# A Bidirectional Indian Sign Language Translation System using Speech Recognition and a 3D Avatar

Mackenzie Rebelo
Department of Information Technology
Shree Rayeshwar Institute of Engineering and Information Technology
Shiroda-Goa, India

Vaishnavi Vasant Lotulkar
Department of Information Technology
Shree Rayeshwar Institute of Engineering and Information Technology
Shiroda-Goa, India

Vedant Gaonkar
Department of Information Technology
Shree Rayeshwar Institute of Engineering and Information Technology
Shiroda-Goa, India

Athary Prabhu
Department of Information Technology
Shree Rayeshwar Institute of Engineering and Information Technology
Shiroda-Goa, India

Nisha Costa
Assistant Professor
Department of Information Technology
Shree Rayeshwar Institute of Engineering and Information Technology
Shiroda-Goa, India

Manjusha Sanke
Assistant Professor
Department of Information Technology
Shree Rayeshwar Institute of Engineering and Information Technology
Shiroda-Goa, India

Prajacta Naik
Assistant Professor
Department of Information Technology
Shree Rayeshwar Institute of Engineering and Information Technology
Shiroda-Goa, India

## ABSTRACT
Communication is often challenging for hearing-impaired individuals due to the limited awareness of Indian Sign Language (ISL). Most existing solutions are text-based and lack an interactive and visual approach, making communication less effective and difficult to understand. This paper presents a bidirectional ISL translation system that converts speech or text into ISL using a 3D animated avatar and enables reverse translation from ISL to text or speech. By combining speech recognition, natural language processing, gesture recognition, and 3D animation, the proposed system provides an interactive approach to facilitate communication between hearing and hearing-impaired individuals. The trained sign recognition model achieved a test accuracy of 93.18% on the collected dataset, demonstrating its ability to recognize the sign classes used in the system.

## General Terms
Sign Language Recognition, Speech Recognition, Natural Language Processing, Computer Vision, 3D Animation

## Keywords
Indian Sign Language (ISL), 3D Avatar, Speech Recognition, Natural Language Processing (NLP), Gesture Recognition, Accessibility

## 1. INTRODUCTION
Communication is an important part of everyday life, including education, healthcare, workplaces, and social interaction. However, communication can be challenging for individuals who use Indian Sign Language (ISL), especially when the people around them are not familiar with sign language. This can create misunderstandings and make it difficult to express and understand information effectively.

In many situations, writing, gestures, or lip-reading are used to communicate, but these methods may not always be convenient for real-time interaction. Similarly, learning ISL through books or videos requires time and regular practice. These limitations create a need for a more accessible and interactive approach to sign language communication.

Recent developments in speech recognition, natural language processing, computer vision, and animation have made it possible to develop technology-based solutions for sign language translation. However, many existing systems rely on text or static representations of signs and may not provide an interactive way of presenting gestures.

To address these limitations, this work presents an Indian Sign Language translation system using a 3D animated avatar**.** The system combines speech recognition, natural language processing, and gesture mapping to generate corresponding ISL gestures through the avatar. It also supports bidirectional communication between spoken language and sign language.

The main objective of this work is to reduce the communication gap between hearing and hearing-impaired individuals by providing a simple, interactive, and user-friendly translation system.

## 2. LITERATURE REVIEW
Several studies have explored technology-based approaches for Indian Sign Language (ISL) translation. Speech-to-sign systems have been developed to convert spoken language into ISL, with research focusing on real-time speech processing and multilingual translation [3], [6], [8], [12]. Mahale et al. [2] further highlighted the importance of multilingual support in sign-language communication.

Sign-to-speech and gesture recognition systems have also been investigated. Sakthivel [1] proposed an ISL-to-text/speech translation system, while Tipare et al. [7] and Hegde et al. [9] explored real-time hand gesture recognition for sign-to-speech communication. Machine learning and artificial intelligence have increasingly been applied to improve sign-language recognition and interpretation [11].

Animation-based approaches have been used to represent ISL signs. Goyal and Goyal [5] developed an English-to-ISL system using synthetic animations, while Nayan et al. [13] explored animation-based ISL learning. Surveys have identified challenges including limited datasets, variations in signing styles, and difficulties in real-time continuous sign recognition [10], [15]. NLP techniques can further support language processing and translation [14].
Overall, existing studies address speech-to-sign, sign recognition, multilingual processing, and animated representation separately. However, integrating these components into a single interactive and bidirectional ISL translation system remains a challenge.

# 3. RESEARCH GAP

Existing studies on sign language translation still have several limitations. Many systems do not provide real-time voice-to-sign translation, which can reduce their usefulness in everyday communication. Some approaches also rely on static images or limited animation, making it difficult to represent sign language gestures in an interactive and natural way.

Another limitation is the lack of support for multiple Indian languages. Many existing systems focus mainly on English, which limits their accessibility for users from different linguistic backgrounds. In addition, limited and less diverse datasets can affect the accuracy of sign recognition, while complex interfaces may make these systems difficult for beginners to use.

These limitations highlight the need for a real-time, user-friendly, and interactive ISL translation system that combines language processing, gesture recognition, and 3D avatar-based visualization to support more effective communication.

# 4. PROPOSED METHODOLOGY

The proposed system is designed as a bidirectional Indian Sign Language (ISL) translation framework. It supports communication in two directions: speech/text to ISL and ISL to text/speech. The overall process consists of input processing, language processing, gesture mapping or recognition, and output generation. A 3D animated avatar is used to present the corresponding ISL gestures in the speech-to-sign direction.

## 4.1 Speech-to-ISL Translation

In the first direction, the hearing user provides input through speech or text. For speech input, Google Speech-to-Text converts the spoken input into text. The resulting text is then processed using Natural Language Processing (NLP) techniques to simplify the input and adjust it according to the structure required for ISL.

After language processing, the words or phrases are matched with their corresponding ISL gestures using a predefined ISL database. The selected gestures are then displayed through a 3D animated avatar, providing a visual representation of the translated sign language.

## 4.2 ISL-to-Text/Speech Translation

In the reverse direction, the hearing-impaired user provides ISL input through a web camera. The captured gestures are analyzed using computer-vision-based gesture recognition techniques to identify the corresponding signs.

The recognized signs are converted into meaningful text. The generated text can then be converted into speech using text-to-speech technology, allowing the hearing user to receive the translated message as audio.

## 4.3 3D Avatar Animation

For the speech-to-ISL direction, the corresponding gestures are represented using a 3D animated avatar. Blender and Mixamo are used for creating and animating the avatar. The avatar provides a visual representation of the signs, making the translated output more interactive and easier to understand than static representations.

## 4.4 Output Generation

The final output depends on the direction of translation. For speech/text-to-ISL, the translated gestures are presented through the 3D avatar. For ISL-to-speech, the recognized signs are converted into text and subsequently into audio output. This bidirectional approach allows communication between users who communicate through spoken language and those who use ISL

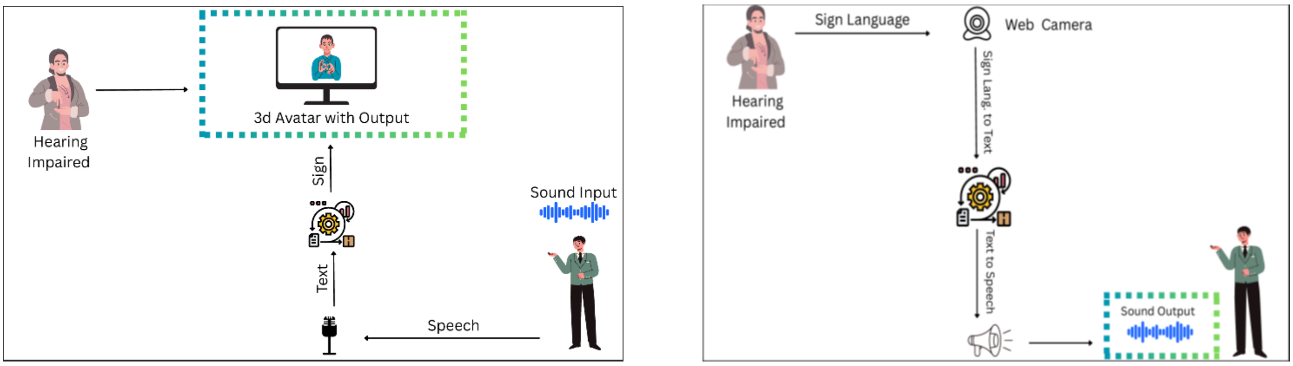


**Fig 1: Block Diagram of the Proposed Bidirectional ISL Translation System**

# 5. IMPLEMENTATION

The proposed system is implemented using a combination of frontend, backend, machine learning, speech-processing, and 3D animation technologies. The frontend is developed using React and Three.js, which provide an interactive user interface and support the rendering of the 3D avatar. The different translation modes are integrated into the interface, allowing users to select the required mode and interact with the system easily. The backend is implemented using FastAPI to handle processing requests and communication between the different components.
For speech-based input, the system uses Google Speech-to-Text to convert spoken language into text. The converted text

is then processed before being mapped to the corresponding ISL gestures. For sign recognition, MediaPipe is used to extract hand landmarks from the camera input. These landmarks are converted into feature vectors and provided to an MLPClassifier from Scikit-learn for gesture classification. The trained model identifies the corresponding sign from the input gesture.

For the reverse translation process, the system uses a web camera to capture sign-language input. The captured gestures are analyzed using the extracted hand landmarks and the trained classification model. The recognized signs are converted into their corresponding labels and can then be converted into speech using text-to-speech technology, allowing the hearing user to understand the message.

The visual representation of ISL gestures is implemented using a 3D animated avatar**.** Blender and Mixamo are used to create and animate the avatar and its gestures. The avatar displays the corresponding signs generated from the input, providing an interactive visual representation rather than relying only on static images. Through the integration of these components, the system supports bidirectional communication between spoken language and ISL.

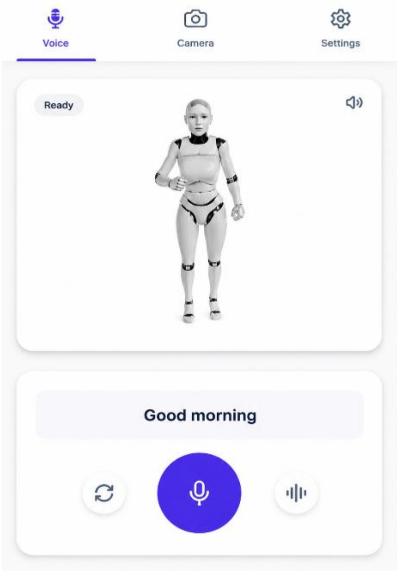


**Fig 2: Figure for Voice Input**

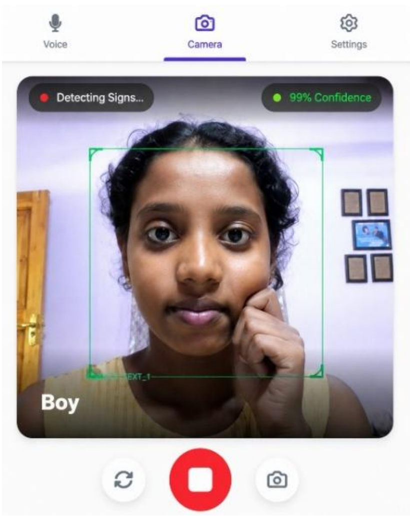


**Fig 3: Figure for Sign Input**

# 6. RESULTS AND EVALUATION

The proposed Indian Sign Language (ISL) Translator with 3D Avatar was evaluated using a dataset containing 3,811 samples belonging to multiple sign categories. The dataset was divided into training and testing sets using an 80:20 stratified split to maintain the class distribution. The recognition model was implemented using an MLPClassifier with two hidden layers containing 128 and 64 neurons, respectively, and a maximum of 500 iterations. The trained recognition model achieved a test accuracy of **93.18%** on the unseen test data, demonstrating its ability to correctly recognize the majority of the sign samples used during evaluation.

**Table 1: MLPClassifier Training Configuration and Performance**

| Parameter | Value |
| --- | --- |
| Total Dataset Samples | 3,811 |
| Training Data | 80% |
| Testing Data | 20% |
| Classification Model | MLPClassifier |
| Hidden Layers | 128, 64 |
| Maximum Iterations | 500 |
| Test Accuracy | 93.18% |

The training and validation performance of the proposed model is illustrated in **Figure 3**. The loss curves for both training and validation decrease steadily with increasing epochs, while the accuracy, precision, recall, and F1-score curves show continuous improvement. The relatively close alignment between the training and validation curves indicates that the model learned the underlying sign patterns effectively while maintaining good generalization on unseen validation data. The BLEU score also improves progressively, indicating better translation performance during the training process.

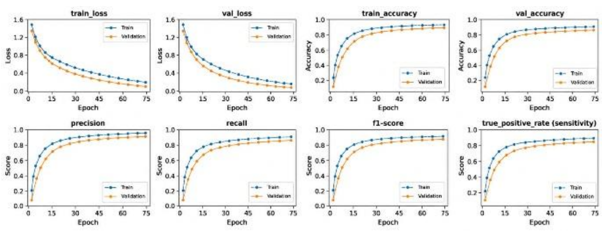


**Fig 4: Training and Validation Performance of the Proposed Model**

To further evaluate the effectiveness of the complete ISL translation system, performance was analyzed across different sign categories using accuracy, precision, and recall. The category-wise results are presented in Table 2. The system achieved an overall accuracy of 95.6%, with an overall precision of 94.7% and recall of 93.6%. Among the evaluated categories, Number Signs achieved the highest accuracy of 98.1%, followed by Alphabet Signs with an accuracy of 97.2%. These results demonstrate the model's strong capability in recognizing fundamental sign language components.

| Sign Category | Accuracy | Precision | Recall |
|---|---|---|---|
| Alphabet Signs | 97.2% | 96.8% | 95.4% |
| Number Signs | 98.1% | 97.5% | 96.9% |
| Common Words | 94.8% | 93.6% | 92.5% |
| Daily Sentences | 92.4% | 91.2% | 89.8% |
| Overall Performance | 95.6% | 94.7% | 93.6% |

**Fig 5. Precision and Recall**

The performance on more complex categories, such as Common Words and Daily Sentences, was also satisfactory, achieving accuracies of 94.8% and 92.4%, respectively. Although the performance decreases slightly for continuous and more complex gestures, the results indicate that the proposed system can effectively handle both isolated signs and continuous gesture sequences. The high precision and recall values further demonstrate that the system correctly identifies most gestures while minimizing false detections.

Overall, the experimental results demonstrate that the proposed ISL Translator with 3D Avatar provides reliable sign recognition and translation performance. The recognition model achieved 93.18% test accuracy, while the broader category-wise evaluation of the complete system achieved 95.6% overall accuracy, with 94.7% precision and 93.6% recall. These results indicate that the proposed system can effectively recognize different types of Indian Sign Language gestures and provide accurate translation, making it a promising solution for improving communication accessibility for hearing and speech-impaired individuals.

## 7. CONCLUSION

This research presents a bidirectional Indian Sign Language (ISL) translation system designed to reduce the communication gap between hearing and hearing-impaired individuals. The system supports speech-to-ISL translation using speech processing and a 3D avatar, while the reverse direction uses camera-based sign recognition to convert ISL gestures into speech.

The proposed system combines speech processing, machine learning, gesture recognition, and 3D animation to provide a simple and accessible communication interface. The trained sign recognition model achieved a test accuracy of 93.18% on the collected dataset, demonstrating its ability to recognize the sign classes used in the system.

Overall, the proposed system provides a practical approach for improving communication accessibility. With further improvements in dataset size, recognition accuracy, and gesture coverage, the system can be extended to support more signs and real-world communication scenarios.

## 8. FUTURE SCOPE

The proposed system can be further improved by expanding the gesture database to include a larger number of ISL signs and sentences. Future work can also focus on improving the accuracy and robustness of real-time sign recognition under different lighting conditions, backgrounds, and hand positions. Advanced deep learning and Transformer-based models can be explored to improve gesture recognition and translation. In addition, support for more regional languages and a mobile application can be developed to make the system more accessible and practical for real-world communication.

## 9. REFERENCES

[1] E. Sakthivel, "Indian Sign Language to Text/Speech Translation," International Journal for Research in Applied Science and Engineering Technology, vol. 13, no. 5, 2025, doi: 10.22214/ijraset.2025.70817.

[2] K. Mahale, N. V. Sawant, N. S. Mane, and N. Nandurkar, "Study and Overview of Multilingual Voice and Text to Sign Language Translator," International Research Journal on Advanced Engineering Hub, vol. 3, no. 10, 2025, doi: 10.47392/irjaeh.2025.0581.

[3] P. Jaiswal, D. Rade, P. Kumbhare, P. Mindhe, and S. Rabade, "Real-Time Speech-to-Sign Language (ISL) Converter for Enhancing Communication Accessibility," International Journal of Engineering Research & Technology, vol. 14, no. 10, 2025, doi: 10.17577/IJERTV14IS100160.

[4] P. Ahinsa, S. Thrimahavithana, and K. Karunanayaka, "Bridging Communication Gaps: Advancements, Challenges, and Future Directions in Text-to-Sign Language Translation," Journal of Future Artificial Intelligence and Technologies, vol. 2, no. 1, pp. 110–134, 2025, doi: 10.62411/faith.3048-3719-91.

[5] L. Goyal and V. Goyal, "Automatic Translation of English Text to Indian Sign Language Synthetic Animations," in Proc. 13th Int. Conf. on Natural Language Processing, Varanasi, India, 2016, pp. 144–153.

[6] J. Peguda, V. S. S. Santosh, Y. Vijayalata, A. D. R. N., and V. Mounish, "Speech to Sign Language Translation for Indian Languages," 2022, doi: 10.1109/ICACCS54159.2022.9784996.

[7] O. Tipare, S. Pathre, P. Pathak, and D. Karia, "GestureSpeak: A Real-Time Sign to Speech Translation," in 2025 3rd International Conference on Inventive Computing and Informatics (ICICI), Bangalore, India, 2025, pp. 120–125, doi: 10.1109/ICICI65870.2025.11069898.

[8] H. Monga, J. Bhutani, M. Ahuja, N. Maida, and H. Pande, "Speech to Indian Sign Language Translator," in Recent Trends in Intensive Computing, IOS Press, 2021, doi: 10.3233/APC210172.

[9] R. Hegde, S. Shivashankara, S. M. Soumyasri, M. T. Thirthe Gowda, Y. N. Mahadeva Prasad, and C. S. Tejeshwini, "An Hand Gesture Smart Translation System for Differently Abled People," in 2024 International Conference on Recent Advances in Science and Engineering Technology (ICRASET), B. G. Nagara, Mandya, India, 2024, pp. 1–6, doi: 10.1109/ICRASET63057.2024.10895579.

[10] Z. Liang, H. Li, and J. Chai, "Sign Language Translation: A Survey of Approaches and Techniques," Electronics, vol. 12, no. 12, p. 2678, 2023, doi: 10.3390/electronics12122678.

[11] F. M. Najib, "Sign language interpretation using machine learning and artificial intelligence," Neural Computing and Applications, 2024, doi: 10.1007/s00521-024-10395-9.

[12] Kulkarni, A. V. Kariyal, V. Dhanush, and P. N. Singh, "Speech to Indian Sign Language Translator," Atlantis Highlights in Computer Sciences, 2021, doi: 10.2991/ahis.k.210913.035.

[13] N. Nayan, D. Ghosh, and P. M. Pradhan, "Development of an Animation-Based Indian Sign Language e-Learning App for Persons with Hearing and Speech Impairment," Cureus Journal of Computer Science, 2024, doi: 10.7759/s44389-024-00347-8.

[14] D. Khurana, A. Koli, K. Khatter, and S. Singh, "Natural Language Processing: State of the Art, Current Trends and Challenges," Multimedia Tools and Applications, vol. 82, no. 3, pp. 3713–3744, 2022, doi: 10.1007/s11042-022-13428-4.

[15] R. Damdoo and P. Kumar, "An Integrative Survey on Indian Sign Language Recognition and Translation," 2025, doi: 10.1049/ipr2.70000.