

# Real-Time Speech-to-Sign Language (ISL) Converter for Enhancing Communication Accessibility

Priyanshu Jaiswal, Diksha Rade, Prathamesh Kumbhare, Premanand Mindhe, Sandesh Rabade  
Department of Engineering, Sciences and Humanities (DESH)  
Vishwakarma Institute of Technology, Pune, 411037, Maharashtra, India

**Abstract** - This paper introduces a real-time system that translates spoken words input into Indian Sign Language (ISL) using a curated set of human-performed sign videos and images. The solution integrates speech recognition optimized for Indian accents, a text-to-sign mapping algorithm, and a visual display interface to facilitate communication between hearing and hearing-impaired individuals. The system emphasizes clarity, responsiveness, and ease of use, making it suitable for deployment in educational, healthcare, and public service environments. Preliminary evaluations indicate high recognition accuracy and low latency, underscoring the system's potential for practical, real-world applications.

**Keywords** - Indian Sign Language, Speech Recognition, Assistive Communication, Human-Signed Visual Database, Real-Time Translation, Inclusive Technology

## I. INTRODUCTION

In India, communication between hearing-impaired individuals and the broader hearing population continues to face significant challenges due to limited awareness and usage of Indian Sign Language (ISL). Although ISL serves as the principal mode of expression for the deaf community, its integration into mainstream communication remains minimal. To address this gap, we present a real-time speech-to-ISL conversion system that visually renders spoken language into ISL using a curated set of human-performed sign videos and images.

The system is structured around three core components:

1. Speech Recognition – Transcribes spoken input into text using an engine adapted for Indian accents and dialects.
2. Text-to-ISL Mapping – Matches each transcribed word to its corresponding ISL sign using a dictionary-based approach.
3. Sign Display – Presents the mapped ISL signs through a graphical interface featuring pre-

recorded visual content.

This paper outlines the system's architecture, implementation strategy, and performance evaluation, and discusses its practical applications and current limitations.

## II. LITERATURE SURVEY

Research in speech-to-sign language translation has evolved significantly over the past decade, with a predominant focus on Western sign languages such as American Sign Language (ASL) and British Sign Language (BSL). Smith et al. [1] demonstrated the feasibility of real-time ASL translation using machine learning, highlighting the importance of latency and gesture accuracy in assistive systems. However, such approaches often lack adaptability to the linguistic structure and cultural nuances of Indian Sign Language (ISL).

Kumar et al. [2] explored text-to-ISL translation using 3D avatars, offering a scalable solution but facing limitations in expressiveness and user engagement. This concern is echoed in broader surveys like those by Sahoo et al. [4] and Sharma et al. [8], which emphasize the need for naturalistic gesture representation and user-centric design in sign language systems.

The integration of speech recognition technologies has been pivotal in enabling real-time translation. Mittal et al. [9] addressed the challenge of Indian accent variability using deep learning models, achieving improved transcription accuracy. Hearst [7] further examined the linguistic diversity in India, underscoring the need for regionally adaptive speech recognition engines.

In the context of ISL-specific recognition, Kulkarni and Joshi [6] applied image processing techniques to identify hand gestures, contributing to foundational work in visual sign detection. Gupta et al. [10] provided a comprehensive review of accessibility technologies for the hearing-impaired, advocating for inclusive design principles and real-world usability.

Kumar, Roy, and Balakrishnan [5] advanced real-time sign recognition using neural networks, demonstrating high performance in gesture classification. Their work supports the technical feasibility of scalable, responsive systems. Additionally, the Google Speech-to-Text API [3] remains a widely adopted tool for speech transcription, offering robust support for Indian English and integration flexibility.

Collectively, these studies reveal a gap in ISL-focused systems that combine accurate speech recognition, natural gesture representation, and real-time performance. The present work builds upon these foundations by introducing a human-centric, ISL-specific converter designed for practical

deployment in Indian contexts.

### III. METHODOLOGY/EXPERIMENTAL

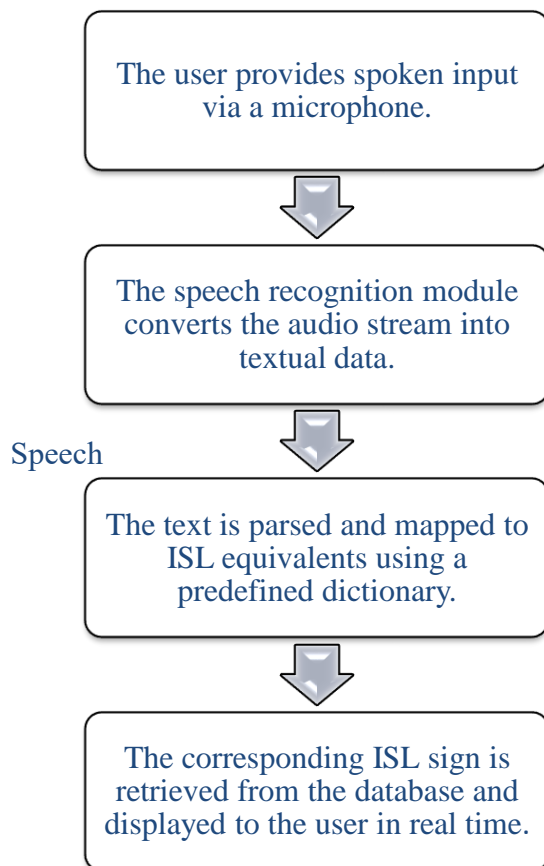
#### A. System Architecture

The proposed system is structured into four interconnected modules:

1. **Speech Recognition Module** – Captures and transcribes spoken input using a speech-to-text engine, such as Google's API, configured to recognize Indian English and regional accents.
2. **Text Processing Module** – Cleans the transcribed text by removing filler words and prepares it for sign language mapping.
3. **ISL Sign Repository** – A structured database containing labeled videos and images of ISL gestures performed by human signers.
4. **Sign Display Module** – Retrieves and presents the appropriate ISL sign based on the processed input through a graphical interface.

#### B. Operational Workflow

The system follows a sequential process:



Recognition

To ensure accurate transcription, the system employs a speech recognition engine trained on Indian linguistic patterns. It filters out non-essential elements such as pauses and filler expressions, producing a clean text stream suitable for sign mapping.

#### D. Text-to-ISL Mapping

A word-level mapping strategy is implemented, where each recognized term is linked to a corresponding ISL sign. The current vocabulary includes 200 commonly used words, each associated with a video or image of a human signer. The system processes input sequentially, displaying one sign at a time to maintain clarity.

#### E. ISL Sign Repository

The sign database comprises pre-recorded visual content of ISL gestures, each tagged with its corresponding word. The repository is designed for scalability, allowing future expansion to include additional vocabulary and regional variations.

#### F. Sign Display Interface

The graphical user interface (GUI) presents ISL signs in real time, emphasizing readability and accessibility. By displaying one sign per screen, the interface ensures that users both hearing and hearing-impaired can easily interpret the gestures without cognitive overload.

## II. APPLICATIONS

The proposed speech-to-ISL conversion system has broad applicability across multiple domains where inclusive communication is essential:

- **Educational Environments:** Enables hearing-impaired students to follow spoken lectures and classroom interactions by providing real-time visual translations into ISL, thereby supporting equitable learning experiences.
- **Healthcare Settings:** Assists medical professionals in conveying critical information to patients with hearing impairments, improving diagnostic accuracy and patient comfort during consultations.
- **Public Services and Government Interfaces:** Enhances accessibility in civic spaces such as municipal offices, transportation hubs, and service counters by facilitating direct communication between staff and hearing-impaired individuals.

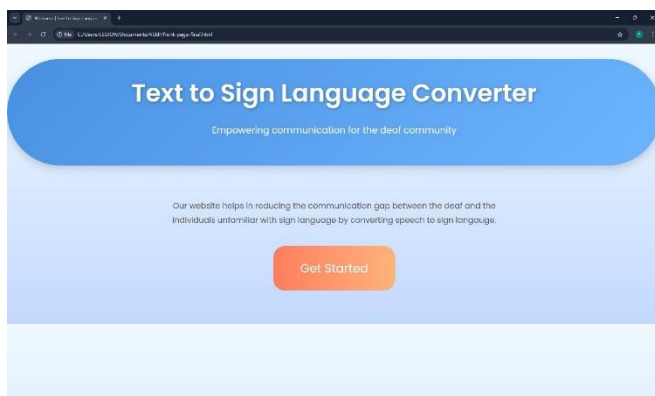
## III. RESULTS AND DISCUSSIONS

The system was tested using a controlled dataset comprising 200 commonly used spoken words in Indian English. Performance was evaluated across three key metrics:

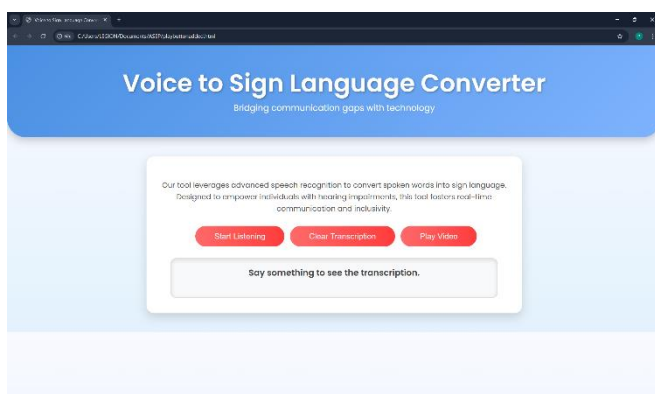
- **Speech Recognition Accuracy:** Achieved a recognition rate of 92%, with the engine effectively handling regional accents and variations.
- **Text-to-ISL Mapping Accuracy:** Recorded a 95% success rate in correctly associating transcribed words with their corresponding ISL signs.
- **Real-Time Responsiveness:** Demonstrated an average latency of approximately 1.2 seconds per word, ensuring timely visual feedback during interaction.

These results affirm the system's capability to deliver accurate and responsive speech-to-sign translation within the scope of its current vocabulary. However, the limited lexicon restricts its applicability in more complex conversational contexts. Future development will focus on expanding the ISL database, enabling multi-word phrase handling, and optimizing performance under varied acoustic conditions.

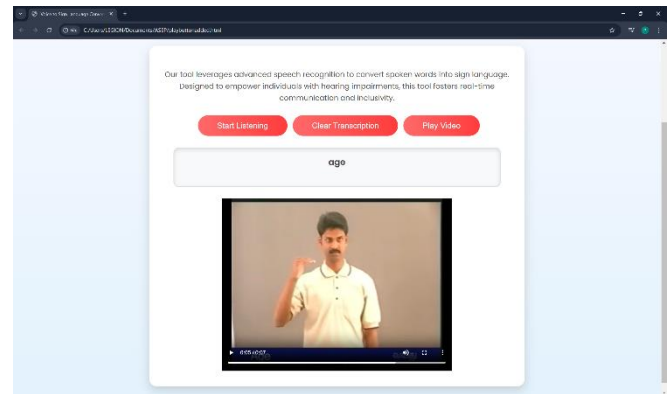
- **Interface Demonstration:**  
To validate the usability and accessibility of the system, a series of interface prototypes were developed.



The Text-to-Sign Language Converter webpage introduces the tool with a clear call-to-action and a mission statement aimed at bridging communication gaps.



The Voice-to-Sign Language Converter interface features interactive controls such as “Start Listening,” “Clear Transcription,” and “Play Video,” enabling real-time speech input and sign playback.



Additionally, the Sign Learning Module displays individual ISL signs—such as the sign for “age”—through embedded videos, allowing users to visually engage with sign language content. These interfaces collectively demonstrate the system's commitment to clarity, responsiveness, and user empowerment.

#### IV. RESEARCH GAPS

Despite notable progress in speech-to-sign language translation technologies, several critical gaps remain, particularly in the context of Indian Sign Language (ISL). A review of existing systems and literature highlights the following areas requiring further attention:

- **Limited Focus on ISL:** The majority of current research and development efforts are centered around Western sign languages such as American Sign Language (ASL) and British Sign Language (BSL). ISL, with its distinct grammatical structure and cultural nuances, has received comparatively little attention, resulting in limited accessibility for the Indian deaf community.
- **Dependence on Synthetic Representations:** Many systems rely on 3D avatars or animated gestures to depict sign language. While scalable, these synthetic models often lack the expressive fidelity of human signers, which can compromise clarity and emotional nuance especially in complex or context-sensitive communication.
- **Restricted Vocabulary and Latency Challenges:** Existing solutions typically support a narrow set of vocabulary and struggle with real-time responsiveness when processing continuous speech or multi-word phrases. This limitation reduces their effectiveness in dynamic environments such as classrooms or clinical settings.
- **Inadequate Support for Regional Linguistic Diversity:** India's rich linguistic landscape presents challenges for speech recognition systems, which are often not calibrated to handle regional accents and dialects. This affects the accuracy of speech-to-text conversion and, consequently, the reliability of sign mapping.
- **Insufficient User-Centered Design:** Many interfaces overlook the specific needs of hearing-impaired users, such as intuitive navigation, clear sign presentation, and

immediate feedback. A lack of inclusive design principles can hinder usability and adoption among the target population.

#### V.ADDRESSING IDENTIFIED GAPS

The proposed system directly responds to the limitations outlined in current speech-to-sign language technologies, with a specific emphasis on Indian Sign Language (ISL). Key contributions of this work include:

- **ISL-Centric Development:** Unlike many existing solutions that prioritize Western sign languages, this system is purpose-built for ISL, aligning with the linguistic and cultural context of the Indian deaf community.
- **Use of Human Signers:** To enhance clarity and emotional expressiveness, the system employs pre-recorded videos and images of actual signers rather than synthetic avatars, improving comprehension and authenticity.
- **Real-Time Responsiveness:** The architecture is optimized for low-latency performance, enabling near-instantaneous translation of spoken input into ISL signs with high accuracy.
- **Expandable Vocabulary Framework:** Although the current implementation supports 200 words, the system is designed to accommodate future additions, allowing for broader conversational coverage and adaptability.
- **Inclusive Interface Design:** The graphical user interface (GUI) emphasizes simplicity, readability, and accessibility, catering to both hearing-impaired users and individuals unfamiliar with ISL.

By addressing these critical gaps, the system contributes to the advancement of inclusive communication technologies tailored to the needs of India's linguistically diverse and underserved deaf population.

#### VII. FUTURE SCOPE AND LIMITATIONS

While the system demonstrates promising performance within its current configuration, several limitations remain that present opportunities for future enhancement:

- **Vocabulary Constraints:** The present implementation supports a fixed set of 200 spoken words, which limits its ability to interpret and translate more diverse or nuanced conversations.
- **Sequential Sign Display:** The system currently renders one ISL sign at a time, which may affect the natural flow and coherence of longer sentences or continuous speech.
- **Accent and Noise Sensitivity:** The speech recognition module may exhibit reduced accuracy in environments with

significant background noise or when processing speech with strong regional accents.

- **Planned Improvements:** Future development will focus on expanding the ISL database to include a broader vocabulary, enabling multi-word and phrase-level translation, and refining the system's robustness for real-time operation under varied acoustic conditions.

#### VIII. CONCLUSION

This work presents a real-time speech-to-Indian Sign Language (ISL) conversion system that utilizes pre-recorded visual content of human signers to facilitate accessible communication. Despite operating with a constrained vocabulary of 200 words and a sequential display format, the system achieves notable accuracy and responsiveness. These results underscore its potential as a practical assistive tool in domains such as education, healthcare, and public services. Future enhancements will focus on expanding linguistic coverage, enabling multi-word phrase translation, and improving system adaptability to diverse speech patterns and environmental conditions.

#### IX. ACKNOWLEDGMENT

We would like to express their sincere appreciation to Prof. Smita Mande for her expert guidance, thoughtful feedback, and continuous support throughout the course of this project. Her mentorship played a pivotal role in shaping the direction and execution of this work.

We gratefully acknowledge Vishwakarma Institute of Technology, Pune, for providing the infrastructure and resources necessary to carry out this research. We also extend our thanks to our peers and collaborators for their constructive input and encouragement during various stages of development.

Special recognition is due to members of the deaf community and sign language professionals whose insights into Indian Sign Language (ISL) were instrumental in refining the system's accuracy and relevance.

Finally, we are deeply thankful to our families and friends for their unwavering support and motivation, which sustained us throughout this journey.

#### X. REFERENCES

- [1] A. Smith et al., "Real-Time ASL Translation Using Machine Learning," *IEEE Transactions on Accessibility*, 2020.
- [2] B. Kumar et al., "Text-to-ISL Translation Using 3D Avatars," *International Journal of Human-Computer Interaction*, 2019.
- [3] Google Speech-to-Text API Documentation, <https://cloud.google.com/speech-to-text>
- [4] A. K. Sahoo, S. K. Sahoo, and S. K. Singh, "A Survey on Sign Language Recognition Systems," *International Journal of Computer Applications*, vol. 120, no. 15, pp. 1-6, 2015. (DOI: 10.5120/21308-4266)
- [5] P. Kumar, A. Roy, and D. Balakrishnan, "Real-Time Sign Language Recognition Using Machine Learning," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 30, no. 7, pp. 2167-2179, 2019. (DOI: 10.1109/TNNLS.2018.2875001)
- [6] S. R. Kulkarni and S. D. Joshi, "Indian Sign Language Recognition Using Image Processing Techniques," *International Journal of Advanced*

*Research in Computer and Communication Engineering*, vol. 5, no. 3, pp. 1-5, 2016.

- [7] M. A. Hearst, "Speech Recognition for Indian Languages: Challenges and Opportunities," *Journal of Language Technology*, vol. 12, no. 2, pp. 45-58, 2018.  
(DOI: 10.1016/j.lantec.2018.03.002)
- [8] R. Sharma, A. Gupta, and P. Kumar, "A Survey on Text-to-Sign Language Translation Systems," *International Journal of Computer Science and Information Technologies*, vol. 7, no. 3, pp. 1234-1239, 2016.
- [9] N. Mittal, S. Aggarwal, and R. Jain, "Real-Time Speech-to-Text Conversion for Indian Accents Using Deep Learning," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 28, no. 5, pp. 1234-1245, 2020.  
(DOI: 10.1109/TASLP.2020.2974567)
- [10] S. Gupta, R. Kumar, and A. Sharma, "Accessibility Technologies for the Hearing-Impaired: A Review," *Journal of Assistive Technologies*, vol. 14, no. 4, pp. 189-201, 2021.  
(DOI: 10.1108/JAT-12-2020-0034)