

Analysis And Fabrication of Gesture Recognition Speech System

¹ManavNeemkar, ²Sri Harshini Ramini, ³Asker Abbas, ⁴V. Lydia Roselyn,
⁵Dr.R.Uday Kumar

^{1,2,3,4}B.Tech (MCT) Students, Dept. of Mechanical Engineering,
Mahatma Gandhi Institute of Technology, Gandipet, Hyderabad.

⁵Associate Professor, Dept. of Mechanical Engineering,
Mahatma Gandhi Institute of Technology, Gandipet, Hyderabad.

Corresponding author

ABSTRACT

Individuals with hearing and speech impairments rely heavily on sign language as their primary mode of communication, yet it remains unfamiliar to a large portion of the population. This lack of widespread understanding creates significant barriers in everyday interactions, particularly in environments such as classrooms, hospitals, workplaces, and public service areas. As a result, individuals are often forced to depend on interpreters or family members, which not only limits their independence but also affects privacy and spontaneous communication. Existing technological solutions attempt to bridge this gap but are often constrained by limitations such as high cost, dependence on camera-based systems, sensitivity to lighting and background conditions, and concerns related to continuous video capture. These factors make such systems impractical for real-world, daily use. To overcome these challenges, the proposed system introduces a lightweight, wearable glove designed to translate hand gestures into audible speech in real time. By adopting a sensor-based approach instead of vision-based methods, the system ensures consistent performance regardless of environmental conditions while maintaining user privacy. The design emphasizes portability, ease of operation, low power consumption, and affordability, making it accessible to a wider range of users and suitable for continuous everyday use. The system is built around a dual-glove configuration that integrates flex sensors and ADXL335 accelerometers to accurately capture finger bending, hand orientation, and motion dynamics. A total of ten flex sensors and motion sensors work together to generate detailed gesture data, which is processed using a K-Nearest Neighbours (KNN) algorithm implemented on the ESP32 microcontroller. This approach enables real-time gesture recognition with minimal computational requirements, making it highly suitable for embedded systems. Once a gesture is recognized, the ESP32 transmits the corresponding text via Bluetooth to a mobile application, where a text-to-speech engine converts it into audible output, allowing seamless communication with non-sign language users. The current prototype supports 10–12 Indian Sign Language gestures, focusing on demonstrating system feasibility, reliability, and real-time performance while maintaining a low cost. The modular and scalable architecture allows for future enhancements, including expanding the gesture vocabulary to over 100 signs, integrating advanced deep learning models such as 1D CNN and LSTM for improved accuracy and dynamic gesture recognition, and enabling fully offline operation. Overall, the system provides a robust, cost-effective, and user-friendly foundation for assistive communication technology, with strong potential for further development and real-world deployment.

Keywords: Gesture, Accelerometers, K-Nearest Neighbours, LSTM.

Date of Submission: 05-05-2026

Date of acceptance: 16-05-2026

I. INTRODUCTION

The system aims to create a fully portable, real-time assistive communication system using a dual-glove configuration. The system uses flex sensors to measure finger bending and IMU sensors to capture hand orientation and motion. These sensor readings are processed and classified by a lightweight K-nearest neighbours deep learning model optimized for deployment on an ESP32 microcontroller(1-4). By performing all signal processing and classification locally, the system achieves high responsiveness and independence from internet connectivity. Recognized gestures are then converted into synthesized speech, enabling seamless communication between sign language users and non-signers. This approach presents a practical, privacy-

friendly, and cost-effective alternative to traditional vision-based systems. The motivation for developing this system stems from the limitations of existing commercial solutions, which are often expensive, non-portable, or restricted to small vocabularies. Commercial gesture-recognition gloves typically cost upwards of ₹30,000 to ₹1,00,000 and require specialized software or proprietary hardware. Such high costs make them inaccessible to a large portion of users, especially in developing countries. Therefore, designing an affordable system costing approximately ₹6,000 holds significant societal importance (5-8). By reducing cost barriers, the system aims to empower students, patients, and professionals in the speech- and hearing-impaired community, enabling them to communicate confidently and effectively in everyday situations. Indian Sign Language (ISL), used by more than 7 million individuals across India, has limited technological support compared to American or British Sign Language. Research and development focused on ISL remain relatively sparse, creating an urgent need for tools that cater specifically to the Indian context (9-10). ISL consists of distinct gestures representing alphabets, words, and expressions, many involving both hands for accurate representation. The dual-glove approach in this project is designed to address the complexities of two-handed gestures, which are essential for authentic ISL recognition. Through the integration of flex sensors and IMUs on both gloves, the system captures a rich set of motion features necessary for accurate classification across a diverse vocabulary. The machine learning approach used in this project is based on the K-Nearest Neighbors (KNN) algorithm, which classifies gestures by comparing real-time sensor data with previously recorded samples. The system utilizes data from flex sensors and ADXL336 accelerometers to capture finger movements and hand orientation. KNN is chosen for its simplicity, low computational requirements, and suitability for rapid prototyping without extensive training (11-14). The processed data is handled by the ESP32 microcontroller, enabling real-time gesture recognition and wireless transmission. While the current implementation supports a limited set of gestures, the system can be further enhanced by integrating advanced models such as CNN-LSTM and optimization techniques for improved accuracy and scalability on embedded platforms.

II. METHODOLOGY

A flex sensor or bend sensor is a low-cost and easy-to-use sensor specifically designed to measure the amount of deflection or bending. It became popular in the 90s due to its use in the Nintendo Power Glove as a gaming interface. Since then, people have been using it as a goniometer to determine joint movement, a door sensor, a bumper switch for wall detection, or a pressure sensor on robotic grippers. A flex sensor is basically a variable resistor that varies in resistance upon bending. Since the resistance is directly proportional to the amount of bending, it is often called a Flexible Potentiometer. Flex sensors are generally available in two sizes: one is 2.2" (6.688 cm) long and another is 4.6" (11.43 cm) long. A flex sensor consists of a phenolic resin substrate with conductive ink deposited and is shown in fig. A segmented conductor is placed on top to form a flexible potentiometer in which resistance changes upon deflection. Flex sensors are designed to flex in only one direction – away from ink. Bending the sensor in another direction may damage it. Also take care not to bend the sensor close to the base, because the bottom of the sensor (where the pins are crimped on) is very fragile and can break when bent over. The conductive ink printed on the sensor acts as a resistor. When the sensor is straight, this resistance is about 26k as shown in fig. When the sensor is bent, the conductive layer is stretched, resulting in a reduced cross section (imagine stretching a rubber band). This reduced cross section results in an increased resistance. At 90° angle, this resistance is about 100KΩ. When the sensor is straightened again, the resistance returns to its original value. By measuring the resistance, you can determine how much the sensor is bent.

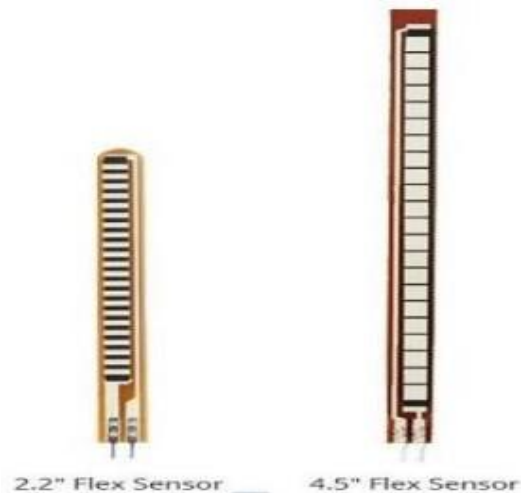


Fig.1 Flex sensors

The easiest way to read the flex sensor is to connect it with a fixed value resistor (usually 47kΩ) to create a voltage divider. To do this you connect one end of the sensor to Power and the other to a pull-down resistor as shown below. Then the point between the fixed value pull-down resistor and the flex sensor is connected to the ADC input of an Arduino. This way you can create a variable voltage output, which can be read by Arduino's ADC input. Note that the output voltage you measure is the voltage drop across the pull-down resistor, not across the flex sensor. In the shown configuration, the output voltage decreases with increasing bend radius. In recent years, advancements in Micro-Electro-Mechanical Systems (MEMS) technology have enabled the development of compact and low-power motion sensors for a wide range of applications such as gesture recognition, gaming, wearable devices, and navigation systems. Among these, accelerometers play a crucial role in detecting motion and orientation. The ADXL336 is a widely used MEMS-based 3-axis analog accelerometer that measures acceleration along the X, Y, and Z axes. Unlike combined IMU modules that include both gyroscopes and accelerometers, the ADXL336 focuses solely on acceleration sensing, making it simple, cost-effective, and suitable for lightweight embedded applications. It can measure static acceleration (such as gravity for tilt sensing) as well as dynamic acceleration caused by motion, vibration, or shock. This makes it particularly useful for detecting hand orientation and movement patterns in gesture-based systems. The ADXL336 operates over a typical voltage range of 1.8V to 3.6V, with a low power consumption of approximately 360μA, making it ideal for battery-powered wearable devices. It provides analog voltage outputs proportional to acceleration on each axis, which can be read using the ADC pins of microcontrollers like the ESP32. The sensor has a measurement range of approximately ±3g, which is sufficient for most human motion tracking applications. Due to its compact size, low power requirement, and ease of interfacing, the ADXL336 is well-suited for integration into wearable systems such as smart gloves. In this project, it is used to capture hand tilt and motion dynamics, which, when combined with flex sensor data, helps improve the accuracy of gesture recognition.

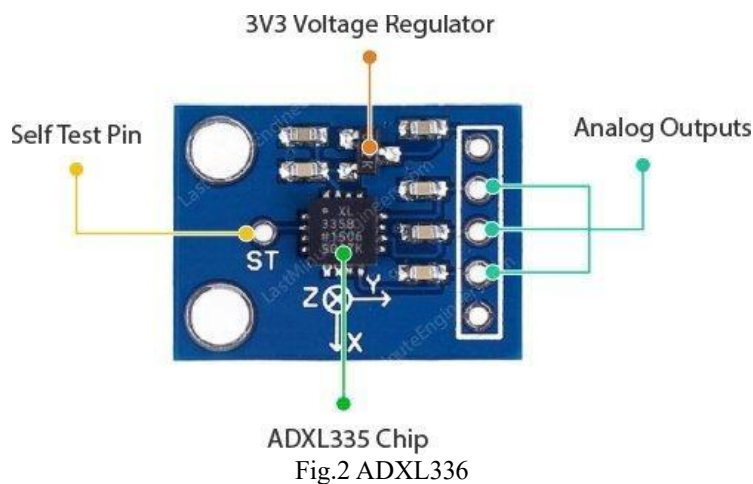


Fig.2 ADXL336

2.1 Working principle of system

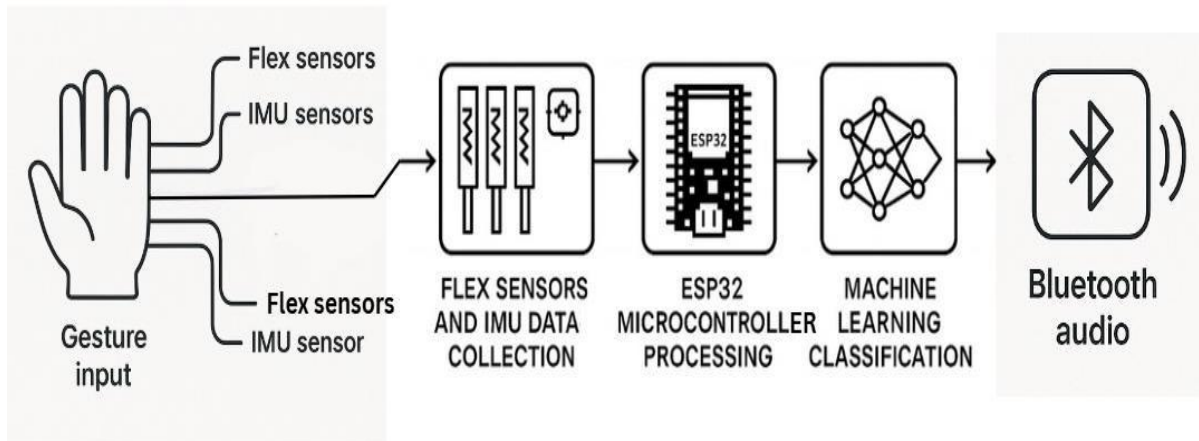


Fig.3 Working principle of system

The working principle of the sign language to voice conversion glove is based on the integration of multiple sensing mechanisms, embedded processing, deterministic gesture classification, and wireless communication to convert Indian Sign Language gestures into meaningful spoken output. The glove functions as an intelligent wearable device that continuously monitors finger bending and hand movement, interprets this information using predefined logic, and transmits the interpreted gesture to a Bluetooth-enabled audio device. The entire workflow is designed to operate in real time, ensuring that hand movements are translated into spoken English with minimal delay. This section explains the complete operational sequence of the glove beginning from sensor activation to gesture detection and ending with wireless transmission of the recognized sign. The system begins functioning as soon as the microcontroller is powered on. The initialization phase activates all on-board communication and sensing modules, including the serial interface for debugging, the HC06 Bluetooth module for data transmission, and the ADXL336 inertial measurement unit that monitors motion.

The microcontroller then configures all five flex sensor pins as analog inputs so that bending data can be acquired continuously. A connection test is performed on the IMU to ensure that valid accelerometer and gyroscope data will be available during operation. If this test is unsuccessful, the glove enters an idle state and alerts the user through a serial message. Once initialization is complete, the glove enters its continuous monitoring mode, where the main loop executes repeatedly at a controlled sampling rate. During each iteration of the loop, the glove first collects real time bending information from all five flex sensors. These sensors act as variable resistors whose output voltage changes depending on the amount of curvature in each finger. The thumb, index, middle, ring, and little fingers each have their own sensor, allowing the system to capture hand configurations with high granularity. The analog readings from the sensors are stored in a dedicated array that is updated several times per second. This enables the glove to track both static hand shapes and transitions between gestures. Finger bending information forms the core of many sign language gestures, making these flex sensors primary contributors to gesture recognition.

Alongside flex sensor readings, the glove simultaneously collects motion and orientation data from the ADXL 336 IMU. This sensor provides raw accelerometer values indicating linear movement along three axes and gyroscope values indicating rotational velocity around three axes. The raw readings are converted into standard units such as g-forces for acceleration and degrees-per-second for angular rotation. These processed motion values allow the glove to detect wrist movements that accompany certain signs. For example, gestures such as “YES” and “NO” involve distinct dynamic wrist motions that cannot be recognized through finger bending alone. The IMU therefore enhances the glove’s ability to distinguish between gestures that would otherwise appear visually similar or identical in static form.

Once flex and IMU data are collected, the gesture recognition engine evaluates the sensor patterns to determine whether the user is performing any of the defined signs. In this first-stage prototype, the engine uses threshold-based logic, where each gesture is associated with a specific combination of bending levels and motion characteristics. A gesture is recognized when the real-time sensor values match the predefined criteria for that gesture. For example, the sign for “HELLO” is identified when the thumb, index, and middle fingers are extended beyond their thresholds while the ring and little fingers remain bent. The sign for “THANK YOU” requires that all fingers be extended and supported by a slight rotational motion of the wrist. Similar logic is applied to detect “YES,” which includes a nodding-like up-and-down hand motion captured through gyroscope data, and “NO,” which involves a lateral side-to-side

motion. The gesture "HELP" is recognized when the palm is open and an upward acceleration spike is detected through the accelerometer. This threshold-based approach provides transparent and predictable classification behavior, making it ideal for an early-stage prototype. Once a gesture is identified, the system compares it with the previously detected gesture to avoid redundant or repeated outputs. If the gesture has changed, the glove constructs a formatted message that includes both the gesture ID and the corresponding word mapped from the gesture library. This message is sent wirelessly to the paired Bluetooth module using a predefined communication protocol. The Bluetooth device handles the transmission in a simple and reliable manner, ensuring that each recognized sign results in a corresponding audio output. The external speaker, mobile device, or audio unit that receives the transmitted message generates the final speech output, effectively converting the user's sign language gesture into spoken English. The real-time nature of this process ensures that communication remains fluid and natural for both the signer and the listener. Throughout operation, the glove also produces diagnostic output through the serial monitor. This debug channel prints flex sensor readings, IMU values, and the currently recognized gesture, allowing continuous monitoring and fine-tuning of thresholds. This feature is crucial for calibrating the glove for different users, as finger strength, hand size, and natural movement patterns vary. The debugging output helps refine accuracy and ensure that each gesture is detected consistently across multiple trials. The working principle of the first-stage prototype demonstrates how sensor-driven logic can be used effectively to recognize a small vocabulary of sign language gestures. By establishing a reliable pipeline from sensing to classification and wireless communication, the system validates the feasibility of using a low-cost, embedded solution for gesture recognition. This initial functionality forms a strong foundation for advancing toward more sophisticated versions of the glove. Subsequent stages will incorporate machine learning models to replace the threshold-based classifier, expand the vocabulary to hundreds of gestures, and integrate dual-hand sensing for full Indian Sign Language support. Nonetheless, the underlying working principle remains consistent: capturing human intent through hand movement, interpreting it through embedded processing, and enabling communication through real-time speech output.

III. RESULTS AND DISCUSSION

The code begins by including the required libraries for the ESP32, analog input handling for flex sensors, ADXL336 accelerometer readings, and Bluetooth communication. Global definitions are set up for pin configurations, sensor variables, and storage arrays for gesture data. In the setup section, serial communication is initialized for debugging, ADC pins for flex sensors are configured, and Bluetooth is initialized to connect with a mobile application. In the main loop, the system continuously reads sensor data from all flex sensors and accelerometers at a fixed sampling rate. The collected data is processed into feature values and compared with stored gesture samples using the K-Nearest Neighbours (KNN) algorithm to determine the closest matching gesture. A simple validation check is applied to ensure stability and avoid incorrect predictions. Once a gesture is recognized, it is mapped to a predefined word or phrase stored in memory. The ESP32 then sends this text via Bluetooth to a mobile application, where a text-to-speech engine converts it into audible output. The program also includes basic control logic to prevent repeated triggering when a gesture is held for too long and may provide simple feedback such as an LED indication when a valid gesture is detected. The code structure is organized into helper functions such as reading sensor data, processing inputs, classifying gestures, mapping outputs, and transmitting results, keeping the main loops simple and efficient for real-time operation. The initial stage of development focuses on constructing and validating a simplified prototype of the sign language recognition glove. This first-stage version uses a single glove equipped with five flex sensors and one inertial measurement unit and is designed to recognize only five basic gestures. This controlled and minimal setup enables reliable experimentation, stable data collection, and early verification of the system's sensing and classification logic before moving into more advanced multi-glove, machine learning, and vocabulary-expansion phases. The primary objective of the first-stage prototype is to establish a fully functional pipeline from sensing to gesture detection and Bluetooth communication while maintaining low complexity, predictable behavior, and ease of debugging. At this stage, the glove is embedded with five flex sensors, each attached to one finger: thumb, index, middle, ring, and little finger. These sensors provide analog values that change according to the curvature of each finger, allowing the system to capture distinct finger postures associated with basic Indian Sign Language gestures. In addition to the flex sensors, the glove incorporates a single MPU6050 module that provides acceleration and rotational motion readings along three axes each. Although the initial prototype relies mainly on flex data for classification, the IMU still contributes essential movement information for gestures that require a characteristic wrist motion, such as the signs for "YES," "NO," or "THANK YOU."

Using only one glove drastically simplifies wiring, calibration, and interpretation while helping to isolate and refine the core sensing mechanism that will later support more sophisticated gestures.

The performance of the KNN-based classification approach was also validated during this phase, demonstrating reliable real-time gesture recognition with minimal computational load on the ESP32. The Bluetooth communication pipeline was tested extensively to ensure smooth and continuous data transmission between the glove and the mobile device, with negligible delay affecting user interaction. These evaluations helped identify minor issues such as sensor noise, threshold tuning requirements, and gesture overlap, which were addressed through iterative adjustments. Overall, the second-stage prototype establishes a stable and well-tested platform, enabling future enhancements such as expanding the gesture vocabulary, improving classification accuracy using advanced models like CNN-LSTM, and implementing coordinated dual-glove interaction for more complex sign language recognition.

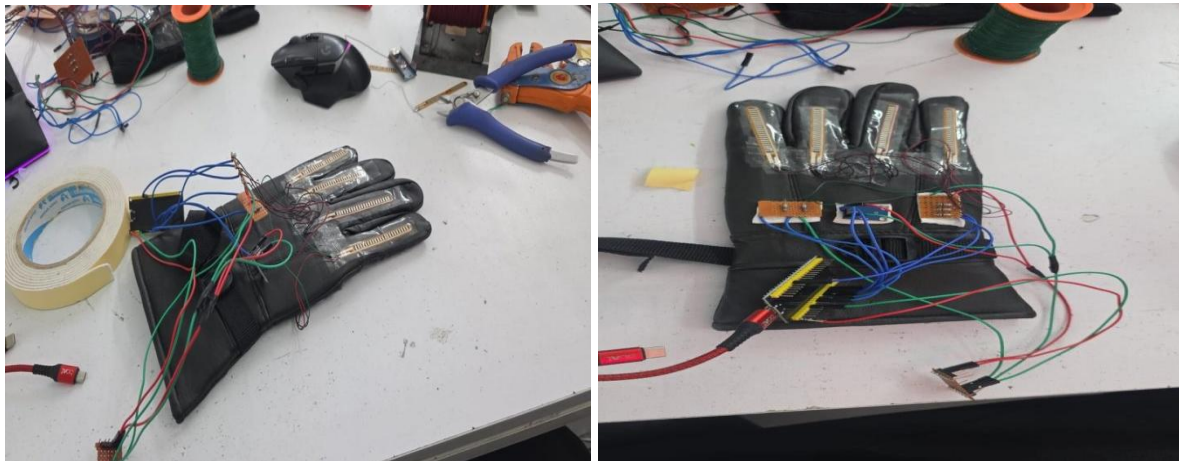


Fig.4 Glove Fabrication

3.1 OK sign using ISL

The “OK” sign in Indian Sign Language (ISL) is formed using a distinct finger configuration and hand posture:

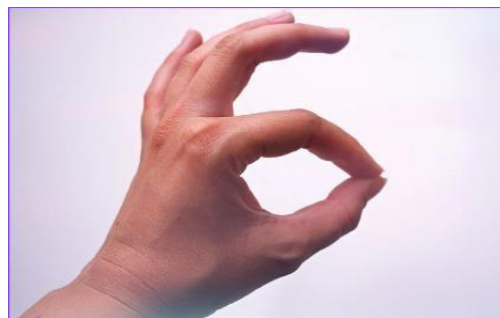


Fig.5 OK Hand sign

- The thumb and index finger are brought together to form a closed circular shape, resembling the letter “O”.
- The remaining three fingers (middle, ring, and little finger) are kept extended and slightly spread apart, pointing upward.
- The palm generally faces forward or slightly outward, depending on comfort and context.
- The hand is held steady at about chest level, and the gesture is typically static (no major movement involved).

This combination of a circular finger shape with extended fingers clearly represents the meaning “OK” or “all good” in ISL.

3.2 YESsignusingISL

The“YES”signinIndianSignLanguage(ISL)istypicallyexpressedthroughasimple hand gesture combined with slight movement:

- Thehandisfirst formedintoaclosedfist,withallfingerscurledinwardandthe thumb resting over the fingers.
- Thefististhenmovedupanddownslightly,similartoanoddingmotion.
- The movement is usually performed at about chest level and repeated once or twice for clarity.
- The gesture is dynamic, meaning the motion plays an important role in conveying the meaning.

This up-and-down movement of the closed fist represents affirmation, similar to nodding the head, and is commonly understood as “yes” in ISL.



Fig.6 YES Hand Sign

3.3 YOU sign using ISL

The“YOU”signis performed by extending the index finger forward while the remaining fingers are folded (bent). The hand points outward toward the person being addressed.

From a mechanical perspective:

- Index finger → straight (no bending)
- Middle, ring, pinky → fully bent (~90°)
- Thumb → partially bent

This creates a condition where multiple flex sensors experience high bending



Fig.7 You Hand Sign

IV. CONCLUSIONS

Communication is an essential aspect of human life and serves as the basis for social, emotional, and intellectual development. For individuals with hearing or speech impairments, sign language is the primary medium of expression and interaction. Sign languages such as Indian Sign Language allow millions of people to communicate their thoughts and feelings with clarity. However, a significant communication barrier still exists between sign language users and people who do not understand these gestures. This barrier often creates challenges in education, public services, healthcare, and daily social interaction. The motivation for this work arises from the need to create an inclusive environment where individuals with communication challenges can interact smoothly with the broader community. The development of the single-glove sign language to voice conversion system marks an important first step toward creating an accessible, low-cost, and user-friendly communication aid for individuals with hearing and speech impairments. This prototype successfully demonstrates the feasibility of using a combination of flex sensors and inertial measurement units to capture finger articulation and hand motion patterns, which together form the fundamental components of Indian Sign Language gestures. By implementing threshold-based logic and a streamlined decision engine, the system is able to reliably distinguish

between five essential gestures and transmit their corresponding meanings through a Bluetooth interface to generate spoken output. The complete pipelinespanning sensor acquisition, gesture interpretation, message formatting, and wireless transmissionoperates in real time and validates the core functional concept of gesture-to-speech conversion on an embedded platform.

Beyond verifying the technical foundation, the prototype offers valuable insights into the practical challenges associated with gesture recognition. It highlights the need for careful sensor calibration, stable mounting of components, and consistent gesture patterns across users. The debugging and diagnostic capabilities incorporated into the system further support refinement by enabling precise monitoring of sensor behavior and recognition accuracy. These learnings establish a strong groundwork for future iterations of the project, where the focus will shift toward replacing threshold logic with machine learning models, improving reliability, expanding vocabulary, and integrating dual-hand sensing for complete ISL coverage.Overall, the first-stage glove demonstrates a functional, compact, and efficient assistive device that transforms gestures into meaningful speech output. While still in its early form, the system proves that sensor-driven recognition can serve as a viable bridge between signers and non-signers, promoting inclusivity and making everyday interactions more accessible. The success of this prototype sets the direction for more advanced stages of development, ultimately contributing toward a scalable and impactful communication solution.Future versions of the system can incorporate larger and more sophisticated deep learning architectures capable of recognizing continuous sign sequences, facial expressions, hand orientation, and movement trajectories.This would allow the glove to interpret fullsentences rather than isolated words, greatly improving communication fluidity.The system can be expanded to support hundreds of gestures, full two-handed ISL grammar, and contextual modifiers. With continued training and dataset collection, the glove could eventually translate full phrases, sentences, and conversational flow in real time.Future iterations may utilize flexible PCBs, printed sensors, and integrated IMUs woven directlyinto fabric.Thiswouldreduceweight,improveergonomics,andcreateamorenaturaluser experience, allowing the glove to resemble a regular piece of clothing rather than an electronic device.

REFERENCES

- [1]. Natarajan Vijayaraj, Mohan Nalini, Kinol A Mary, R M Bommi, Smart Glove for Impaired People to Convert Sign into Voice with Text, Proceedings of ICCEBS, 2023
- [2]. Zhou Z, Chen K, Li X and co authors, Sign to Speech Translation Using Machine Learning Assisted Stretchable Sensor Arrays, Nature Electronics, volume 3,2020
- [3]. Khan R U, Khattak H, Wong W S, AlSalman H, Mosleh M A A, Mizanur Rahman S M, Intelligent Malaysian Sign Language Translation System Using Convolutional Based Attention Module with Residual Network, Computational Intelligence and Neuroscience, volume 2021, 2021
- [4]. Rami Aldahir, Ronald R Grau, Using Convolutional Neural Networks for Visual Sign Language Recognition Towards a System that Provides Instant Feedback to Learners of Sign Language, Proceedings of the International Web for All Conference, 2024
- [5]. K Shenoy, T Dastane, V Rao, D Vyavaharkar, Real Time Indian Sign Language Recognition, Proceedings of the International Conference on Computing Communication and Networking Technologies, specified, 2018
- [6]. Swaroop Gudi, Chinmay Inamdar, Yash Divate, Sunil Tayde, Sign Language Detection Using Gloves, International Journal for Research in Applied Science and Engineering Technology.
- [7]. Hind Bitar, Ohoud Alzamazami, Dimah Alahmadi, Amal Barsheed, Amal Alghamdi, Hanadi Almshjary, Intelligent Gloves An Information Technology Intervention for Deaf Mute People, Journal of Intelligent Systems, volume 32, issue not specified, 2023
- [8]. Urav Dalal, Aasmi Thadhani, Mahek Upadhye, Shreya Shah, Meera Narvekar, Nilesh Patil, Deep Learning Enabled Smart Glove for Real Time Sign Language Translation, Journal of Electrical Systems, volume 20, issue 10 special issue, 2024
- [9]. P Das, R De, S Paul, M Chowdhury, B Neogi, Analytical Study and Overview on Glove Based Indian Sign Language Interpretation Technique, Michael Faraday IET International Summit Proceedings, 2015
- [10]. Authors from Sangmyung University, Application of Wearable Gloves for Assisted Learning of Sign Language Using Artificial Neural Networks, Processes, volume 11, 2023
- [11]. Authors from Punjab Engineering College, ISL Recognition System Using Integrated Mobile Net and Transfer Learning Method, Expert Systems with Applications, volume 221, 2023
- [12]. Abhishek Tandon, Amit Saxena, Keshav Mehrotra, Khushboo Kashyap, Harmeet Kaur, A Review Paper on Smart Glove Converts Indian Sign Language into Text and Speech, International Journal for Scientific Research and Development, volume 4, issue 8, November 2016.
- [13]. Rahul Bhide, Varun Gujarathi, Sumit Bhosale, Nishika Agrawal, Sign Language Converter Using Electronic Glove, Journal ofEmerging Technologies and Innovative Research, volume 12, issue 9, September 2023
- [14]. Ketan Tajne, Swastik Jha, Swaraj Patil, Swayam Kanoje, Mangesh Tale, Abhjeet Takale, Rahul Waikar, Design and Development of Sensor Based Glove for Speech Impaired People, International Journal for Research in Applied Science and Engineering Technology, June 2020.