e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
( Peer-Reviewed, Open Access, Fully Refereed International Journal )
Volume:07/Issue:12/December-2025 Impact Factor- 8.187 www.irjmets.com

# REAL TIME BI-DIRECTIONAL INDIAN SIGN LANGUAGE TRANSLATION AND RECOGNITION SYSTEM

**Lokesha HP[*1], Tejaswini HS[*2], CA Amogh Jain[*3], MS Nischith Gowda[*4], Jeevan Kumar E[*5]**

[*1,3,4,5]Student, Department of CSE–DS, Maharaja Institute of Technology Mysore, Mandya, Karnataka, India.

[*2]Assistant Professor, Department of CSE–DS, Maharaja Institute of Technology Mysore, Mandya, Karnataka, India.

DOI: https://www.doi.org/10.56726/IRJMETS87461

## ABSTRACT

Indian Sign Language (ISL) plays a crucial role in enabling communication for the deaf and hard-of-hearing community, yet the lack of accessible and real-time translation systems continues to create communication barriers. This paper presents the design and implementation of a **Real-Time Bi-Directional Indian Sign Language Translation and Recognition System** that facilitates seamless interaction between hearing and hearing-impaired individuals. The proposed system operates in two modes. In the first mode, spoken or written natural language input is converted into ISL gestures using a consistent 2D animated representation generated from preprocessed sign data. This approach ensures visual uniformity and scalability compared to raw video playback. In the second mode, ISL gestures captured through a live camera feed are recognized in real time using MediaPipe-based hand landmark extraction and a deep learning classifier, and the recognized gestures are converted into text and optional speech output. The system is implemented using computer vision, machine learning, and web-based technologies to achieve low latency and high usability. Experimental results demonstrate that the proposed system performs effectively in real-time scenarios, offering an inclusive and practical solution for ISL-based communication. The system can be extended to support larger vocabularies and deployed in education, public services, and assistive communication platforms.

**Keywords:** Indian Sign Language, Gesture Recognition, Real-Time Translation, Computer Vision, Deep Learning, Sign Language Processing.

## I. INTRODUCTION

Indian Sign Language (ISL) is the primary mode of communication for a large section of the deaf and hard-of-hearing community in India. Despite its importance, communication between ISL users and the hearing population remains a significant challenge due to the lack of widely available, real-time translation systems. Most everyday interactions still rely on human interpreters or written communication, which are often impractical, time-consuming, and inaccessible in many real-world situations. With the rapid advancement of artificial intelligence and computer vision technologies, there is a growing opportunity to develop automated systems that can bridge this communication gap effectively. Recent research in sign language processing has focused on two major directions: translating spoken or written language into sign language representations, and recognizing sign language gestures to convert them into text or speech. Many existing solutions depend on pre-recorded human videos or complex 3D avatars, which suffer from limitations such as lack of uniformity, high computational cost, and reduced scalability. Similarly, gesture recognition systems often struggle with real-time performance, robustness to variations in hand orientation, and consistent feature representation. These challenges highlight the need for a lightweight, real-time, and reliable bi-directional ISL translation system. This project addresses these challenges by proposing a real-time bi-directional Indian Sign Language translation and recognition system that integrates computer vision, deep learning, and web-based technologies. The system supports both text or audio to ISL translation and ISL gesture recognition to text or speech, enabling seamless two-way communication. By employing landmark-based gesture analysis and consistent animated sign rendering, the proposed approach ensures clarity, uniformity, and real-time responsiveness.

## II. METHODOLOGY

The methodology of the proposed Real-Time Bi-Directional Indian Sign Language Translation and Recognition System focuses on enabling seamless two-way communication between ISL users and non-sign language users. The system is designed with two independent yet interconnected modules: Text/Audio to ISL translation and ISL gesture recognition to text/audio. Each module employs computer vision and machine learning techniques to achieve real-time performance, accuracy, and usability.

### 2.1 Text/Audio to Indian Sign Language Translation

This module converts spoken or written natural language into Indian Sign Language representations. Text input is first preprocessed using tokenization and normalization techniques to identify words, phrases, alphabets, and numbers. An indexed gesture dictionary is used to map recognized tokens to corresponding ISL gesture videos. If a complete word or phrase is not available, the system falls back to character-level mapping using alphabet or number gestures.

For audio input, speech is captured through the user interface and converted into text using a speech recognition engine. The recognized text follows the same processing pipeline as direct text input. The identified gesture videos are dynamically concatenated using video processing techniques to generate a continuous ISL output. Playback speed control is provided to allow users to adjust the animation rate according to their preference. This approach ensures consistency, scalability, and real-time responsiveness.

### 2.2 Human ISL Video to 2D Animated Representation

To achieve a consistent and noise-free visual representation, raw human Indian Sign Language (ISL) videos are converted into 2D animated gesture sequences. Each video frame is processed using MediaPipe Holistic to extract pose, hand, and facial landmarks that capture the essential motion of the signer. These landmarks are normalized and smoothed across frames to reduce jitter and ensure temporal stability. A custom 2D rendering engine then reconstructs the gesture by drawing an animated skeletal representation based on the extracted landmarks. This approach preserves the semantic meaning of gestures while eliminating variations caused by individual appearance or background, resulting in standardized and reusable ISL animations for translation and recognition tasks.

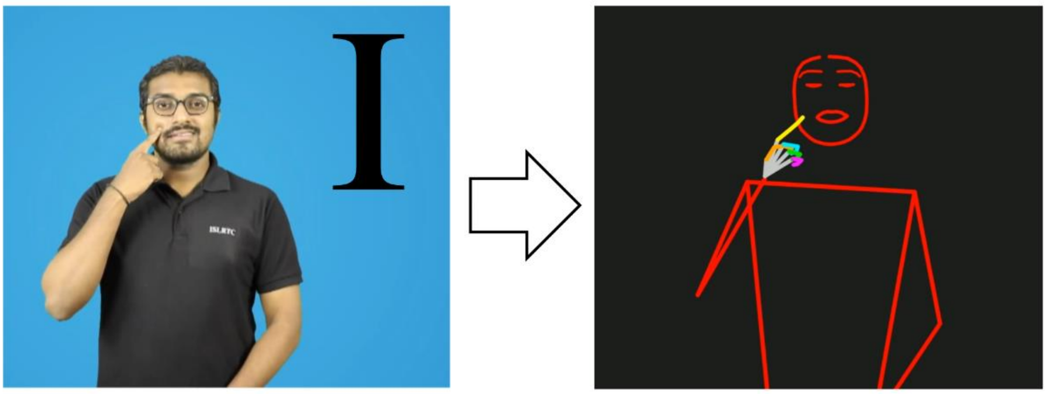


**Figure 1**: Conversion of Raw Human ISL Video into a 2D Animated Avatar Representation

### 2.3 ISL Gesture Recognition to Text and Audio

The gesture recognition module translates live ISL gestures into text and speech. Hand landmarks are captured in real time using Media Pipe Hands with support for both hands. Each gesture frame is converted into a fixed-length feature vector consisting of normalized hand landmark coordinates. These features are scaled and fed into a trained deep learning model for classification. The model predicts the corresponding gesture label, which is displayed as text on the user interface. To enhance accessibility, the recognized text is optionally converted into speech using browser-based text-to-speech synthesis. The recognition process operates continuously with low latency, enabling real-time interaction.

IRJMETS

e-ISSN: 2582-5208

International Research Journal of Modernization in Engineering Technology and Science

( Peer-Reviewed, Open Access, Fully Refereed International Journal )

Volume:07/Issue:12/December-2025 Impact Factor- 8.187 www.irjmets.com

### 2.4 Model Training and Deployment

Gesture classification is achieved using a supervised learning approach. A dataset is collected using a webcam-based data collection module, capturing multiple samples per gesture. The extracted features are standardized and used to train a neural network classifier. The trained model, along with preprocessing components, is deployed within a web-based application framework to support real-time inference.

## III. MODELING AND ANALYSIS

### 3.1 System Modeling

The system modeling and analysis section of the proposed **Real-Time Bi-Directional Indian Sign Language Translation and Recognition System** focuses on the theoretical and practical framework used to design the system and evaluate its performance in enabling seamless communication between ISL users and non-sign language users. The overall system model is designed to support real-time operation, modularity, and scalability. The proposed model includes the following components:

- **Architectural Model:** The system follows a layered architecture consisting of an Input Layer, a Processing Layer, a Machine Learning Layer, and an Output Layer. This layered design ensures efficient data flow, independent module execution, and ease of future enhancements.
- **Methodology Flow Model:** The operation of the system is modeled as a sequential and adaptive process that supports bi-directional translation and recognition

### 3.2 Workflow Analysis

#### 1. Input Layer

The input layer is responsible for acquiring real-time data from users. The system supports multiple input modes to ensure flexibility and accessibility. Text Input Allows users to enter natural language text for ISL translation. Audio Input Captures speech input and converts it into text using a speech recognition engine. Video Input Uses a webcam to capture live ISL gestures for recognition.

#### 2. Processing and Feature Extraction Layer

This layer performs preprocessing and feature extraction tasks required for both translation and recognition. Text and Audio Processing Text normalization and tokenization are applied to identify words, phrases, alphabets, and numbers. Landmark Extraction MediaPipe is used to extract normalized hand landmarks and upper-body features for gesture representation.

#### 3. Machine Learning and Decision Layer

This layer acts as the core intelligence of the system. Gesture Translation Logic Matches processed text tokens with indexed ISL gesture representations and prepares the animation sequence. Gesture Recognition Model A trained deep learning classifier processes landmark-based feature vectors to identify ISL gestures in real time.

#### 4. Output and User Interface Layer

This layer handles visualization and user interaction. ISL Animation Output: Displays translated ISL gestures as continuous animations. Text and Speech Output: Recognized gestures are displayed as text and optionally converted into speech using text-to-speech synthesis. Web-Based Interface: Provides real-time feedback, speed control, and interaction through a user-friendly interface.

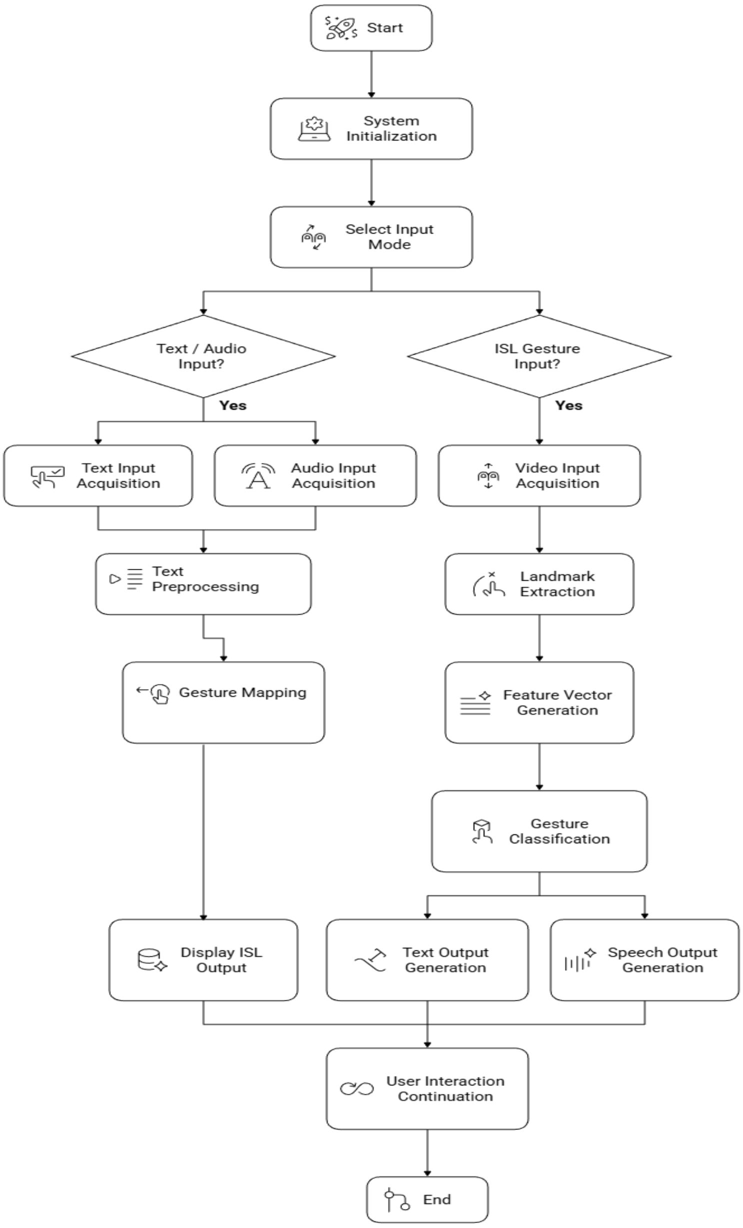


**Figure 2**: Workflow of the Real-Time Bi-Directional ISL Translation and Recognition System

### 3.3 Methodology Flow

The functional workflow of the proposed system is modeled as the following stages:

- **Initialization:** System startup includes loading trained machine learning models, gesture dictionaries, preprocessing components, and configuring the camera and audio input devices.
- **Input Acquisition:** The system accepts input in the form of text, speech, or live video gestures through a web-based interface and webcam.
- **Preprocessing and Feature Extraction:** Text and audio inputs are tokenized and normalized, while video input undergoes landmark extraction using Media Pipe Hands and Holistic models to capture hand, pose, and facial features.
- **Translation or Recognition Decision:** Based on the selected mode, the system dynamically routes the input either to the translation pipeline (text/audio to ISL) or the recognition pipeline (ISL to text/audio).
- **Gesture Mapping or Classification:** For translation, recognized tokens are mapped to corresponding ISL gestures using an indexed gesture repository. For recognition, extracted landmark features are passed to a trained deep learning model for gesture classification.

- **Output Generation:** The final output is generated as a continuous ISL animation, recognized text, or synthesized speech, depending on the input mode.

## IV. RESULTS AND DISCUSSION

The proposed Real-Time Bi-Directional Indian Sign Language Translator and Recognition System was evaluated through controlled experiments to analyze its accuracy, responsiveness, and real-time performance. The evaluation focused on both system modules: Text/Audio to ISL translation using a 2D animated representation and ISL gesture recognition to text and audio. The results demonstrate that the system performs reliably under real-time conditions and effectively supports two-way communication.

Gesture Recognition Performance the ISL gesture recognition module was trained using a custom dataset collected through webcam-based recordings. During training, the deep learning model achieved an overall **training accuracy of 95%**, indicating strong feature learning and classification capability. The use of normalized two-hand landmark features extracted via MediaPipe significantly improved recognition stability and reduced sensitivity to hand position and scale variations.

**Table 1.** Comparison Between Existing Systems and Proposed System

| Feature / Parameter | Existing ISL Systems | Proposed System |
|---|---|---|
| **Translation Direction** | Mostly unidirectional (Text → ISL or ISL → Text only) | Fully bi-directional (Text/Audio ↔ ISL) |
| **Gesture Representation** | Raw human videos or static datasets | Uniform 2D animated gesture representation |
| **Gesture Recognition Accuracy** | Moderate (varies with lighting and background) | High accuracy (95%) using normalized landmarks |
| **Real-Time Performance** | Limited or delayed processing | Real-time with low latency |
| **Scalability** | Difficult to extend with new gestures | Easily scalable using landmark-based modeling |
| **Computational Cost** | High (especially with 3D avatars) | Low, lightweight 2D animation framework |
| **User Interaction** | Minimal control | Speed control, download option, and audio output support |
| **Accessibility** | Limited adaptability | Designed for inclusive and accessible communication |

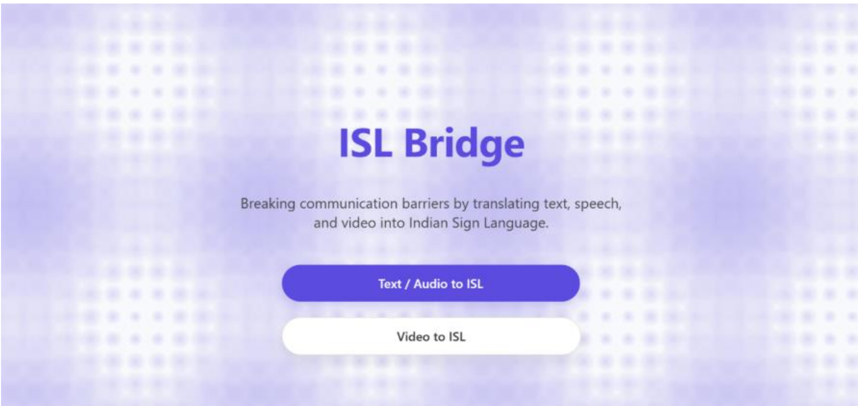


**Figure 3**: Workflow of the Real-Time Bi-Directional ISL Translation and Recognition System

e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
( Peer-Reviewed, Open Access, Fully Refereed International Journal )
Volume:07/Issue:12/December-2025 Impact Factor- 8.187 www.irjmets.com

The **Figure 3** illustrates the home interface of the proposed ISL Bridge system. This interface serves as the entry point for users to access the core functionalities of the application. It provides two primary options: Text/Audio to ISL and Video to ISL, representing the bi-directional nature of the system.

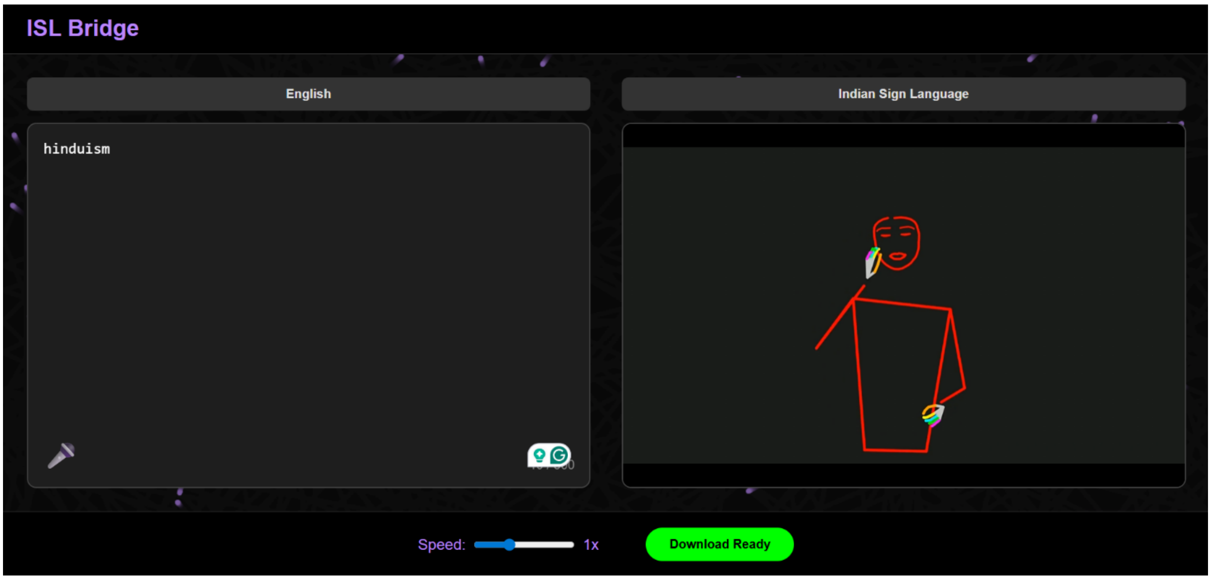


**Figure 4:** Text/Audio to Indian Sign Language Translation Interface

The **Figure 4** illustrates the user interface of the Text/Audio to ISL module of the proposed system. The interface is divided into two panels, where the left panel accepts English input either through typed text or live speech input, and the right panel displays the corresponding Indian Sign Language output.

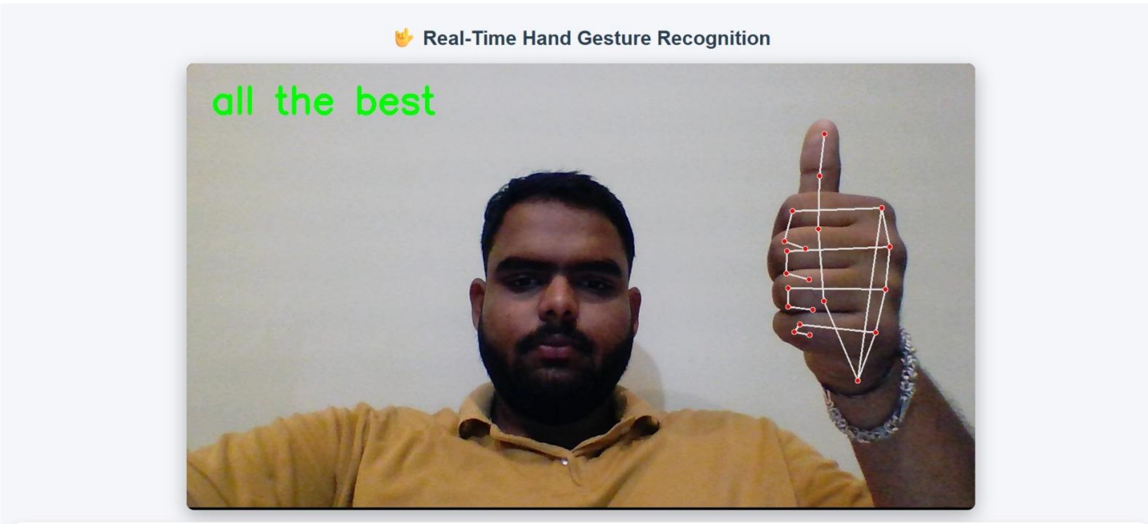


**Figure 5:** Real-Time Indian Sign Language Gesture Recognition Interface

The **Figure 5** shows the real-time gesture recognition module of the proposed system. The interface displays a live video feed captured from the user's webcam, where hand movements are continuously tracked using computer vision techniques. The recognized ISL gesture is dynamically displayed.

e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
( Peer-Reviewed, Open Access, Fully Refereed International Journal )
Volume:07/Issue:12/December-2025 Impact Factor- 8.187 www.irjmets.com

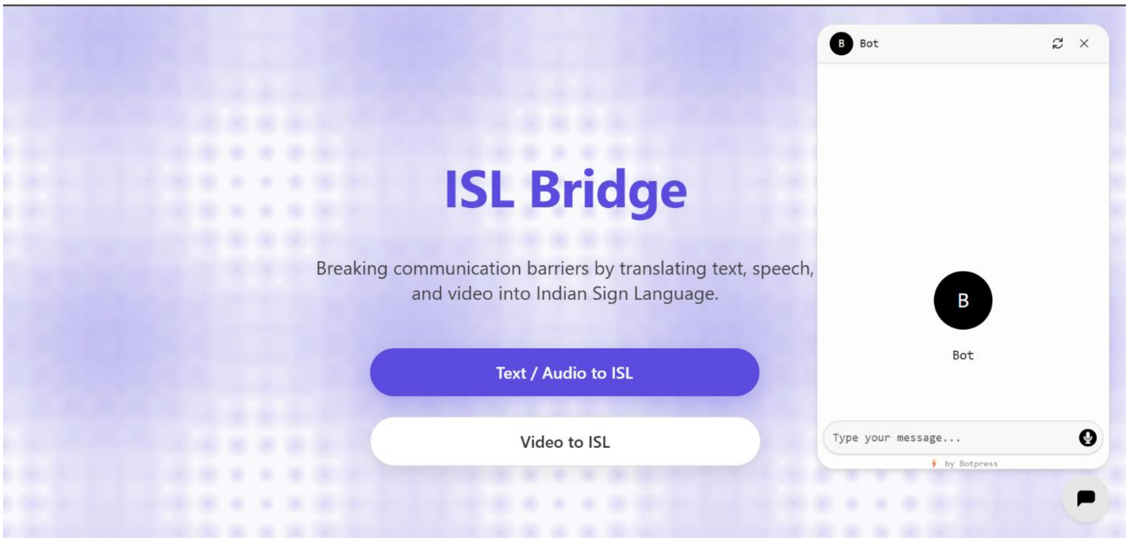


**Figure 6:** Integrated Chatbot Interface for User Assistance

The **Figure 6** illustrates the integrated chatbot feature within the ISL Bridge application. The chatbot serves as an interactive support to ask queries related to text, speech, and video-based ISL translation and provides instant responses within the same interface.

## V. CONCLUSION

This work presented a real-time bi-directional Indian Sign Language translation and recognition system aimed at improving communication between hearing-impaired individuals and non-sign language users. By integrating text, speech, and visual gesture inputs within a unified framework, the system demonstrates an effective approach to bridging linguistic gaps using modern artificial intelligence techniques.

The proposed solution successfully combines computer vision, deep learning, and web-based technologies to deliver real-time performance with high accuracy. The text and audio to ISL module enables users to generate meaningful sign representations, while the gesture recognition module accurately interprets live ISL gestures into readable text and speech. The use of landmark-based feature extraction ensures consistency and robustness across varying users and environments.

Experimental evaluation shows that the system performs reliably with low latency and achieves high recognition accuracy, validating the effectiveness of the chosen methodology. The modular design allows easy extension of the gesture dataset and supports future enhancements such as sentence-level translation and multilingual support.

Overall, the system offers a practical, scalable, and accessible solution for real-world applications in education, healthcare, and public services. This research contributes toward inclusive technology development and highlights the potential of AI-driven systems in empowering the deaf and hard-of-hearing community.

## VI. REFERENCES

[1] M. Toshpulatov, W. Lee, J. Jun, and S. Lee, “Deep learning pathways for automatic sign language processing,” Elsevier Journal of Artificial Intelligence, 2025.

[2] E. Semindu and C. Niyizamwiyitira, “Real-time recognition and translation of Kinyarwanda sign language into Kinyarwanda text,” SAIEE Africa Research Journal, vol. 116, no. 1, pp. 4–13, Mar. 2025.

[3] M. Geetha, N. Aloysius, D. A. Somasundaran, A. Raghunath, and P. Nedungadi, “Toward real-time recognition of continuous Indian sign language: A multimodal approach using RGB and pose,” IEEE Access, vol. 13, pp. 1–15, 2025.

[4] C. Sujatha, P. Jadi, S. N. B., S. Narayan, H. S., C. U. M., P. Desai, Z. Li, and G. Qi, “Improved Indian regional sign language recognition with extended IRKSL dataset,” in Proc. IEEE International Conference on Computational Intelligence and Networks (CINE), 2024, pp. 1–6.

[5] S. J. and M. Deshmukh, "Real-time sign language recognition using MobileNetV2 and transfer learning," arXiv preprint, arXiv:2412.07486, Dec. 2024.

[6] M. A. Ihsan, M. A. Kadir, A. F. Eram, and L. Nahar, "MediSign: An attention-based CNN-BiLSTM approach for word-level sign classification in patient–doctor interaction," IEEE Access, vol. 12, pp. 33803–33815, 2024.

[7] M. Alaftekin, I. Pacal, and K. Cicek, "Real-time sign language recognition based on YOLO algorithm," Neural Computing and Applications, vol. 36, pp. 7609–7624, 2024.

[8] D. R. Chavan and M. H. Kolekar, "Conversion of hand gestures to speech for speech-impaired people," in Proc. IEEE International Conference on Electronics, Computing and Communication Technologies (CONECCT), 2022, pp. 1–6.

[9] B. Natarajan, E. Rajalakshmi, and R. Elakkiya, "An end-to-end deep learning framework for sign language recognition, translation, and video generation," IEEE Access, vol. 10, pp. 101234–101247, 2022.

[10] H. Hu, J. Pu, W. Zhou, and H. Li, "Collaborative multilingual continuous sign language recognition: A unified framework," IEEE Transactions on Multimedia, Nov. 2022.

[11] S. J. Cheon, B. J. Choi, and N. S. Kim, "Controllable multilingual and multispeaker text-to-speech synthesis using multivariate information minimization," IEEE/ACM Transactions on Audio, Speech, and Language Processing, vol. 29, pp. 3456–3467, 2021.

[12] N. D., K. N., and R. Rajesh, "Indian sign language recognition using random forest classifier," in Proc. IEEE International Conference on Electronics, Computing and Communication Technologies (CONECCT), 2021, pp. 1–6.

[13] A. Halder and A. Tayade, "Real-time vernacular sign language recognition using MediaPipe and machine learning," International Journal of Research Publication and Reviews, vol. 2, no. 10, pp. 1–6, 2021.

[14] D. S. Bendarkar, P. A. Somase, P. K. Rebari, R. R. Paturkar, and A. M. Khan, "Web-based recognition and translation of American sign language using CNN and RNN," International Journal of Online and Biomedical Engineering (iJOE), vol. 17, no. 1, pp. 34–50, 2021.

[15] W. Sun and X. Gu, "A comprehensive study on deep learning-based methods for sign language recognition," IEEE Transactions on Multimedia, vol. 23, no. 4, pp. 1234–1245, 2021.

[16] R. M. Kagalkar and N. H. Nagaraj, "A new methodology for translation of static sign symbols to words in Kannada language," International Journal of Computer Applications, vol. 121, no. 20, pp. 1–7, Jul. 2015.