RESEARCH ARTICLE OPEN ACCESS

Wearable Device for Real Time Indian Sign Language to Speech Translation using Machine Learning Techniques

Dinesh. P [a], Vanmathi. P [b], Kamalakannan. R [c], Kaviyasri. G [d], Govindaraj. S [e]

[a] *(Department of Medical Electronics, SRM Valliammai Engineering College, Chennai, India)*
[b] *(Department of Medical Electronics, SRM Valliammai Engineering College, Chennai, India)*
[c] *(Department of Medical Electronics, SRM Valliammai Engineering College, Chennai, India)*
[d] *(Department of Medical Electronics, SRM Valliammai Engineering College, Chennai, India)*
[e] *(Department of Medical Electronics, SRM Valliammai Engineering College, Chennai, India)*

Date of Submission: 02-04-2025 Date of acceptance: 12-04-2025

ABSTRACT
People with speech impairments often find it difficult to communicate with people who are not familiar with the sign language which will limit their ability to engage in daily activities. To overcome this challenge a novel assistive device has been developed to convert Indian sign language into audible speech in real time. This wearable glove has ten flex sensors that are integrated on each finger and an MPU6050 gyroscope and accelerometer that together record the hand gestures and spatial orientation. By digitizing the analog signals from these sensors by Arduino micro controller and sends the data to a Raspberry Pi. A machine learning model is implemented on the raspberry pi to classify the hand gestures and map them to predefined speech outputs. These speech outputs are played through a speaker, facilitating effective communication with people who are not familiar with sign language. This real time Indian sign language to speech system provides portable and user-friendly solution that improves communication for speech impaired individuals.

Keywords-Accelerometer, Arduino Mega, Flex Sensors, Gesture Classification, Indian Sign Language (ISL), Random Forest Algorithm, Raspberry Pi, Sign Language to Speech

I. INTRODUCTION

Communication is essential to human interaction, learning, and social interaction; however, the World Health Organization (WHO) estimates that 70 million people all over the world face speech difficulties, making this the most vital aspect in coping with daily challenges. Speech impairments may be due to congenitally disabling conditions or any other illnesses of the nervous system and trauma or even age that obstructs the ability to speak. This affects not only personal expression but extends to involvement in social, educational and professional activities.

For instance, among the estimated 70 million people around the world that use sign language as a primary means of communication. Most of them face challenges in daily interactions because they must often communicate with people unfamiliar with this visual language. Research done by the National Institute on Deafness and Other Communication Disorders (NIDCD) has established nearly 14 million individuals with severe speech impairments (about20% of this group) report frequent misunderstandings in their efforts to communicate.

This system bridges the communication gap faced by speech-impaired individuals by converting Indian Sign Language into audible speech in real time. It helps speech impaired individuals to express themselves clearly and effectively, promoting their independence and self- confidence. By enabling seamless communication, it helps break down barriers faced by speech impaired individuals and makes the way for a world where everyone can interact freely and meaningfully.

II. LITERATURE SURVEY

In Ref [1], a system was designed to overcome barriers that the speech and hearing-impaired individuals encounter with several sign languages. It uses CNN and deep learning to translate between Indian Sign Language, American Sign Language, and various Indian regional languages. Although the platform is quite inclusive, it is very computationally expensive, requires large amounts of data to produce accurate translations, is unable to cope with unusual gestures and struggles with real-time processing.

The paper [2] presents a study which tackled the communication problems of the speech and hearing impaired by translating gestures into text. While the system enables interaction, it does not work well with dynamic gestures, differences among users, and external factors such as lighting and background noise. Moreover, the need for unequally controlled settings considerably restricts its usefulness in actual situations.

In Ref [3], a communication device in the form of a Sign-to-Speech glove was designed to eliminate communication barriers, especially among disabled individuals. The glove incorporates flex sensors for detecting finger movements and an Arduino microcontroller to process the information. While effective, it only translates the alphabet and lacks support for advanced or multiple-hand gestures. Additionally, it requires recalibration after every period of usage, making it less flexible and less effective.

In Ref [4], a real-time system was built that can recognize sign language using a live webcam stream through hand landmarks provided by the Media Pipe library. The extracted landmark data is used to train and classify the sign using a Random Forest algorithm. While this is a capable approach integrating computer vision and machine learning, the system still faces several challenges including occlusions, rapid gesture movements, lighting issues, as well as dependence on high-performance cameras.

In Ref [5], a speech-generating glove was designed to convert ISL gestures to speech. This system uses dedicated pushbuttons for input of specific gestures, the input is processed by Arduino Uno to output text which is further synthesized into speech using a text to speech software written in Python. The use of pushbuttons restricts adaptability, thereby making the system unsuitable for continuous gestures. It also lacks expansion to support more than one language and would be difficult for users with limited ability in using their hands.

In Ref [6], a real time ASL identification is done using CNN system in a dedicated Android application along with reliable hand gesturing to voice conversion. Dynamic gestures pose a challenge as they are still effective for only static gestures hence the system needs a high computational power. The system focuses on ASL only which limits its scalability and adaptability to other sign languages.

A smart glove system is designed as a speech enhancement device that recognizes sign language for deaf and mute people. It uses static gestures, but dynamic inputs and scalability are not supported. Comfort and prolonged wearability are also an issue. However, it enables some form of basic communication [7].

In Ref [8], A system was developed for the interpretation of sign language into both text and speech using deep learning-based LSTM networks for classifying hand gestures. This adopts the use of computer vision techniques and edge- cutting algorithms to identify hand gestures have been trained on a hand landmark pose dataset with key points for face, hand, and pose landmarks.

The system manages to have 90% performance efficiency thus solving the problem of speaking individuals trying to communicate with non-speaking individuals but still some problems are left with dynamic gestures and time efficient processing.

In Ref [9], A real-time system uses CNNs to convert International Sign Language (ISL) to speech has been developed with 97.85% accuracy. This initiates communication between sign language users and non- signers, it struggles in translating intricate gestures, varying light conditions. Additionally, the arrangement is highly computationally intensive and does not support various sign languages.

In Ref [10], a sensor-based system using five flex sensors on each finger converts sign language into speech, offering 32 options. It reduces production costs by eliminating the need for additional sensors and is more accurate than vision-based systems. However, the system may face limitations in recognizing complex gestures.

III. PROPOSED SYSTEM

The objective of the proposed system is to develop a wearable assistive device that converts Indian Sign Language (ISL) to an audible speech so that those with speech impairments can communicate efficiently. This device aims to provide a pocket-friendly way of overcoming the communication problem faced by speech impaired people and enables smooth communication with non-sign language users.

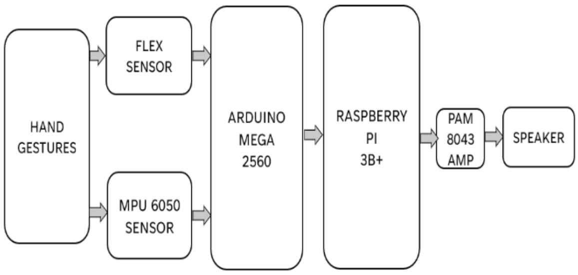

Fig. 1. Block Diagram of the sign language-to-speech translation system

The system shown in fig. 1 is glove-based design and the flex sensors are places over the fingers it senses the bending of the fingers. The sensors are an essential and an integral part of recognizing the Indian sign language through the hand shapes. There is gyroscope and accelerometer MPU6050 sensor module as well, that will help to find out the orientation and the motion of the hand. Using the proposed combination of sensors, the gestures are detected and recognized correctly.

Processing is done by an Arduino Mega and a Raspberry Pi. Arduino Mega mainly works as Analog to Digital Converter which reads the signals from flex sensors and from MPU6050 module and converts it to digital data. A Raspberry Pi acts as central processing unit, receiving data from an Arduino mega.

A Random Forest classifier is a machine learning model that runs on the Raspberry Pi for the accurate recognition and classification of the gestures. This method significantly improves adaptability and reduces errors compared to threshold-based systems, providing stable and reliable performance across diverse user populations and environments.

The audio output of that sign is played via the raspberry pi whenever a sign is detected. The system stores files for all pre-recorded speech segments and plays them out on an external speaker for real-time gesture-to-speech translation. The system provides a way for speech impaired people to communicate with each other easily and efficiently.

This really makes the whole device lightweight, user-friendly and handy for daily use. Rather than relying on sign-to-text solutions, this device directly turns gestures into speech, facilitating a more natural and functional communication mode. The technique is flexible because of machine learning method can be trained to understand even more gestures in the future, if needed.

IV. METHODOLOGY AND IMPLEMENTATION

The methodology of the project is divided into the following major phases which will ensure the instant and accurate translation of Indian Sign Language (ISL) into audible speech addressing aspects of sensor data acquisition, signal processing, gesture identification, and speech output.

In the sensor data acquisition stage, an implemented of wearable glove using ten flex sensors and an MPU6050 IMU sensor to capture finger flexion and hand motion. The bending of the sensors causes them to go from a high to a low state, which generates an input voltage for the Arduino Mega, we have connected the sensors through its A0–A9 pins and the analog voltage is digitized through the 10-bit ADC.

Data acquired from the accelerometer (Ax, Ay, Az) of the MPU6050 is used to monitor the orientation and motion of the hand which is then sent based on the I2C protocol (SCL pin 21, SDA pin 20) and will provide us with extensive data for gesture recognition.

In the signal processing part, the digitized data is sent from Arduino to Raspberry Pi through UART serial communication. The process involves calculating features of the flex sensors (F1, F10) and IMU accelerometer (Ax, Ay, Az) from the input of the sensors, while performing some noise reduction and normalization, to prepare the input for a classification process.

The Random Forest machine learning algorithm is used for gesture classification. The Random Forest algorithm is an ensemble learning method that builds multiple decision trees based on varying subsets of the dataset and combines their output by majority voting in order to minimize overfitting and enhance generalization.

Based on some of the features derived from the ISL dataset, the trained model identifies the incoming data into the ISL gestures based on the mapping of the different hand movements. This makes classification very accurate and robust while also ensuring the classification to be performed in real time.

To evaluate the accuracy of the gesture recognition model, key performance metrics such as Accuracy, Precision, Recall, and F1-score are used. Accuracy is given by:

$$\text{Accuracy} = (TP + TN) / (TP + TN + FP + FN) \quad (1)$$

where TP (True Positives) represent correctly classified gestures, TN (True Negatives) indicates correctly classified non-gesture instances, FP (False Positives) are incorrectly classified gestures, and FN (False Negatives) refer to missed gesture classifications. The Precision metric is defined as:

Precision = TP / (TP + FP) (2)

which measures the proportion of correctly predicted gestures among all classified gestures. The Recall metric is given by:

Recall = TP / (TP + FN) (3)

which evaluates the model's ability to correctly detect actual gestures. The F1-score, which provides a balance between Precision and Recall, is computed as:

F1-score = 2 × (Precision × Recall) / (Precision + Recall) (4)

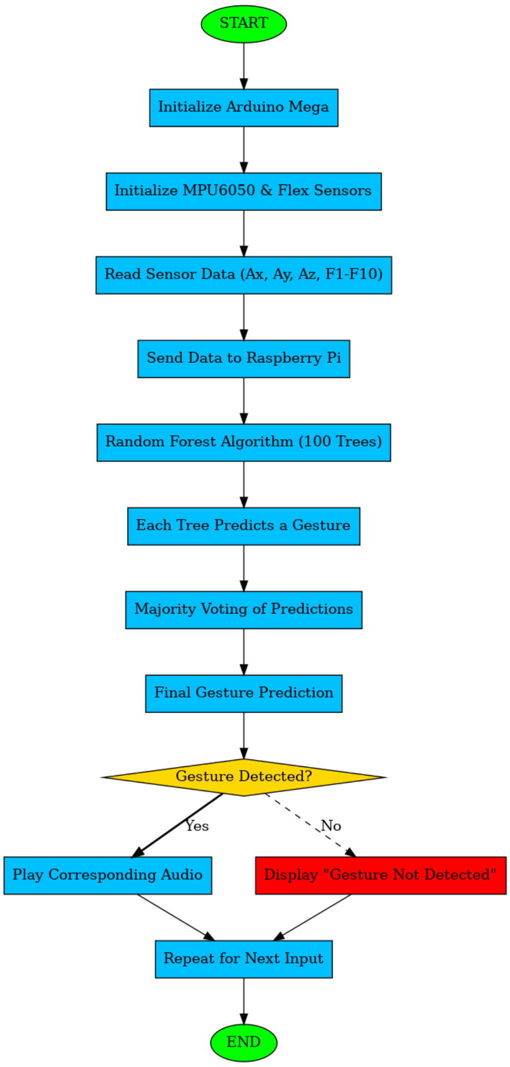

Fig. 2. Flowchart representation of working of Sign language to Speech translation system

This system works by collecting data through sensors, processes the collected data through Arduino mega, pass the data to Raspberry Pi, and then the trained Random Forest model gives classification on detected gesture. The Raspberry Pi plays the respective audio file as part of speech output generated after a gesture is recognized. The PAM8403 amplifier amplifies the audio signal and outputs it powered by a loudspeaker, providing instant feedback for the user.

The working of proposed system shown in fig 2:

1. Flex sensors, a 3-axis accelerometer, and a gyroscope collect hand movement data.
2. The Arduino Mega receives and processes the sensor data.
3. The processed data is sent to the Raspberry Pi for further analysis.
4. A Random Forest-based machine learning model predicts the gesture and triggers speech synthesis.
5. The detected gesture is converted into speech and played through a speaker.

V. RESULT AND DISCUSSIONS

The hardware implementation of the assistive device is shown in Fig. 3 which includes a glove outfitted with ten Spectra Symbol flex sensors (one on each finger) and an MPU6050 gyroscope and accelerometer, specifically designed to recognize and interpret hand gestures in real time.

Fig. 3. Hardware implementation of the sign language-to-speech translation system.

We have developed an efficient system that will convert Indian Sign Language (ISL) gestures into speech using machine learning. A Random Forest algorithm is used to classify gestures based on the real-time sensor data, achieving a high accuracy.

When a gesture is recognized, it is then mapped to the respective word which is then played as speech output. This system was tested on 10 ISL gestures with an accuracy of 97.20%.

Fig. 4. Hand sign for the gesture "HELLO"

To ensure that the system can accurately recognize different sign gestures a series of signs were performed, recorded, and analyzed. Fig.4 shows the sign for "HELLO" gesture.

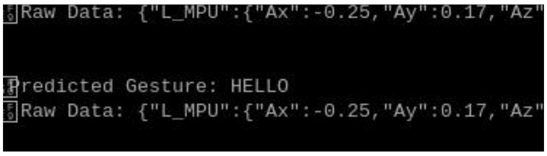

Fig. 5. Output for the gesture "HELLO"

The detected sign shown in fig. 5 is displayed in the Geany's command window and the corresponding speech output is played through the speaker. Then the system continues to monitor in real time for the next gestures.

TABLE1.PERFORMANCEOFTHEMLMODEL

Metric	Performance Score (%)
Overall Accuracy	97.2
Precision	97.5
Recall	97.6
F1-Score	97.3

The system performance metrics are shown in Table 1 indicates that the system achieved 97.2% accuracy ensuring reliability of gesture classification. A precision score of 97.5% indicates very few false positives and a recall score of 97.6% evidence of correctly identifying actual gestures. The 97.3% F1-score indicates a strong balance between precision and recall. This makes the system highly effective for real-time sign language recognition and conversion to speech with these high-performance metrics validating the robustness of the system.

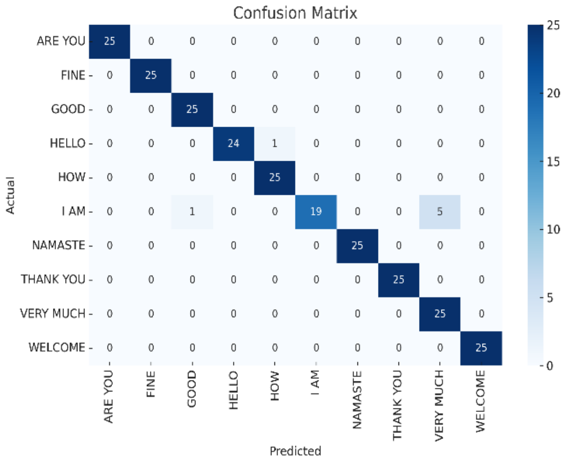

Fig. 6. Confusion Matrix for the Developed Model

The confusion matrix is shown in fig. 6 which provides a detailed visual representation of the model's classification performance. The diagonal entries show correct predictions of gestures, and off-diagonal entries indicate which gestures were confused with one another.

Out of 10 gestures, 8 were classified with 100% accuracy showing the strong predictive ability of the model. The "HELLO" gesture has only one misclassification which occurred because it looks similar to another gesture. The "I AM" gesture had some ambiguity with misclassifications with "FINE" once and "VERY MUCH" five times, indicating some overlap in the sensor patterns for these two gestures.

Although the matrix shows some minor misclassifications, it attests to the overall high accuracy and precision of the proposed real-time gesture recognition system.

VI. CONCLUSION

The proposed Indian Sign Language (ISL) translation system translates hand gestures into speech which can act as a reliable and efficient assistive system for people with speech and hearing impairments. Using a Random Forest classifier, it achieves a significant 97.2% test accuracy, as well as 97.5%, 97.6%, and 97.3% precision, recall, and F1-score, respectively.

These performance metrics confirm the strength and accuracy of the model in recognizing and classifying gestures with a minimal error margin.

Confusion matrix analysis shows the system's excellent classification capability where 8 out of the 10 gestures were recognized at 100% accuracy. While minor misclassifications did occur for a limited number of cases, mostly due to overlapping sensor readings, overall, the system proved its reliability. Its ability to recognize and process communication in real-time only makes it more pragmatic as a tool to fill the communication void for the speech impaired individual.

This work furthers the field of assistive technology with a novel, sensor-based technique for the real-time translation of sign to speech. Future works may include increasing the gesture vocabulary, refining sensor calibration to achieve better accuracy, making use of optimization techniques to reach a greater computational efficiency and integrating the system into a lightweight and low-power wearable device. These upgrades will continually improve usability of the system to also develop a scalable and useful communication tool for persons with speech impairment.

REFERENCES

[1] S. M. Antad, S. Chakrabarty, S. Bhat, S. Bisen, and S. Jain, Sign language translation across multiple languages, Proceedings of the 2024 International Conference on Emerging Systems and Intelligent Computing (ESIC), 2024, pp. 741–746.

[2] Y. Dhamecha, R. Pawar, A. Waghmare, and S. Ghosh, Sign language conversion using hand gesture recognition, Proceedings of the 2023 2nd International Conference for Innovation in Technology (INOCON), 2023, pp. 1-6.

[3] H. T. H, R. S, K. H. Akhil, and K. Deepa, ML assisted sign language to speech conversion gloves for the differently abled, Proceedings of the 2024 IEEE International Conference on Interdisciplinary Approaches in Technology and Management for Social Innovation (IATMSI), 2024, pp. 1-6.

[4] J. P, G. B. A, H. A, and K. G, Real-time hand sign language translation: text and speech conversion, Proceedings of the 2024 7th International Conference on Circuit Power and Computing Technologies (ICCPCT), 2024, pp. 488-493.

[5] C. A. Kumar, V. S. Reddy, N. Kandarpa, and P. Sharma, Smart IoT system for Indian sign language translator, Proceedings of the 2024 Second International Conference on Intelligent Cyber Physical Systems and Internet of Things (ICoICI), 2024, pp. 254-258.

[6] V. Mane, T. Adsare, T. Dharmik, N. Agrawal, D. Patil, and P. Atram, Sign language translator for speech impaired people, Proceedings of the 2023 5th Biennial International Conference on Nascent Technologies in Engineering (ICNTE), 2023, pp.1-5.

[7] P. M, V. R. Balaji, T. R, and S. R, Sign Speaks—An IoT based smart gloves for dumb, Proceedings of the 2021 6th International Conference on Communication and Electronics Systems (ICCES), 2021, pp. 470-475.

[8] A. D. Shetty, J. Shetty, K. K, Rakshitha, and S. S. B, Real-time translation of sign language for speech impaired, Proceedings of the 2023 7th International Conference on Computing Methodologies and Communication (ICCMC), 2023, pp. 570-575.

[9] S. Sivamohan, S. AnslamSibi, T. R. Divakar, and S. Jagan, Hand gesture recognition and translation for international sign language communication using convolutional neural networks, Proceedings of the 2024 2nd International Conference on Advancement in Computation & Computer Technologies (InCACCT), 2024, pp. 635-640.

[10] N. T. Muralidharan, R. R. S., R. M. R., S. N. M., and H. M. E., Modelling of sign language smart glove based on bit equivalent implementation using flex sensor, Proceedings of the 2022 International Conference on Wireless Communications Signal Processing and Networking (WiSPNET), 2022, pp. 99-104.

[11] P. Telluri, S. Manam, S. Somarouthu, J. M. Oli, and C. Ramesh, Low cost flex powered gesture detection system and its applications, Proceedings of the 2020 Second International Conference on Inventive Research in Computing Applications (ICIRCA), 2020.

[12] J. Kunjumon and R. K. Megalingam, Hand gesture recognition system for translating Indian sign language into text and speech, Proceedings of the 2019 International Conference on Smart Systems and Inventive Technology (ICSSIT), 2019, pp. 14-18.

[13] C. U. Om Kumar, K. P. K. Devan, P. Renukadevi, V. Balaji, A. Srinivas, and R. Krithiga, Real-time detection and conversion of gestures to text and speech to sign system, Proceedings of the 2022 3rd International Conference on Electronics and Sustainable

Communication Systems (ICESC), 2022, pp. 73-78.
[14] J. Peguda, V. S. S. Santosh, Y. Vijayalata, A. D. R. N, and V. Mounish, Speech to sign language translation for Indian languages, Proceedings of the 2022 8th International Conference on Advanced Computing and Communication Systems (ICACCS), 2022, pp. 1131-1135.