Real Time Sign Language Conversation Using AI

Bhagyalakshmi V[1]; Chandana S [2]; Himasri J S [3]; Krupashree N R [4]; Muhammed, Sinan K Z [5]

[1]Assistant Professor, Department of Electronics and Communication Engineering, Maharaja Institute of Technology Thandavapura, Nanjangud, Mysore, Karnataka, India

[2,3,4,5] Students, Department of Electronics and Communication Engineering, Maharaja Institute of Technology Thandavapura, Nanjangud, Mysore, Karnataka, India

DOI: https://doi.org/10.51244/IJRSI.2026.1307000420

Received: 06 August 2026; Accepted: 12 August 2026; Published: 22 August 2026

ABSTRACT

This project introduces a real-time Indian Sign Language (ISL) recognition and translation system to improve communication between hearing-impaired and non-signing individuals. Using OpenCV and MediaPipe, the system accurately detects and tracks hand landmarks to recognize gestures made with one or both hands. Each gesture is mapped to a predefined ISL phrase, which is then converted into text and spoken output using a non-blocking text-to-speech engine (pyttsx3) for smooth and simultaneous operation. The system also supports text-to-sign translation, displaying pre-recorded ISL videos with synchronized speech for interactive learning. Developed entirely in Python, the model provides real-time performance and requires minimal hardware. It can be applied in educational, public, and assistive communication systems. Overall, the proposed system offers an efficient, affordable, and inclusive solution that enhances accessibility and bridges communication barriers for individuals with hearing and speech disabilities.

Keywords: Indian Sign Language, gesture recognition, OpenCV, MediaPipe, text-to-speech, accessibility.

INTRODUCTION

Communication plays a vital role in daily life, but individuals with hearing and speech impairments often face barriers in expressing themselves to others who do not understand sign language. Indian Sign Language (ISL) serves as a crucial medium for such individuals, yet the lack of awareness and understanding among the general population limits effective interaction. [1]

With the advancement of computer vision and artificial intelligence, it has become possible to develop systems that can recognize hand gestures and translate them into meaningful text and speech. This project utilizes OpenCV and MediaPipe frameworks for realtime hand tracking and gesture recognition, enabling accurate detection of two-hand ISL gestures. [2-3]

The system converts gestures into text and speech, and vice versa—translating text into sign videos. This dualmode design promotes accessibility, inclusivity.

METHODOLOGY

The proposed system is designed to translate Indian Sign Language (ISL) gestures into text and speech, and vice versa. It operates in two main modes — Sign-to-Text and Text-to-Sign. The project uses Python as the core programming platform, integrating OpenCV and MediaPipe libraries for efficient hand-tracking and gesture recognition.

In the Sign-to-Text mode, live video is captured through a webcam and processed frame by frame. MediaPipe extracts key hand landmarks, and OpenCV assists in visualizing hand positions. The system analyzes the relative positions of fingers to determine gesture patterns. Each recognized gesture corresponds to a predefined phrase stored in a gesture-phrase mapping dictionary. [4]

Once a gesture is recognized, the identified phrase is displayed as text and simultaneously converted into speech using the pyttsx3 text-to-speech engine.

A non-blocking multithreading approach ensures that speech output does not interrupt real-time video processing, maintaining smooth performance.

Input Output Interface

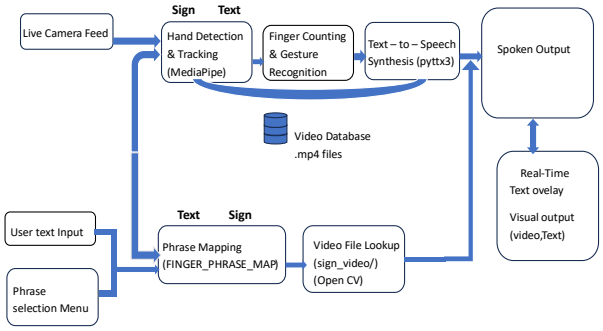

Fig 1: Block diagram of Real Time Sign Language conversation using AI

In the Text-to-Sign mode, users can input or select text, and the system retrieves corresponding prerecorded ISL videos for playback. This enables interactive learning and communication between hearing and speech-impaired individuals and nonsigners. The overall design ensures accessibility, accuracy, and realtime performance in both directions of translation. [5-6]

Implementation and Simulation

The proposed Indian Sign Language Translator System is implemented in Python using OpenCV, MediaPipe, and pyttsx3 libraries. The system consists of two operational modes: Sign-to-Text and Text-to-Sign. In the Sign-to-Text mode, real-time video input is captured through the webcam using OpenCV. MediaPipe's hand tracking model detects hand landmarks and extracts features such as finger count and thumb position. These features are compared against predefined gesture mappings to identify the corresponding ISL phrase. The recognized phrase is displayed on-screen and simultaneously converted into speech through a separate non-blocking pyttsx3 thread to ensure smooth parallel execution. [7-9]

In the Text-to-Sign mode, users select a phrase from a menu. The system searches for the corresponding prerecorded sign video file in the local directory and plays it using OpenCV's video playback function. Simultaneously, the pyttsx3 engine producessynchronized speech output, enhancing the interactive experience. [10-14]

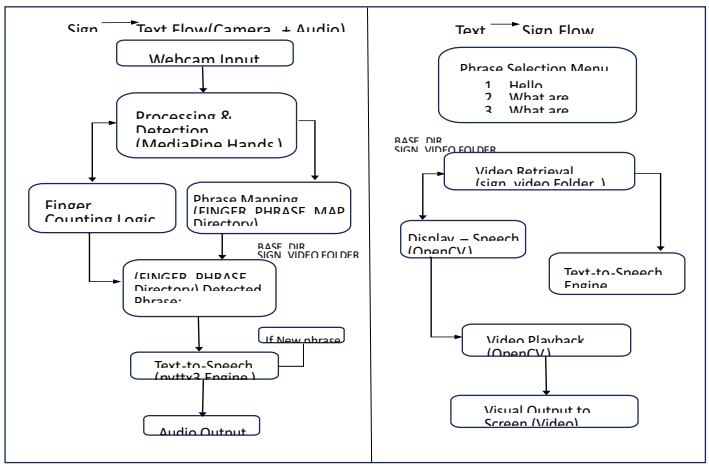

Fig 2: Flow diagram of Real Time Sign Language conversation using AI

ISSN No. 2321-2705 | DOI: 10.51244/IJRSI |Volume XIII Issue VII July 2026

Simulation was performed on a standard Windows system with a webcam. The program achieved stable realtime gesture detection and accurate phrase recognition. The dual-mode functionality ensures bidirectional communication and supports both learning and assistive applications for hearing and speechimpaired individuals. [15-18]

RESULT AND ANALYSIS

The developed Indian Sign Language (ISL) Translator System was tested under various real-time conditions to evaluate its accuracy, performance, and usability. The Sign-to-Text mode successfully detected and interpreted multiple ISL gestures with high reliability using MediaPipe's hand-tracking framework. The system accurately recognized gestures corresponding to

predefined phrases such as "hello, what is your name," "please wait for sometime," and "are you busy." The recognized phrases were displayed as onscreen text andsimultaneously converted into clear audio output through the pyttsx3 engine.[19-24]

TABLE 1: Data Reorganization Table

Left Hand	**Right Hand**	**Sign (Phrase)**
1	1	hello, what is your name
1	2	what are you doing
1	3	are you busy
1	4	let's go for lunch
1	5	please wait for sometime
2	2	any question
2	3	don't worry
2	4	good question
3	3	what is today's date
5	5	nice to meet you
2	2	any question
2	3	don't worry
2	4	good question
3	3	what is today's date
5	5	nice to meet you

The system demonstrated real-time responsiveness, maintaining smooth frame processing at approximately 25–30 frames per second. The non blocking text-to speech implementation ensured that gesture recognition

ISSN No. 2321-2705 | DOI: 10.51244/IJRSI |Volume XIII Issue VII July 2026

continued seamlessly during speech synthesis. In different lighting environments and hand orientations the system maintained consistent detection performance, proving its adaptability for real-world use. **[25-28]**

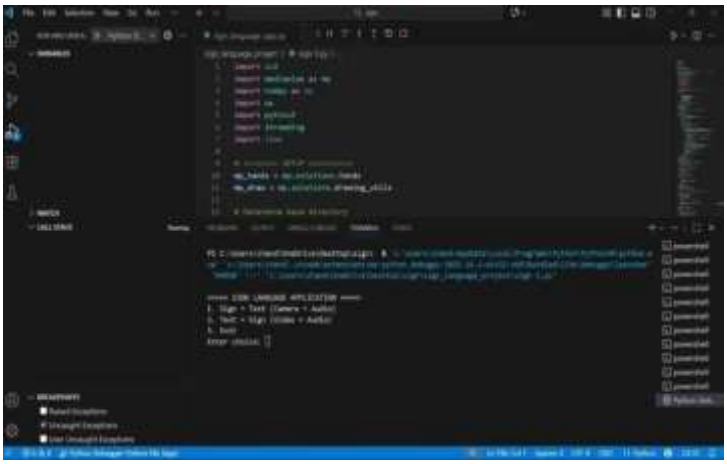

Figure 1: Execution of Sign Language Application in

Visual Studio Code (Python)

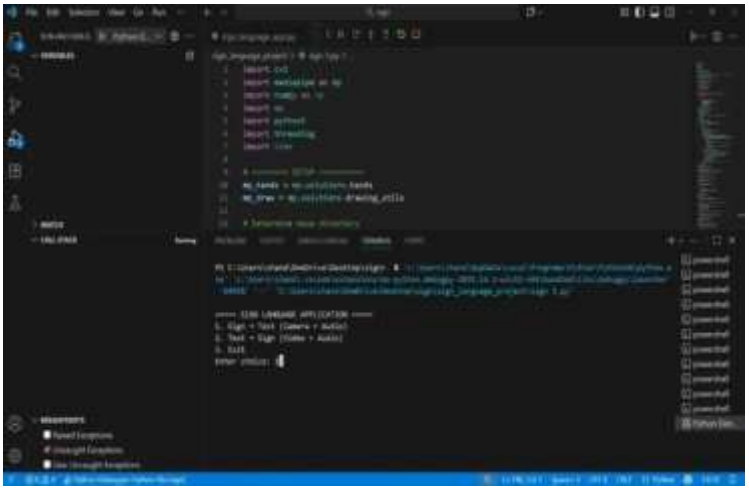

Figure 2: System Output – Sign to Text and Audio Translation entering the input 1 for sign to text conversion

Figure 3: Output image – sign to text and audio

ISSN No. 2321-2705 | DOI: 10.51244/IJRSI |Volume XIII Issue VII July 2026

Figure 4: input image – entering input 2 for text to sign conversation

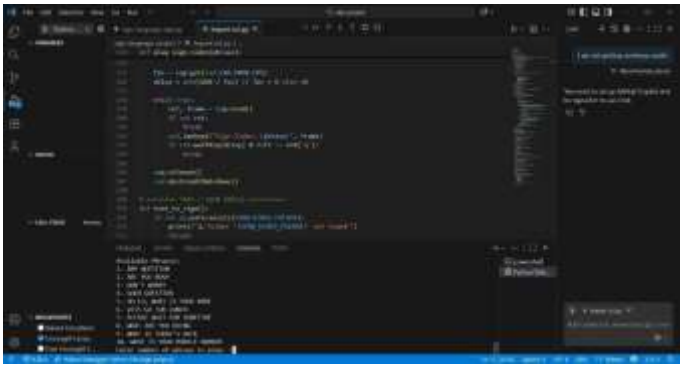

Figure 5: Giving input to convert text to sign

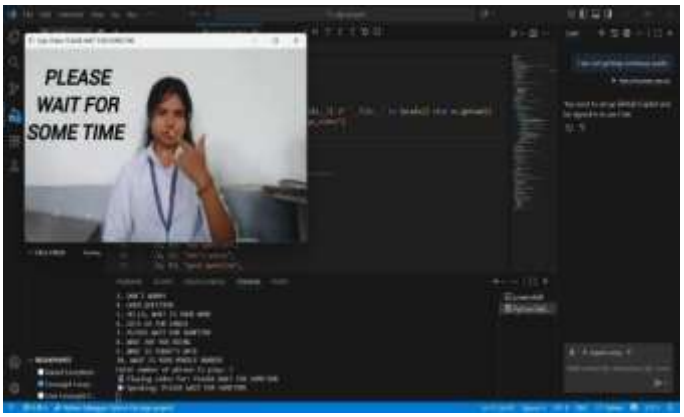

Figure 6: Output image – text to sign

In the Text-to-Sign mode, all pre-recorded sign videos were played correctly with synchronized speech output. The user interface responded promptly to input selections, providing an interactive and accessible to

Communication experience. Overall, the system achieved reliable recognition accuracy and effective dual-mode operation, validating its potential for educational, assistive, and public service applications aimed at bridging communication barriers for the hearing and speech impaired. [**29-30]**

CONCLUSION, CHALLENGES AND FUTURE ENHANCEMENTS

Conclusion

The proposed Indian Sign Language (ISL) Translator System successfully enables two-way communication between hearing-impaired individuals and non-signers. The Sign-to-Text mode effectively recognizes hand gestures using MediaPipe and OpenCV, translating them into meaningful text and speech. Likewise, the

Text-to-Sign mode plays corresponding sign videos with synchronized audio, offering an interactive learning experience. The system's real-time performance, high accuracy, and smooth non-blocking speech generation demonstrate its practical potential in classrooms, public service areas, and assistive communication devices. Overall, it provides a low-cost, efficient, and accessible solution to bridge communication barriers and promote inclusivity**.**

Challenges

During implementation, a few challenges were encountered. Gesture recognition accuracy occasionally fluctuated under poor lighting or when hands overlapped. The system required proper background contrast and consistent camera positioning for reliable detection. Synchronization between video playback and speech output also needed fine-tuning. Additionally, the predefined gesture set limited vocabulary coverage, restricting the translation of complex or dynamic signs.

Future Enhancements

Future improvements can include integrating machine learning models for dynamic gesture recognition and expanding the ISL vocabulary. Adding facial expression detection and emotion-based translation could improve context understanding. Cloud or mobile integration may enable IoT-based accessibility, allowing real-time communication through smart phones or kiosks. Multilanguage speech output can further enhance usability and inclusivity in diverse environments.

REFERENCES

1. S. Kumar, A. Singh, and R. Patel, "CNN-based Indian Sign Language Recognition,"
2. International Journal of Computer Applications,
3. R. Sharma and P. Gupta, "Real-Time Indian Sign Language Recognition Using MediaPipe and OpenCV," International Research Journal of Engineering and Technology
4. P. Gupta and N. Kaur, "Vision-Based Gesture Translation with Text-to-Speech Output for ISL," International Journal of Innovative Science and Research Technology (IJISRT), 2020.
5. M. Raj and D. S. Reddy, "Hand Gesture Recognition for Indian Sign Language Translation," International Journal of Advanced Research in Computer and Communication Engineering, 2021.
6. K. S. Ramesh and P. Kumar, "Design and Implementation of Real-Time ISL to Speech Conversion System," International Journal of Engineering Trends and Technology (IJETT), 2022.
7. Z. Yan et al., "Spatial-Temporal Graph Convolutional Networks for Action Recognition," IEEE TPAMI, , 2019.
8. F. Zhang and S. Chen, "Real-Time Hand Pose Estimation Using Deep Learning," IEEE Access, 2020.
9. T. Starner and A. Pentland, "Real-Time American Sign Language Recognition Using Hidden Markov Models, 1997.
10. MediaPipe Hands, Google AI, "Real-Time Hand Tracking," 2022.
11. G. Bradski, "The OpenCV Library," Dr. Dobb's Journal of Software Tools, 2000.
12. M. S. Hossain and G. Muhammad, "Hand
13. Gesture Recognition Using CNN and Transfer
14. Learning," IEEE Access, 2019 Kumar and S. Mehta, "Machine Learning Techniques for Gesture Recognition," IJERT, 2020. Graves, "Speech Synthesis with Deep Neural Networks, 2013.
15. D. P. Kingma and J. Ba, "Adam: A Method for Stochastic Optimization," ICLR, 2015.
16. J. B. Tiwari and R. Shukla, "A Review on AIbased Assistive Systems for Hearing-Impaired Users," IJCSE, 2021.
17. A. Mittal and V. Bhatia, "Assistive Communication System for Deaf and Dumb Using Computer Vision," IJACSA, 2019
18. S. Rautaray and A. Agrawal, "Vision-Based Hand Gesture Recognition for Human–Computer Interaction: A Survey," Artificial Intelligence Review, 2015.
19. M. Elmezain, A. Al-Hamadi, and B. Michaelis, "Robust Methods for Hand Gesture Recognition Using Vision-Based Systems," IEEE Transactions on Systems, Man, and Cybernetics, 2009.

20. J. Shotton et al., "Real-Time Human Pose Recognition in Parts from Single Depth Images," IEEE CVPR, 2011.
21. D. Molchanov, S. Gupta, K. Kim, and J. Kautz, "Hand Gesture Recognition with 3D Convolutional Neural Networks," IEEE CVPR Workshops, 2015.
22. T. Ong and S. Ranganath, "Automatic Sign Language Analysis: A Survey and the Future Beyond Lexical Meaning," IEEE TPAMI, , 2005.
23. P. Premaratne and P. Meegama, "Hand Gesture Recognition Using Feature Extraction," International Journal of Computer Science and Network Security, 2006.
24. R. Poppe, "A Survey on Vision-Based Human Action Recognition," Image and Vision Computing, 2010.
25. S. Mitra and T. Acharya, "Gesture Recognition: A Survey," IEEE Transactions on Systems, Man, and Cybernetics, 2007.
26. H. Cooper, B. Holt, and R. Bowden, "Sign Language Recognition," IEEE Signal Processing Magazine, 2012.
27. M. Koller, H. Ney, and R. Bowden, "Deep Learning of Mouth Shapes for Sign Language," IEEE ICCV,2015.Chaudhary, J. Raheja, and K. Das, "Intelligent Approaches for Hand Gesture Recognition," Pattern Recognition and Image Analysis,2013
28. J. F. C. Mendes et al., "Vision-Based Assistive Technology for Deaf People," Sensors, 2019
29. S. Bheda and A. Radpour, "Using Deep Learning for Sign Language Recognition," arXiv preprint arXiv:1708.07687, 2017.
30. K. Sahoo and S. Mishra, "Computer Vision Based Sign Language Recognition System," IJCSI, 2018.

AUTHORS PROFILE

Mrs. Bhagyalakshmi V received her M.Tech degree in Industrial Electronics from Visvesvaraya Technological University (VTU), Belagavi, and is currently pursuing her Ph.D. at VTU, Belagavi. She is serving as an Assistant Professor with over 14 years of teaching experience in engineering education. She has Completed PG Diploma in Child and School Counselling. Her teaching and research interests Electronic Circuits and emerging engineering technologies. She has actively contributed to academic development through teaching, curriculum delivery, and student mentoring. Her research focus is directed toward advanced applications in electronics and technology-driven engineering solutions, with a strong commitment to academic excellence, research innovation, and professional development.

Ms Chandana S, Ms Himasri J S, Ms Krupashree N R, and Mr Muhammed Sinan K Z are undergraduate students of the Department of Electronics and Communication Engineering, Maharaja Institute of Technology, Thandavapura. This work was carried out as part of their academic project under faculty guidance.