IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 4, April 2026

DOI: 10.17148/IJARCCE.2026.154131

Gesture Talk

Tejas Sawant[1], Gaurav Singh[2], Shubham Tiwari[3], Aman Verma[4], Dipali Shende[5]

Dept. of AI & Data Science, A. C. Patil College of Engineering, Navi-Mumbai, India[1-5]

Abstract: Communication is one of the most basic needs for humans, but millions of people who are deaf or hard of hearing struggle to get information and connect with the world around them. The tools available to help them are often limited, providing only pictures that don't show the full, natural way sign language works. Many platforms also don't respond quickly or allow people to use their voice, making them hard to use for those with different needs. This paper introduces a web application that helps bridge the communication gap for the deaf and hard-of-hearing community by converting text and speech into real-time, animated Indian Sign Language (ISL). The app lets users type or speak naturally, and then turns that into smooth, animated signs. It uses speech recognition to turn spoken words into text, which is then mapped to an animation system that shows the signs accurately. Instead of using still images, the system creates smooth animations that help people understand better and stay engaged. The web-based design works on any device without needing extra software. By combining speech recognition, language processing, and animation all in one place, the app offers a simple and inclusive way for the deaf community to communicate more easily.

Index Terms: Indian Sign Language, speech recognition, real-time conversion, web application, deaf accessibility, animation rendering.

I. INTRODUCTION

Effective communication is essential for people to fully engage in learning, work, and social activities. However, for those who are deaf or have hearing difficulties, communicating with others who can hear remains a big challenge. Indian Sign Language (ISL) is the main way millions of hearing-impaired people in the country express themselves, but it is not well understood by most people. In the past, efforts to help bridge this gap have used human interpreters, fixed sign charts, or recorded videos. These methods are limited in how well they can scale, how available they are, and how quickly they can respond in real time. Without better, technology-based tools, many hearing-impaired people face difficulties in everyday digital interactions [1].

Also, most previous research and existing apps have mainly focused on recognizing sign language and turning gestures intotext, rather than the other way around — converting spoken or written words into signs. Although these recognition systems are useful, they place the responsibility of communication en-tirely on the hearing-impaired person. Many of these platforms also don't support voice input, use still images instead of smooth animations, and aren't easy to use through standard web browsers without special hardware or software.

To fix these problems, this paper introduces a web-based system that converts both text and speech into real-time animated Indian Sign Language. Instead of using static images, the system uses an animation pipeline to create smooth and natural sign sequences, making it easier for users to under-stand.

Unlike traditional assistive tools, this system combines speech recognition, text processing, and animation in one plat-form that can be used through a web browser [2]. This design removes the need for complex setup, supports various input methods, and ensures the system is inclusive and practical for use in real-life settings like schools, public services, and healthcare facilities.

The system also integrates Google Translate API to support input from any regional Indian language, including Hindi, Marathi, Tamil, and Bengali, among others. This multilingual capability ensures that the platform is not restricted to English-speaking users, significantly widening its reach and making it genuinely accessible to communities across the country where regional languages are the primary mode of communication. Furthermore, the system incorporates SignTube for retrieving reference sign videos and a fingerspelling fallback mechanism for words not present in the ISL database [2]. Together, these features ensure complete coverage of user input regardless of vocabulary limitations, making the system robust, reliable, and well-suited for real-world deployment in diverse communica-tion environments.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 4, April 2026

DOI: 10.17148/IJARCCE.2026.154131

TABLE I
LITERATURE SURVEY SUMMARY

Sr. No.	Title	Author	Objective	Methodology	Benefits	Drawbacks
1	Sign Language Generation System Based on Indian Sign Language Grammar	Sugandhi, P. Kumar, S. Kaur	Translate English text into ISL animations using ISL grammar rules	ISL parser, HamNoSys notation, SiGML conversion, avatar-based animation	Grammatically accurate ISL output; covers alphabets, numbers, academic and legal terms	Supports only English input; no speech input; no regional Indian language support
height2	Audio or Text to Sign Language Converter	Ankita Harkude et al.	Convert audio or text to ISL using NLP and 3D animation	Speech recognition, NLTK, Django framework, Blender 3D animation	Bridges communication gap; produces 3D animated ISL output	Complex Blender pipeline increases latency; no multilingual or regional language support
3	Harnessing AI to Generate Indian Sign Language from Natural Speech	Kulkarni et al.	Develop a speech/text-to-ISL conversion system using 3D modelling	Automatic Speech Recog-nition (ASR), ISL gram-mar analysis, 3D model-based gesture animation	High accuracy (99.2%); user-friendly interface; supports voice input	Relies on complex 3D modelling; limited to English speech; no regional language translation
4	Enhanced Sign Language Translation between ASL and ISL Using LLMs	Anonymous et al., arXiv 2024	Bridge communication between ASL and ISL users using deep learning	Random Forest Classifier for ASL gesture recognition, LLM for ASL-to-ISL translation	Context-aware translation; handles linguistic nuances between sign languages	Requires ASL input first; does not accept text or speech directly; no regional language input
5	An Integrative Survey on Indian Sign Language Recognition and Translation	Damdoo and Kumar, IET 2025	Survey ISLRT techniques and identify research gaps	Review of ISL datasets, recognition models, and translation approaches	Comprehensive overview of ISL challenges and directions	Highlights that most systems handle only isolated signs or limited vocabulary; no real-time animation or multilingual input
6	Sign Language Interpretation using Machine Learning and AI	Springer Neural Computing 2024	Survey sign language recognition and speech-to-sign translation systems	Review of ML, image pro-cessing, and animation-based SL systems includ-ing Android apps	Covers both recognition and generation; surveys popular assistive apps	Majority of surveyed systems focus on ASL; ISL and regional language systems remain underrepresented
7	Challenges and Innovative Solutions in Sign Language Technology	JETIR 2025	Identify barriers in SL technology and propose inclusive solutions	Survey of ASL, BSL, ISL systems; analysis of lin-guistic and technological challenges	Highlights key barriers including linguistic diversity and lack of standardization	Most reviewed systems support only ASL or BSL; no support for regional Indian languages or animated ISL output
8	From Lab to Field: Real-World Evaluation of an AI-Driven Smart Video Solution to Enhance Community Safety	S. Yao, B. Rahimi Ardabili, et al.	Evaluate the performance and scalability of AI-driven video surveillance systems in real-world community safety deployments	Deep learning-based video analytics tested in real-world environments, focusing on object detection, event recognition, and system deployment challenges	Demonstrates real-world feasibility; highlights practical performance considerations	Limited discussion on pixel-level analysis; focuses on evaluation rather than novel algorithms

II. RELATED WORK

A. *Literature Survey*

B. *Research Gap*

Existing approaches are largely restricted to ASL or BSL, support only English text input, and lack a unified platform capable of real time multilingual conversion, animated ISL output, speech recognition, and accessible web-based deploy-ment within a single scalable system. [3]

III. PROPOSED SYSTEM

The proposed system is an AI-powered real-time web ap-plication that turns text and speech into Indian Sign Language animations, helping the deaf and hard-of-hearing community

communicate more effectively. It is built on a simple web-based setup, so it works smoothly without needing special

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 4, April 2026

DOI: 10.17148/IJARCCE.2026.154131

equipment or a human interpreter.

The system uses several integrated modules that work together to produce accurate ISL signs:

Speech and Text Input Processing: The system accepts either typed text or spoken words as input. When using speech, the browser's microphone picks it up and converts the words into text right away using a speech recognition engine. This text is then passed to the next stage of the pipeline.

Multilingual Translation: The system uses the Google Translate API to handle input in any Indian regional language. No matter which language the user types or speaks in, the input is translated into a standardized English form before

being mapped to ISL signs, making the system accessible to a much wider range of users.

Animation Rendering: To display the signs, the system uses MediaPipe, which generates smooth, real-time skeletal animations of ISL gestures. Instead of relying on still images or pre-recorded videos, these animations are fluid and natural, making the signs far easier to understand.

The system also addresses diverse accessibility needs through dedicated components:

Sign Video Integration: The system connects to SignTube to retrieve reference sign videos for words, giving users additional visual context alongside the generated animations. **Web Dashboard:** Users interact with the system through a simple, browser-based dashboard that displays live animation output, accepts text or speech input, and providesreal-time feedback. Since it is entirely web-based, it works on any device
without requiring additional software.

Algorithm 1: ISL Web Application: Text and Speech to Sign Language Conversion

IV. System Architecture

The proposed system is an AI-driven real-time web ap-plication built on a lightweight browser-based architecture that automates sign language conversion using integrated translation and animation intelligence. The system captures user input either as typed text or live speech through the browser interface, where input undergoes preprocessing steps such as language detection and text normalization to ensure consistent performance across varying regional languages and accents. A lightweight Flask based backend enables real time processing, translation, and animation streaming through a web-based dashboard. Multiple conversion stages [4] are handled seamlessly — for speech input, the system employs a browser-integrated speech recognition engine that transcribes spoken words into text in real time. The transcribed or directly typed text is further processed using the Google Translate API, which detects the input language and translates it into a standardized form before mapping it to the ISL sign database. Sign generation is achieved using MediaPipe, which renders skeletal keypoint-based animations for each recognized word or phrase. These animations are displayed fluidly on the web dashboard, where gesture sequences are rendered frame by frame to ensure smooth and natural ISL representations that are easy to comprehend [5]. Instead of relying on pre-recorded static image mappings, the system incorporates a dedicated Animation Pipeline based on real time keypoint rendering. Sign video references are fetched through SignTube integra-tion to provide supplementary visual context alongside the generated animations. Accessibility is ensured by comparing input words against a structured ISL vocabulary database to confirm valid sign mappings. Upon successful conversion, the system activates the output display mechanism [6]. Animations are rendered in real time on the dashboard, containing the corresponding ISL gesture sequence along with the original
input text for quick comprehension and user verification.

Algorithm 1 ISL Web Application: Text and Speech to Sign Language Conversion

Initialize web application, speech engine, translation API,
and ISL sign database
Load modules: Speech Recognition, Google Translate API, MediaPipe, SignTube
for each user session S_t in Web Application W **do**
    Accept user input: text typed or speech via microphone
    if input mode is speech **then**
    Capture audio using browser speech recognition en-gine
    Transcribe audio to raw text T_{raw}
    else
    Accept typed text as raw text T_{raw}
    end if
    Detect language of T_{raw} using Google Translate API
    if language $\neq$ English **then**

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 4, April 2026

DOI: 10.17148/IJARCCE.2026.154131

13: Translate T_{raw} to English using Google Translate API
14: Store translated text as T_{en}
15: **else**
16: Set $T_{en} \leftarrow T_{raw}$
17: **end if**
18: Normalize T_{en}: lowercase, remove punctuation, tok-enize into words
19: **for** each word W_i in T_{en} **do**
20: **if** W_i exists in ISL sign database **then**
21: Fetch corresponding ISL sign mapping for W_i
22: Retrieve reference sign video from SignTube
23: Render MediaPipe skeletal keypoint animation for W_i
24: **else**
25: Perform fingerspelling animation for W_i letter by letter
26: **end if**
27: **end for**
28: Sequence all rendered animations in word order
29: Stream final ISL animation to web dashboard in real time
30: **if** animation rendering successful **then**
31: Display ISL output alongside original input text T_{raw}
32: Log session: input text, detected language, timestamp
33: **end if**
34: **end for**

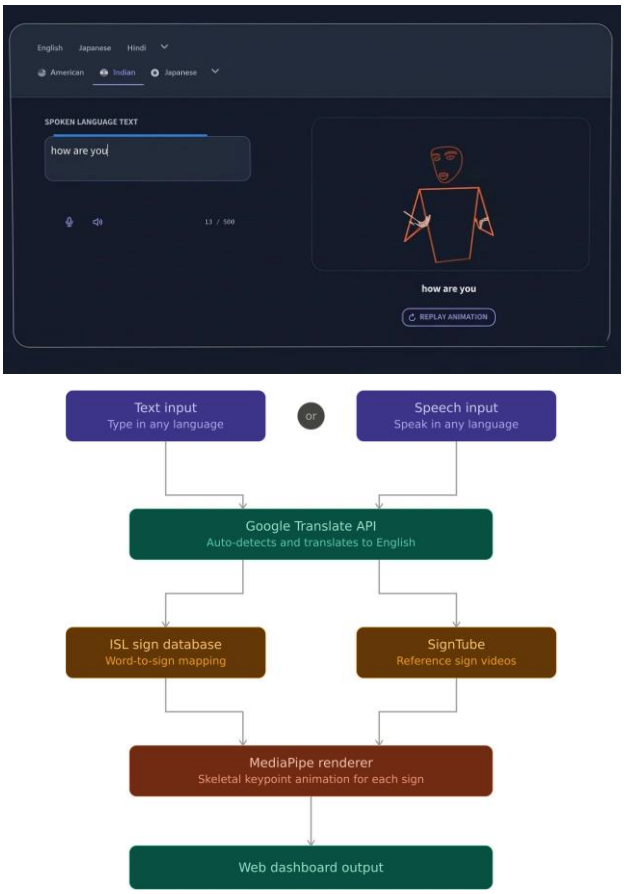

Fig. 4. Architecture of Gesture Talk

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 4, April 2026

DOI: 10.17148/IJARCCE.2026.154131

This diagram shows how the ISL web application works, demonstrating the process of handling text or speech input from the user [7] through a single browser-based system. Translation models take the input, which can be in any language, and convert it into a standard format before linking it to signs in the ISL sign database. The system allows for different ways to input information, such as typing text or speaking live, and uses fingerspelling when there is no direct sign available [8]. MediaPipe creates smooth hand movement animations for each word recognized and shows them in real time on the web interface. Processing is done entirely in the browser, ensuring quick performance and accessibility without needing special equipment or cloud services.

V. RESULTS AND DISCUSSION

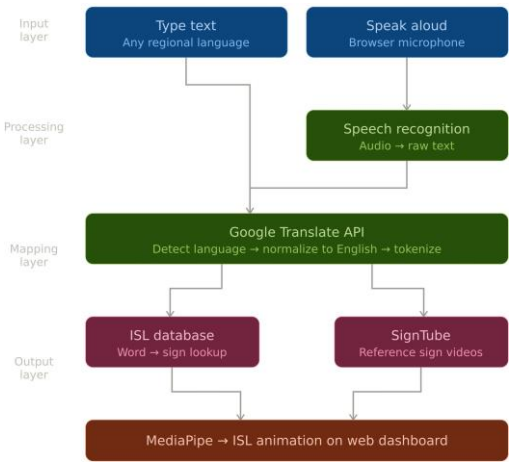

Fig. 5. System Workflow

Fig. 6 shows the input interface of the Gesture Talk web application, demonstrating how users can enter text or speechFig. 6. System Demonstration 1 – Input Interface

in both English and Devanagari script within a single browser-based dashboard [7]. The input panel displays a live text field where the user has typed “hindi, ?” in Hindi, confirming multilingual support without any manual language selection. A microphone button enables direct voice input, while the Translate button triggers the conversion pipeline. The interface also provides quick-access phrase suggestions in both Hindi and English at the bottom of the panel, allowing users to test common conversational inputs instantly [8]. The clean, minimal design ensures accessibility across all user groups, including those unfamiliar with technology, making the sys-tem practical for real-world assistive communication without requiring specialized hardware or software installations.

Fig. 7. System Demonstration 2 – Interpreter Panel

Fig. 7 shows the Interpreter panel of the Gesture Talk system, demonstrating the real time hand animation output generated for the converted ISL signs. This module utilizes MediaPipe [10] to render skeletal hand keypoint animations, displaying both the left and right hand movements separately within a structured four-quadrant canvas. The Sign and Spell toggle at the top allows users to switch between full word sign mode and letter-by-letter fingerspelling mode, providing flex-ibility for words that may not exist in the ISL database. Each hand is color coded with distinct LEFT and RIGHT labels, ensuring visual clarity during multi-hand signs that require simultaneous gestures. Navigation controls at the bottom allow users to step through sign sequences word by word using the previous and next buttons,

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 4, April 2026

DOI: 10.17148/IJARCCE.2026.154131

while the play button animates the

complete sign in a continuous fluid motion. This approach provides a lightweight yet visually accurate alternative to pre-recorded video demonstrations. This feature is particularly essential in educational institutions, public service counters, and healthcare environments, where clear and immediate sign language communication is critical for enabling inclusive interaction.

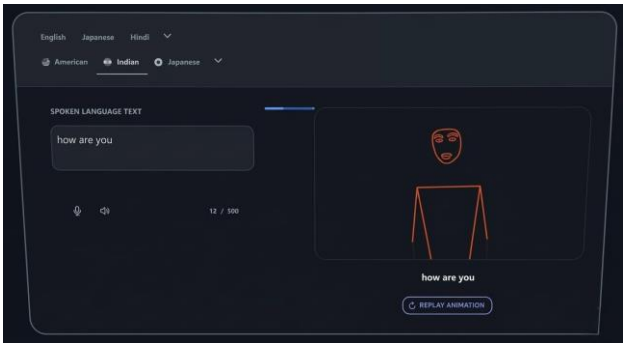

Fig. 8. System Demonstration 3 – Full Dashboard

Fig. 8 presents the complete Gesture Talk dashboard where the end-to-end conversion pipeline is demonstrated in a single unified interface. The user has typed “where are you” in the Input Source panel, and the system has processed the phrase in real time using Flash Engine V3.2 [11]. The ASL Sequence panel on the left displays the tokenized output, breaking the sentence word by word — with “YOU” highlighted as the currently active sign token at position one, and “WHERE” queued at position two. Simultaneously, the Interpreter panel on the right renders both left and right hand skeletal animations for the sign “YOU”, with a Sign Token description below confirming the hand shape as index finger extended pointing forward. The system maintains word-by-word navigation so users can step through each sign independently [12]. This recorded sequence facilitates efficient learning and communi-cation, helping users understand the structure of ISL signs and verify the accuracy of each gesture generated by the MediaPipe animation engine.

CONCLUSION

This paper presented an AI powered web application de-signed for real time conversion of text and speech into Indian Sign Language animations, making communication accessible for the deaf and hard-of-hearing community. By integrating multiple technologies within a unified browser-based pipeline, including Google Translate API for multilingual input normal-ization, MediaPipe for skeletal keypoint animation rendering, SignTube for reference sign video retrieval, and a structured ISL sign database for accurate word-to-sign mapping, the system effectively converts input across any regional Indian language into fluid and comprehensible ISL animations. The rule-based fingerspelling fallback further enhances reliability by ensuring that words without direct sign mappings are still communicated accurately. The lightweight, browser-based architecture enables low latency processing and real time

animation output without reliance on specialized hardware or complex software installations. Experimental observations indicate that the proposed system provides improved accessi-bility and practical usability compared to existing ASL or BSL focused approaches. The proposed system offers a scalable and inclusive solution for assistive communication and can be further extended with continuous sign language support, expanded ISL vocabulary, and gesture-based input modules to address evolving accessibility challenges.

ACKNOWLEDGMENT

We sincerely thank our guide, Dr. (Mrs.) Jayaprabha V. Terdale, for her valuable guidance, feedback, and continuous support during the course of this research. We are grateful to Prof. Archana P. Haral, Major Project Coordinator, for her encouragement and coordination throughout the project. We also thank Prof. Shilpali P. Bansu, Head of the Department of Artificial Intelligence and Data Science, for providing essential resources and support toward the development of this system. Our sincere gratitude goes to Dr. Sawata R. Deore, Principal of A. C. Patil College of Engineering, for fostering an environment of academic excellence and innovation. Finally, we thank all team members for their dedicated collaboration and contributions in building this assistive communication solution.

REFERENCES

[1] P. Sugandhi, P. Kumar, and S. Kaur, “Sign Language Generation System Based on Indian Sign Language Grammar,” *ACM Transactions on Asian and Low-Resource Language Information Processing*, vol. 19, no. 4, Article 54, 2020.

[2] J. Yi, Z. Zhang et al., “T-CPAD: A Transformer-Based Approach for Crowd Flow Prediction and Anomaly Detection,” *IEEE Access*, vol. 12, [Article in press], 2024.

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 4, April 2026

DOI: 10.17148/IJARCCE.2026.154131

[3] A. Harkude et al., "Audio or Text to Sign Language Converter," *Journal of Emerging Technologies and Innovative Research (JETIR)*, vol. 10, no. 10, 2023.

[4] H. Jeon, H. Kim et al., "Proactive Anomaly Surveillance System for CCTV Footage Analysis in Adverse Environmental Conditions," *Expert Systems with Applications*, vol. 254, p. 124391, 2024.

[5] R. Kulkarni et al., "Harnessing AI to Generate Indian Sign Language from Natural Speech," *International Journal of Advanced Computer Science and Applications (IJACSA)*, vol. 15, no. 4, 2024.

[6] K. Haritha, N. Sai Geethika et al., "Enhancing Public Safety with AI & ML-Based CCTV Surveillance," *International Journal for Modern Trends in Science and Technology*, vol. 11, no. 03, pp. 322–331, 2025.

[7] R. Damdoo and P. Kumar, "An Integrative Survey on Indian Sign Language Recognition and Translation," *IET Image Processing*, 2025. DOI: 10.1049/ipr2.70000

[8] S. M. Mahesh Kumar, "Conversion of Sign Language into Text," *International Journal of Applied Engineering Research*, vol. 13, no. 9, pp. 7154–7161, 2018.

[9] Z. An, Y. Ando, and R. Kawakami, "A Review of Deep Learning-Based Approaches to Sign Language Processing," *Advanced Robotics*, Taylor & Francis, 2024. DOI: 10.1080/01691864.2024.2442721

[10] J.-W. Kim et al., "Human Pose Estimation using MediaPipe Pose and Optimization Method based on a Humanoid Model," *Applied Sciences*, vol. 13, no. 4, p. 2700, 2023.

[11] F. Camastra, A. Ciaramella et al., "An Integrated Intelligent Surveillance System for Industrial Areas," in *Ital-IA 2024: 4th National Conference on Artificial Intelligence*, Naples, Italy, 2024.

[12] S. Yao, B. Rahimi Ardabili et al., "From Lab to Field: Real-World Eval-uation of an AI-Driven Smart Video Solution to Enhance Community Safety," arXiv preprint arXiv:2312.02078v3, 2025.