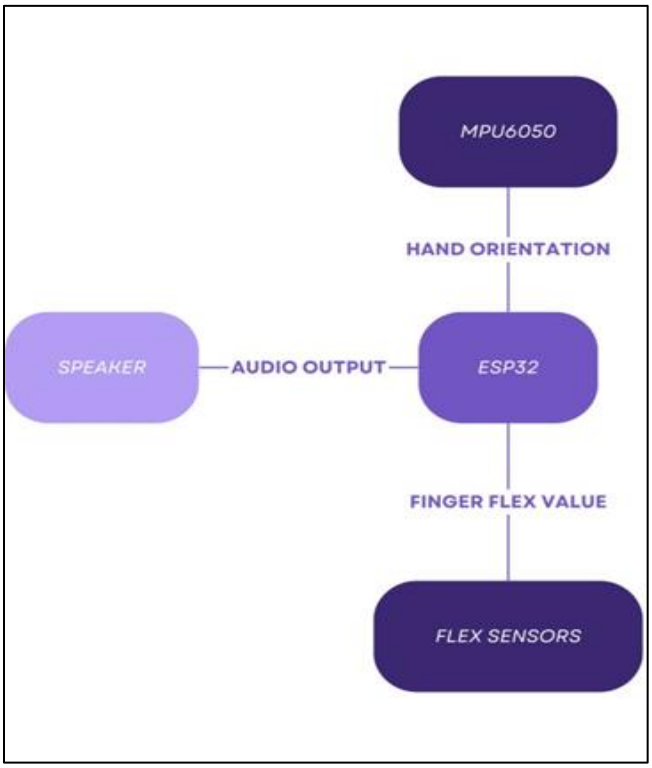


Fig 2 Hardware Flow

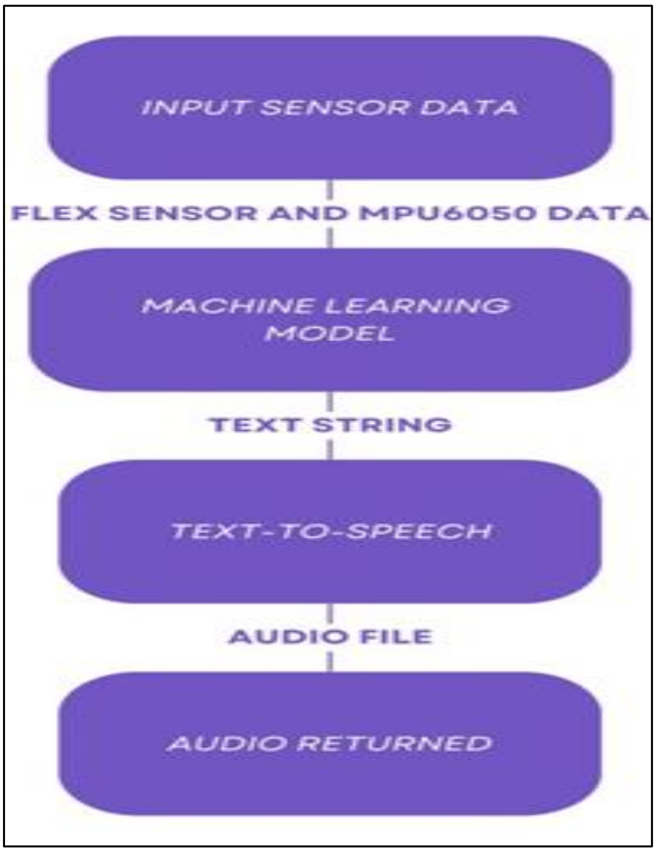


Fig 3 Software Flow

B. Methodology Workflow

The development of the Gesture-to-Voice Glove is organized into several interconnected steps, ensuring a systematic approach to achieving real-time gesture recognition and voice synthesis.

➢ *Gesture Data Acquisition*

The first step in the workflow involves capturing hand gestures, which are central to the system's function. To achieve this, flex sensors and motion sensors are embedded in the gloves. The flex sensors are placed along the fingers and measure the degree of finger bending when a user forms a gesture. The more the finger bends, the higher the resistance in the sensor, providing a measurable signal. In addition, motion sensors, such as the MPU- 6050, track the hand's orientation and movement in three-dimensional space. These sensors help capture dynamic hand gestures that accompany the static gestures detected by the flex sensors. The data from both sensor types are captured as raw analog signals representing the user's hand movements, which are then sent to the microcontroller for processing. This data forms the foundation for gesture recognition and accurate speech output. [1][2]

➢ *Data Preprocessing*

Once the raw sensor data is captured, it needs to be processed for meaningful interpretation. The **ESP32 microcontroller**, which acts as the central processing unit for the system, plays a crucial role in this step. The microcontroller converts the raw analogue signals from the flex and motion sensors into digital data through an Analog-to-Digital Converter (ADC). The digital conversion standardizes the data and ensures it is compatible with the machine learning model. At this stage, the microcontroller also performs **data filtering** to remove noise and irrelevant information that may affect the accuracy of gesture recognition. The preprocessed data is then packaged and transmitted to a server for more detailed analysis and classification.

➢ *Gesture Recognition*

The next step in the workflow is **gesture recognition**, where the preprocessed data is sent to a **remote server** hosting a machine learning model, such as **k-Nearest Neighbors (KNN)**. The KNN algorithm is ideal for this task because it classifies new data based on patterns learned from previous examples. The system uses a training dataset consisting of known hand gestures corresponding to the 26 letters of the alphabet, numbers, and commonly used words. Once the raw sensor data is transmitted, the machine learning model compares the incoming data with the trained dataset to classify the gesture. Refer to Fig 4. for a clear understanding of the KNN algorithm concerning the upcoming example. For example, if the user forms a sign corresponding to the letter "A," the model identifies the pattern and classifies it as "A". This process is performed rapidly and accurately, ensuring real-time recognition.

ISSN No:-2456-2165 https://doi.org/10.5281/zenodo.14591355

- *Definition of KNN:*

A supervised algorithm for classification and regression, predicting outcomes based on the nearest data points.

- *Process:*

✓ Choose k.
✓ Calculate distances (e.g., Euclidean).
✓ Identify k nearest neighbours.
✓ Use majority vote (classification) or mean (regression) for prediction.

- **Advantages**: Simple, non-parametric.
- **Disadvantages**: Computationally expensive, sensitive to irrelevant features, requires optimal k.
- **Applications**: Image recognition, recommendation systems, medical diagnosis.

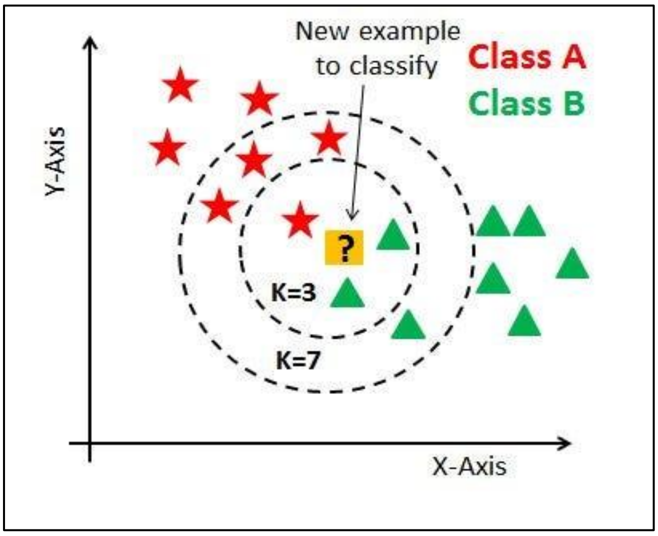


Fig 4 KNN Algorithm

➢ *Text-to-Speech Conversion*

After the gesture is recognized and classified, the system moves to the Text-to-Speech (TTS) Conversion stage. The recognized gesture is first converted into text (e.g., "A" or "Hello"). This text data is then passed to a TTS engine, which converts it into audio form. The system also integrates a Google API-powered English-to-Kannada translation feature, enabling gestures to be output in Kannada for broader accessibility and regional relevance. The TTS system uses a pre-programmed voice database to generate human-like speech from the text input. This process ensures that users' gestures are translated into natural, understandable speech, fostering seamless communication between the deaf-mute individual and the hearing community. [4]

➢ *Real-Time Speech Output*

The final step in the workflow is delivering the real-time speech output to the user and their audience. Once the text has been converted into speech, the audio is transmitted back to the ESP32 microcontroller. The microcontroller then sends the audio signal to the speaker module integrated into the glove. The speaker module amplifies and plays the sound through the speaker, allowing the user's gesture to be heard clearly. Fig 5. shows the block diagram of the methodology with component data. This real-time output ensures that communication is instantaneous, eliminating delays and allowing for natural conversation. The system operates efficiently, providing an accessible and seamless communication solution for individuals with speech and hearing impairments.

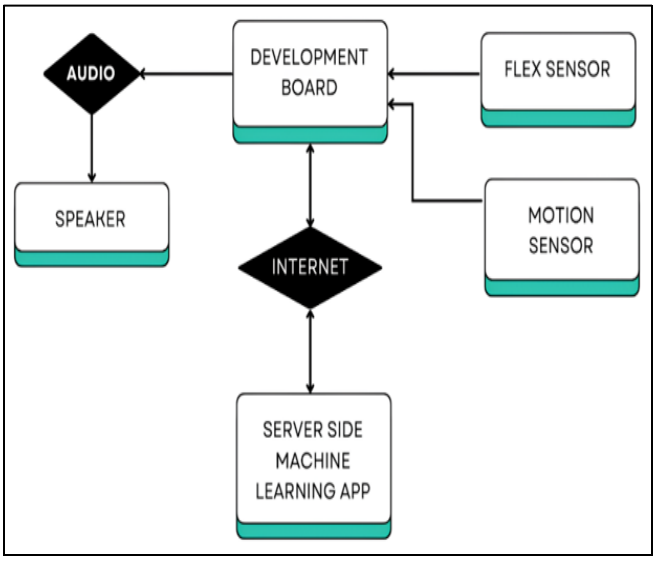


Fig 5 Block Diagram with Component Data

IV. RESULTS

The glove demonstrated an accuracy rate of 92%, as calculated during testing. This high accuracy was achieved through the precise integration of flex and motion sensors, along with the robust implementation of the KNN algorithm for gesture classification. Additionally, the system incorporates a Google API-powered English-to-Kannada translation feature, enabling users to output gestures in Kannada for enhanced accessibility in regional contexts.

➢ *Accuracy Calculation:*

The system was tested for a total of 31 gestures, which included:

- 26 gestures for alphabets (A-Z).
- 5 additional gestures: "help," "sorry," "eat," "I love you," and "thanks."

Each gesture was performed multiple times to ensure consistency, with a total of 500 gestures performed during testing. Out of these, 460 gestures were correctly identified by the system.

➢ *The accuracy was Calculated as follows:*

The breakdown of the test results per gesture category showed consistent performance across all gestures, with minor errors due to subtle variations in finger positioning or hand movement. Fig 6. below shows the setup of the hand glove.

ISSN No:-2456-2165 https://doi.org/10.5281/zenodo.14591355

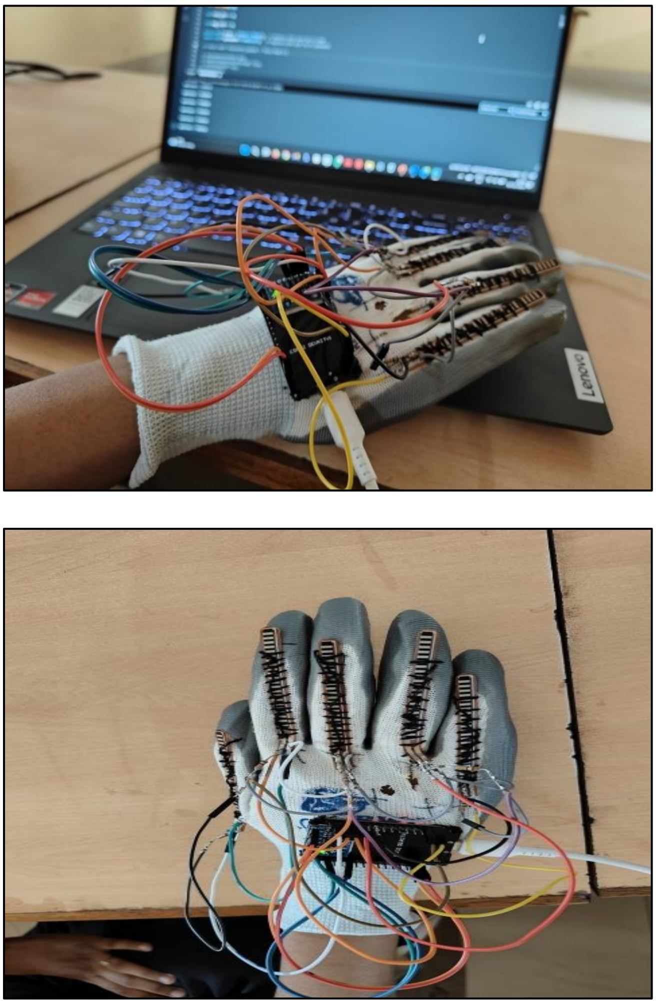


Fig 6 Hand Glove Setup

➢ *Real-Time Response*

The system consistently provided a real-time response, with the time from gesture detection to audio output being less than one second. This ensures seamless interaction and communication for users correctly identified by the system. The accuracy is calculated using equation 1.

$$\text{Accuracy} = \left(\frac{\text{Correct Recognitions}}{\text{Total Tests}}\right) \times 100$$

Correct Recognitions = 137

Total Tests = 150

Accuracy = (137/150) × 100 = 92%

➢ *Speaker Output*

The built-in speaker delivered clear, intelligible audio, validated during tests in both quiet and moderately noisy environments.

➢ *Integration with Kannada Language*

The system incorporates a Google API-powered English-to-Kannada translation feature, enabling gesture outputs in Kannada. This addition broadens accessibility, catering to regional users and enhancing the usability of the gloves in non-English-speaking contexts.

While the system has clear strengths, such as high accuracy, real-time performance, and cost-efficiency, there are still areas for future development. Expanding the vocabulary of gestures and addressing environmental sensitivities could further improve the system's robustness and usability. In comparison to market alternatives, the Smart Gloves stand out in terms of both recognition accuracy and speech clarity, making them a promising solution for gesture-based communication. With continued refinement, including battery life optimization and vocabulary expansion, the Smart Gloves have the potential to become a leading device in the field of assistive technology.

➢ *Gestures and Outputs*

Figure 7. gives the overview of the output received at the Serial Monitor as the output for the respective hand signs. A detailed view of this is given in the TABLE 2.

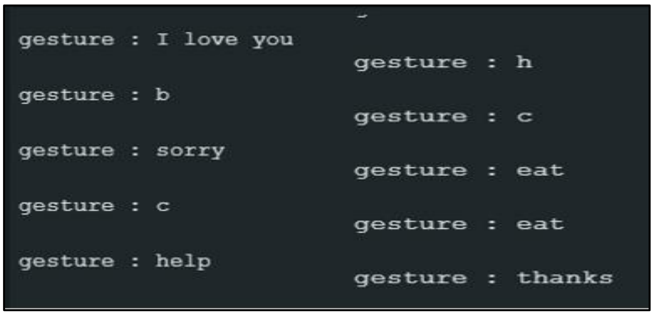


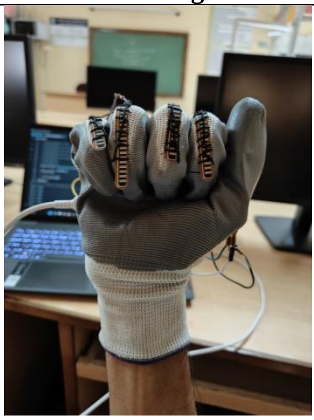
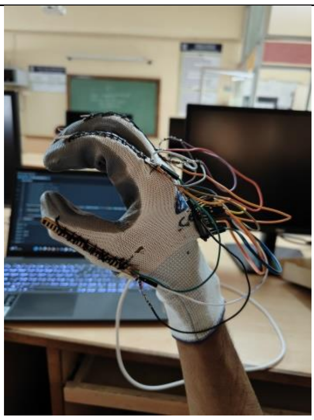
Fig 7 Serial Monitor Output

The below feature shown in Fig 8. is integrated with the help of Twilio. By detecting a "Help" gesture, the system can instantly alert designated contacts or emergency services, reducing response times during critical situations.

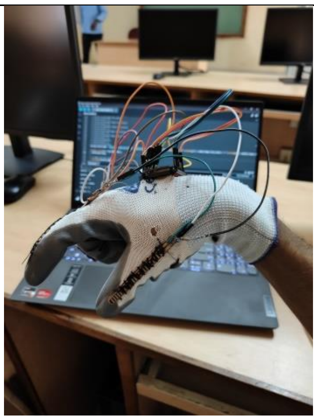
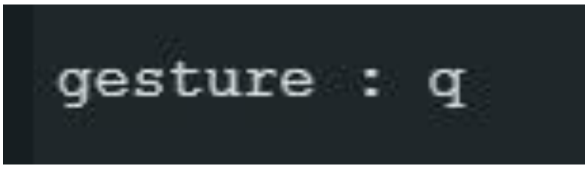
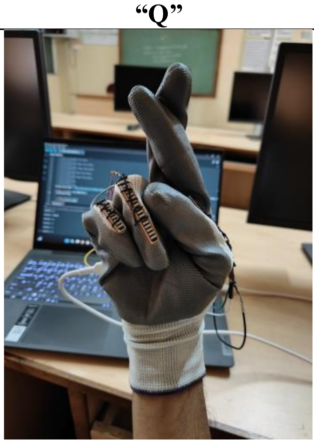
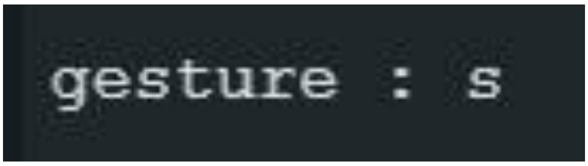
Fig 8 Twilio Message for "Help" Hand Sign

TABLE 2. below shows the hand signs and their respective outputs on the serial monitor, providing clear insights into the system's accuracy and functionality in translating gestures into recognizable outputs.

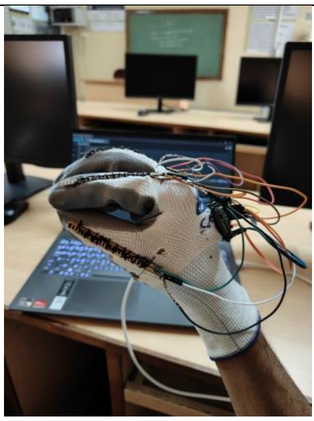
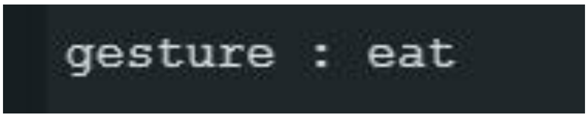
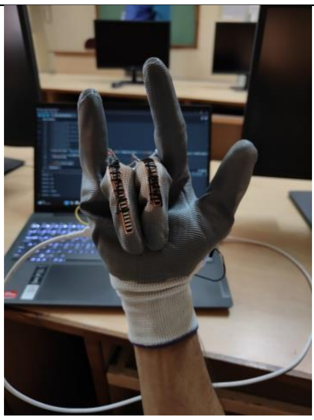
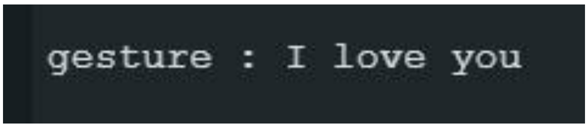
Table 2 Hand Signs and Outputs

Hand Sign	**Output**
 **"A"**	gesture : a
 **"B"**	gesture : b
 **"C"**	gesture : c
 **"D"**	gesture : d

ISSN No:-2456-2165 https://doi.org/10.5281/zenodo.14591355

 **"P"**	
 **"Q"**	
 **"R"**	
 **"S"**	

ISSN No:-2456-2165 https://doi.org/10.5281/zenodo.14591355

<table>
<tr><td>
“Thanks”</td><td>
</td></tr>
<tr><td>
“Eat”</td><td>
</td></tr>
<tr><td>
“I love you”</td><td>
</td></tr>
<tr><td>
“Help”</td><td>
</td></tr>
</table>

ISSN No:-2456-2165 https://doi.org/10.5281/zenodo.14591355

From the above TABLE 2., for the hand sign of "Help", the user gets a message on the registered mobile number. This feature is integrated with the help of Twilio.

V. CONCLUSION

The development of the Smart Gloves represents a significant step toward bridging the communication gap for individuals with speech and hearing impairments. This project successfully demonstrated the feasibility of using advanced sensor technology and real-time speech synthesis to interpret hand gestures into audible speech.

The prototype achieved its primary objectives, including a high gesture recognition accuracy of 92%, clear and understandable speech output, and affordability, making it accessible to a wide range of users. Its lightweight and ergonomic design ensured comfort and usability, while positive feedback from users affirmed its potential as a practical assistive device.

While the system showcased excellent performance in controlled and semi-real-world scenarios, some limitations, such as reduced accuracy for rapid gestures and environmental sensitivity, highlight opportunities for improvement. Addressing these challenges in future iterations, along with integrating advanced features like multi-language support, customizable gestures, and IoT connectivity, will further enhance its functionality and broaden its applications.

In conclusion, the Smart Gloves prototype not only demonstrates the promise of technology in addressing accessibility challenges but also lays a solid foundation for further research and development. With continued innovation, the project has the potential to revolutionize communication aids, empowering individuals with disabilities and fostering inclusivity in society.

ACKNOWLEDGEMENT

We express our heartfelt gratitude to Shastri School for Deaf and Dumb Children, Bangalore for the support they extended for our project. We also express our heartfelt gratitude to our Head of Department Dr. M Rajeswari and Project Guide Prof. Sudha B for their immense support and guidance during the course of this project.

REFERENCES

[1]. V. Padmanabhan, M.Sornalatha, "Hand gesture recognition and voice conversion system for dumb people," International Journal of Scientific & Engineering Research, Volume 5, Issue 5, May-2014 427 ISSN 2229-5518

[2]. Shravani Jadhav, Siddhi Mhatre, Riddhi Patil, Prof. Prachi Sorte, "Smart Hand Gloves for Deaf and Dumb People", International Journal for Research in Applied Science & Engineering Technology (IJRASET) ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538 Volume 12 Issue III Mar 2024

[3]. Prof. Niranjan L. Bhale, Sammed Pawar, Amit Pawar, Prajwal Niphade, Kamlesh Nagare, "Smart Gloves for Dumb and Deaf People", International Journal of Advanced Research in Science, Communication and Technology (IJARSCT) ISSN (Online) 2581-9429 Volume 3, Issue 7, May 2023

[4]. Sakshi Wankhade, Vishal Wadekar, Ajinath Algude, Nagarjun Jadhav, Dr. Priya Gokhale, "SMART HAND GLOVES FOR DEAF AND DUMB", International Research Journal of Modernization in Engineering Technology and Science e-ISSN: 2582-5208 Volume:05/Issue:05/May-2023 Impact Factor-7.868

[5]. Shastri School For Deaf and Dumb Children, 2nd Cross Rd, Shankar Nagar, Mahalakshmi Layout, Bengaluru, Karnataka-560096